A SURVEY ON ENERGY HOLE IN CLUSTERING ROUTING PROTOCOL OF WIRELESS SENSOR NETWORKS

Oluwadara J. Odeyinka 1,2,3,4
Michael C. Ndinechi 1
Onyebuchi C. Nosiri 1
Nnaemeka C. Onuekwusi 1

ABSTRACT

Wireless Sensor Networks (WSNs) has become an active branch in the field of Communication Engineering. Owing to energy demand of wireless sensor nodes and the entire network structure, different routing protocol had been developed to address the constraint on sensor nodes battery. Clustering Routing Protocol (CRP) being the most effective approach used could suffer a major drawback as nodes close to the Sink depletes their energy first, thereby resulting into energy hole whereby data routed to the Sink through the affected nodes become interrupted. This paper takes a sample of wireless sensor networks on selected perimeters and the possibility of energy hole on the sampled networks using a single and different hierarchical level metrics simulated in Matlab environment. The result revealed that nodes far away from the Sink still retain a large joules (J) of Energy above the mean energy residue values even though the structured network had failed. Hence, possible solutions to alleviate the challenges posed due to energy hole in similar networks are recommended.

Contribution/ Originality: This study is one of very few studies which have investigated the existence of energy hole in WSNs by deploying sensor nodes on selected perimeters and modifying the existing energy models to present a logical analysis.

1. INTRODUCTION

There had been increased interest in the possible use of Wireless Sensor Networks (WSNs) in wide range of application, making WSNs an attractive research area. Wireless Sensor Networks, consist of variable number of static or mobile nodes deployed across an area of interest to obtain, process and transmit relevant data. The sensor nodes can be uniformly or randomly deployed depending on the area of application. A wireless sensor node typically consist of several parts powered by battery [1]. These includes: (a) Sensing unit; this is capable of sensing temperature, pressure, humidity, visual, acoustic, location and many more. (b) Micro-controller for processing the obtained data and (c) a radio transceiver for receiving the raw data and transmitting the processed data to the Sink through a radio frequency channel. Sink node being a resourceful node with un-restricted communication, computational capability and additional energy source acts as an interface between wireless sensor networks and resource management center [2].

Energy being a vital issue in WSNs attracts the need for a routing protocol to set up optimal paths between sensor nodes and the Sink. In transmitting data in WSNs, it is equally essential to ensure that a good proportion of data sent is received at the Sink. However, energy depletion of the nodes can interrupt communication, affect...
reliability of the network especially its sensed data \[5\] and in worst case, could cause network partitioning which in turn interrupt monitoring of the phenomenon of interest.

Based on network structure, routing protocol in WSNs can be divided into three categories: Flat routing protocol, Location based routing protocol and Clustering routing protocol. Clustering being the most popular energy efficient routing technique provides various advantages like longer life time, scalability and less delay; but it leads to hot spot problem \[4\]. Clustering routing protocol consist of four stages: Cluster formation, Cluster Head selection, Data aggregation and data communication. Cluster Head (CH) as the local coordinator or Sink of the cluster coordinates the activities among Sensor nodes, collection of information within the cluster, data aggregation and transmission of the aggregated data to the global Sink in a single hop or multi-hop scheme.

As a means to ensure energy conservation in WSNs, residual energy of nodes and proximity to the Sink are always taken into consideration in CH selection technique. These however make nodes closer to the Sink to deplete their energy thereby resulting into unsuccessful delivery of data to the Sink. Even though, one of the advantages of clustering routing protocol is energy-hole avoidance. Energy hole in WSNs is either that nodes cannot relate directly to the Sink because of limited radio range or directly sending the data will limit the energy and reduce the network lifetime \[5, 6\]. By sending data from the far nodes from the Sink to the near nodes which works like a funnel and all the gathered data to the lower side of the funnel will send towards the Sink which will cause energy consumption and after a while, the far nodes’ energy will remain and the network will be out of work and that is energy hole to this phenomenon \[7-9\].

This paper exemplified the possibility of energy-hole in sampled WSNs deployed along rectangular shaped perimeters and suggest possible solutions to mitigate the occurrence of energy hole on similar samples.

