An Energy-Efficient Re-Clustering and Data Reporting Algorithm for Energy Constrained Wireless Sensor Network

Hemappa B
Research Scholar,
Department of E&E,
SMIT, Sikkim, India

Manjaih D.H.
Department of Computer Science,
Mangalore University, India

Rabindranath Bera,
Department of E & C
Engg, SMIT,
Sikkim, India

Shylaja B.T.,
Consultant,
Parivartan Software Solutions,
Hubli, India

ABSTRACT
Wireless sensor networks consist of tiny sensors powered by low power battery and in most of the applications sensors are deployed to monitor the unattended areas. As sensor nodes are energy constrained in the network, energy efficiency is an important factor and should be considered during the protocol design phase. In this article we propose a robust, power-aware in-network clustering protocol to increase the network lifetime, which is based on parameters like residual energy, distance to the base station and proximity of the base station. The proposed cluster head selection protocol offers an efficient way of balancing the energy among member nodes of the cluster with considerably lesser energy consumption. Proposed novel algorithm is experimented in Matlab simulator and results shows significant improvements when compared to LEACH algorithm. The performance is evaluated by considering the impact of energy balancing among the nodes and network lifetime efficiency in WSN.

General Terms
Cluster Head Selection, Re-clustering, Wireless sensor network.

Keywords
Clustering Algorithm, Energy efficiency, In-network processing, Network lifetime.

1. INTRODUCTION
In recent years WSN has emerged as a strong monitoring and sensing paradigm, especially in fields like forest fire detection, target tracking in the battlefield, habitat monitoring and other similar fields of interest. A WSN is a one hop or multihop network of nodes, each with a short-range radio, limited sensing and on-board processing capability[8]. Sensor nodes are enabled by small batteries and have a cost conscious energy budget as they cannot be recharged after the deployment [11]. Therefore, it is obvious that specialized energy-aware routing and data gathering protocols offering high scalability should be applied in order that network lifetime is preserved acceptably high in such environments [14]. WSNs are usually deployed in large scale to monitor static or dynamic events. Static events (temperature and humidity) are easy to capture compared to dynamic events, as dynamic events come and go quickly. In an event driven sensor network, each and every sensor starts to communicate and engage in data transmission and it leads to network congestion and data collisions. If the base station is far away from the nodes, direct communication will require a large amount of transmit power from each node [13]. To avoid such scenarios, group of sensors form a cluster and sensor nodes communicate to the cluster head and the cluster head communicates the data to the sink. Clustering helps in managing the energy levels efficiently. The in-network data processing within the cluster improves network performance by reducing the amount of data to be delivered [12] and