2. REVIEW OF RELATED STUDIES

Maik and Agnew \[10\] proposed the non-uniform node distribution strategy in WSNs. Due to the unbalanced energy depletion among nodes in the networks, more nodes are included to the affected areas. However, the authors only considered the energy consumption of sending data without taking into account the energy consumption of nodes in the receiving mode. Simulation results show that when the network lifetime is over, up to 90% of the total initial energy of the nodes is left unused if the nodes are normally distributed in the network. Since the nodes closest to the Sink tend to deplete their energy faster than other nodes.

He and Xu \[11\] proved that in order to achieve a maximum use of energy, the best position for a Sink in a circular sensor network is the center of the network. So, their traffic follow a many-to-one communication pattern. Nodes belonging to inner rings not only transmit data sensed by them but also forward data generated by outer rings. As a result, the energy depletion for these nodes is much faster than their counterparts farther away from the Sink, and so the energy holes are created.

Morteza and Mortaza \[12\] in their method to prevent overload and unbalanced energy consumption near the Sink as well as lack of energy hole, proposed the following duties.

- Clustering is done in equal in all layers, except the first.
- The best node is selected as head cluster by considering remaining energy and the nodes distance.
- In the near the Sink layer, the nodes transfer their own data and receive the data from cluster head of high layers to the Sink directly. The network was formed in a concentric circles.

In Lian, et al. \[13\] in order to solve the problem of energy hole, distribution of more nodes had been used near the Sink. In Alipour, et al. \[14\] in order to solve the problem of energy hole, two nodes with more energy were been placed in the middle of the structured network to primarily mitigate the challenges of energy hole. The nodes which are near the Sink will send the data directly to the Sink after clustering. The nodes which are far from the Sink will first send data to the two nodes at the middle of the network. Since the energy of these two nodes is high and both act as bond, they do not deplete their energy and network life time is enhanced. The near nodes, since they are located in radio range, will not consume extra energy. Thus, energy hole will be prevented in the network.

Odelyinka, et al. \[15\] developed a model of an energy-efficient wireless sensor networks suitable for perimeter monitoring and suggested a little or no modification for similar network. However, at the end of the life-time of the sampled networks, there exist a variation in mean residue energy of the deployed samples.

3. METHODOLOGY

The methodology adopted in this research work entails: Sensor node deployment on the experimental perimeters, Cluster formation and Cluster head selection, model for sensor nodes communication, hierarchical routing protocol and validation by simulation in Matlab. The simulation parameters are listed on Table 1.

3.1. Sensor Nodes Deployment

A total of 306 nodes were uniformly deployed along the perimeter of rectangular shaped experimental locations, with dimension (L by B) square meter for a 5000m by 50m and 2700m by 40m perimeters; a modification to the work of Odelyinka, et al. \[15\]. The arrangement of sensor nodes assumed a clockwise direction starting from the least co-ordinates such that nodes with id 1-50 and 250-306 are the closest to the Sink. The following assumptions are made.

1. All Sensor nodes have the same initial energy of 10J.
2. Sensor nodes per cluster in the deployment is uniform.
3. Sensor nodes transmit directly to their respective CHs within a particular cluster.
5. The Sink is located at \((-30, -10)\) of the experimental location and has the information about the identity (id), energy content and location of each node as displayed on Figures 1-6.

3.2. Parameters for Cluster Head Election

Based on the parameters used for CH election, clustering approaches can be categorized as deterministic, adaptive and random ones. In deterministic schemes, special inherent attributes of the sensor nodes are considered, such as the identifier (ID) and number of neighbors they have. In adaptive manners, CHs are elected from the deployed sensor node with higher weights, which includes residual energy, communication cost etc. In random modes, mainly used in secure clustering algorithms, CHs are elected randomly without regard to any other metrics like residual energy and communication cost (distance to the Sink). Considering the mode of sensor node deployment in the sampled network in this paper being uniformly deployed, adaptive approach is utilized in the selection of Cluster Heads.

3.3. Model for Sensor Nodes Communication

A simple model for the radio hardware energy dissipation were developed where the transmitter dissipates energy to run the radio electronics and the power amplifier, and the receiver dissipates energy to run the radio electronics. As displayed on Table 1, $E_{TX(elec)}$ and $E_{RX(elec)}$ were set to 50nJ/bit while $E_{amp}$ was set to 100pJ/bit/m$^2$ for the transmit amplifier to achieve an acceptable $E_b/N_0$ (energy per bit over noise spectral density) \cite{9}. The cost to transmit a message depends on the distance between the transmitter and receiver. Thus, to transmit a $n$-bit message a distance $r$, the radio expends energy:

$$E_{TX}(n, r) = n * E_{TX(elec)} + n * r^2 * E_{amp}$$

(1)

$E_{TX}(n, r)$ is the energy required to transmit a $n$-bit data through a distance $r$, $E_{TX(elec)}$ is the transmitter electronics energy, $E_{amp}$ is the free space transmit amplifier.

To receive this message, the radio expends energy:

$$E_{RX}(n) = n * E_{RX(elec)}$$

(2)

Where ‘r’ is the distance between the sources and destination. Nodes route their packet to the Sink through intermediate nodes. Thus nodes act as routers for other nodes in addition to their sensing functions.

3.4. Clustering Routing Protocol

The proposed routing protocol was based on hierarchical data transfer along any shortest path through the cluster heads to the Sink. Cluster head selection was based on the energy levels of the nodes and proximity to the Sink, and Sensor nodes are uniformly distributed over the entire perimeter. The Sink starts the process by asking nodes to form clusters. The CH election phase proceeds after the cluster formation phase. The selection of CH(s) within each cluster formed is carried out by electing a Sensor node that required least transmission energy to the Sink to be the CH for a particular transmission round. Due to draining activities being constraint on a Cluster Head during data aggregation and transmission phase, Cluster Head role is rotated among the sensor nodes of each cluster at every transmission round. A completely new estimation of energy was carried out at the beginning of every transmission round to elect a new CH for the cluster and thereby energy wastage is being reduced to its minimum and utilization of each Sensor node's energy is being maximized.

The Clustering Routing Protocol developed is summarized as follows:

1. Sensor node deployment and cluster formation phase.
2. Selection of cluster head in each cluster.
3. Data aggregation phase which involves the gathering of collected data by the cluster head from the sensor nodes within its cluster, the data aggregation phase was based on energy prediction technique using the first order radio energy model \cite{9}. This is described thus:
   a. The initial energy, $E_{in}$ of each node (RN or CH) was measured.
   b. The distance, $r_x$ between each regular node (RN) and the CH, where, $x = 1, 2, 3, \ldots$ was measured.
   c. Energy required by each node for transmission within the cluster was estimated using the formula:

$$E_{rn} = \epsilon_{amp} * n * r_x^2$$

(3)

Where $E_{rn}$ is the energy required by regular nodes for intra-cluster communication, $\epsilon_{amp}$ is the free space transmit amplifier and $r_x$ is the distance between regular node and CH.
d. The maximum energy after the subsequent transmission round for each node (RN) was estimated and selection of CH was done using the formula:

\[ E_{rn}(\text{max}) = E_{in} - (\varepsilon_{amp} \times n \times r_n^{-2}) \]

\[ E_{rn}(\text{max}) \] is maximum energy of regular node, \( E_{in} \) is the initial energy of nodes, \( \varepsilon_{amp} \) is the free space transmit amplifier, \( n \) is the data size in bit and \( r_n \) is the distance between regular node and CH, and the next CH selection will take place after the completion of the current round.

Data transmission phase which involves the transfer of all data from the nearest cluster head(s) to the Sink.

**Table 1. Simulation parameters.**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of sensor nodes</td>
<td>306</td>
</tr>
<tr>
<td>Experimental Location dimensions (L by B)</td>
<td>(300m by 50m) &amp; (2700m by 40m)</td>
</tr>
<tr>
<td>Sensor per Length (L)</td>
<td>150</td>
</tr>
<tr>
<td>Sensor per breath (B)</td>
<td>3</td>
</tr>
<tr>
<td>Initial sensor node energy</td>
<td>10</td>
</tr>
<tr>
<td>Transmit electronic energy (nJ)</td>
<td>50</td>
</tr>
<tr>
<td>Receive electronic energy (nJ)</td>
<td>50</td>
</tr>
<tr>
<td>Transmit power amplifier energy (pJ)</td>
<td>100</td>
</tr>
<tr>
<td>Packet size in (bits)</td>
<td>2000</td>
</tr>
<tr>
<td>Sink coordinates</td>
<td>((-30, -10))</td>
</tr>
<tr>
<td>Cluster Size</td>
<td>1, 6, 18</td>
</tr>
</tbody>
</table>

**3.5. Simulation Processes in Matlab Environment**

During the simulation processes, snap-shots were taken intermittently to survey the possible outcome and the choice of cluster head before data aggregation, after data aggregation and during multi-hop transition phases at different data transmission rounds. These processes occurred within micro seconds while the snap shots were been taken intermittently to carry out the survey. The emphasis is on the choice of CHs which are majorly those closer to the Sink as revealed from the various samples taken. The outcome of this selection processes of CHs on the overall energy usage in the networks are revealed in the sub-sections under section 4.0

![Figure 1. Transmission round 22 before aggregation.](image)

**Source:** Simulation Snap-shot in Matlab (R2013a) version.

In the transmission scene, node with identity 12 (id:12) was among the Sensor nodes selected as the cluster head with a residual energy of 8.4517J. Even though there are other sensor nodes with higher residual energies but the residual energy criteria of CH selection has been traded for proximity to the Sink making a sensor node closer to the Sink and possessing a fairly higher residual energy to be selected as one of the CH for the sampled intra-cluster transmission. During this process being a process before data aggregation, sensor nodes forward their data to the cluster head for aggregation.
During this phase, sensor node id:301 is selected as cluster head, with residual energy 9.9803J. The sensor node being a node with high residual energy in the series happened to be one among the sensor nodes closer to the Sink. The task laden on the particular node serving as one of the CHs during this phase resulted in a sharp drop in its energy from 9.9803J to 9.1405J as displayed during multi-hopping transition in Figure 3.

During this phase, sensor node id:15 with residual energy 9.8647J is selected as one of the CH for this particular phase. It would be recalled in the deployment scheme that nodes with lower identities such as nodes id:1,2,3,4,5,6,… and nodes with higher identities such as nodes id: 300, 301,302,…306 are closer to the Sink compared to nodes with middle range identity such as nodes id: 150,151,152… thereby enhancing the frequent selection of nodes with lower and higher identities as CH and secluding the middle range node from this responsibility.
Figure 4. Transmission round 578 before aggregation.

Source: Simulation Snap-shot in Matlab (R2013a) version.

Figure 4 illustrates a typical middle stage transmission compared to round 22 where the simulation process is still at early stage. Here, some of the sensor nodes are beginning to die as their energy content could no longer initiate transmission. However, the emphasis is in the choice of CH as most of these choices are centered on nodes closer to the Sink. The sampled screen shot of Figure 4 is for transmission round 578; a typical example of processes before data aggregation. Before aggregation, the communication is intra-cluster when sensor nodes send raw data to their respective Cluster heads. In Figure 4 node id: 305 with residual energy 6.5475J is selected as one of the CHs. The set of Cluster heads selected during this stage performs aggregation. Thereafter, the stage transit to after aggregation stage as typified in Figure 5 when CHs must have removed all redundant data.

Sensor node id:9 is selected as a CH with residual energy 6.5591J during one among many aggregation phases of round 578. However, some sensor nodes are dead which implied that during data Aggregation phase of this transmission round, data routed through the Dead Sensor nodes may be lost due to energy hole as it could be observed that zero packet size is being routed.
This is a sampled transition phase between multi-hopping and after multi-hopping of round 578, though the sampled transition phase is one among the numerous phases of the particular transmission round. The node with id:3 selected as one of the cluster head signified that priority is placed on nodes closer to the Sink in the choice of Cluster Head selection. Even though there exist a node with residual energy a bit higher than that of node id:3 and also among the closest to the Sink (node id: 304 with residual energy 6.5486J) and there could also be similar cases which are not displayed during this particular scene, this proved that there is a trade-off between residual energies of nodes and proximity to the Sink.

4. RESULTS

All the results were obtained by simulation in Matlab (R2013a) version at varied parameters of change in Cluster sizes (single, six and eighteen) and network sizes (3000m by 50m and 2700m by 40m). The residual energy metrics used for the survey in the research work is the energy contents of all Sensor nodes at the end of the networks lifetime. The result of simulation are displayed on Figures 7-12.

4.1. Residual Energy of Single Cluster Networks

Figure 7 is a residual energy of a typical single cluster deployment of a sampled perimeter of about 3000m by 50m. The displayed graph is the energy contents remaining of all nodes when transmission to the Sink could no longer be established. Even though transmission to the Sink is lost, it could be observed from the graph that about 40% of the sensor nodes ranging from id:75-225 retained their energies above the mean energy residue value of 0.3J.
Similar observation appears in Figure 8 relative to Figure 7. The network sample is equally a single cluster network on an approximate dimension: 2700m by 40m. Similarly, about 40% of sensor nodes retained energy above the mean energy residue value of 0.24J. Mostly, Sensor nodes with identities ranging from; 100-225 are within the category. The uneven distribution of loads in the network could be the resultant effect of having nodes with high energy content after the expiration of the network.

4.2. Residual Energy of 6 Cluster Hierarchical Networks

Figure 9 is a residual energy of a network sample of 6 clusters on the perimeter of a 3000m by 50m field. After the expiration of the network (when the Sink node assumed that all nodes is dead as a result of inability to receive data from the sensor nodes), the graph displayed revealed that about 50% of the nodes with identities ranging from 51 – 230 retained energies above the mean energy residue value of 5.2J while some of the Sensor nodes retained their energy at the initial value of 10J.
Figure 9. Energy Residue versus Sensor Nodes for 6 cluster deployment on a 3000m by 50m perimeter.
Source: Simulation Snap-shot in Matlab (R2013a) version.

Similarly, Figure 10 is a residual energy of a network sample of 6 clusters on a perimeter of approximately 2700m by 40m field. After the expiration of the network (when the Sink node assumed that all nodes is dead as a result of inability to receive data from the sensor nodes), the graph displayed revealed that about 50% of Sensor nodes with identities ranging from 55–230 retained energies above the mean energy residue value of 4.8J, while few of the nodes energy remain unused.

Figure 10. Energy Residue versus Sensor Nodes for 6 cluster deployment on a 2700m by 40m perimeter.
Source: Simulation Snap-shot in Matlab (R2013a) version.

4.3. Residual Energy of 18 Cluster Hierarchical Networks

Figure 11 represents the residual energy of a sampled 18 clustered hierarchical network on a perimeter of about 3000m by 50m. the residual energy being the energy content of node after the network failure indicate what joules of energy each node has. Few Sensor nodes between identities: 50 – 100 retained their energy above the mean value of 3J. While some Sensor nodes ranging from identities:100 -200 remained at the initial value of 10J.
Figure 11. Energy Residue versus Sensor Nodes for 18 cluster deployment on a 3000m by 50m perimeter.  
Source: Simulation Snap-shot in Matlab (R2013a) version.

Similarly, Figure 12 illustrates the residual energy of a sampled 18 clustered hierarchical network on a perimeter of about 2700m by 40m. The residual energy being the energy content of node after the network failure indicate what joules of energy each node has. Few Sensor nodes with identities ranging between; 70 – 90 and 225-240 has their residue energy above the mean energy residue value of 2.6J. while some nodes ranging from id:100 - id:200 retained their energies at the initial value of 10J. These, in a similar situation to Figure 11 implied that the energy of the affected Sensor nodes were never or sparingly utilized.

Figure 12. Energy Residue versus Sensor Nodes for 18 cluster deployment on a 2700m by 40m perimeter.
Source: Simulation Snap-shot in Matlab (R2013a) version.

5. CONCLUSION

Observations and inferences deduced from this survey revealed the possibility of energy depletion hole in clustering routing protocol of WSNs. Owing to the fact that data routing in clustering protocol is hierarchical such that data are directed towards the global Sink in a multi-hop style. This results in unbalanced distribution of load as nodes far away from the Sink are less tasked while sensor nodes closer to the Sink are laden with data routing activities, thereby leading to the possibilities of such nodes draining their energies. The outcome of the survey in this paper carried out on selected perimeters for single cluster size and multi-cluster sizes with a Static Sink outside the area enclosed by the perimeter revealed that after the network could no longer established data transmission to the Sink; assuming that all sensor nodes is dead, the residual energy displayed showed in all cases that while some sensor nodes were confirmed dead (energy is 0 Joules), some nodes far away from the Sink retained energies above the mean residue energy values while others maintained the initial value of 10 Joules.
6. RECOMMENDATIONS

The increasing challenge of energy hole in Wireless Sensor Networks could pose a constraint in the quality of data delivery as vital information may not reach the Sink. Thus, the following are possible recommendations for optimal energy usage and to mitigate the effect of energy hole on similar networks deployed for perimeter surveillance:

(i) Non uniform cluster size (Cluster size of nodes closer to the Sink should be larger than far away nodes).
(ii) A center based Sink.
(iii) Sink mobility where applicable.
(iv) Adoption of intermediate or repeater Sink.

Funding: This study received no specific financial support.
Competing Interests: The authors declare that they have no competing interests.
Acknowledgement: All authors contributed equally to the conception and design of the study.

REFERENCES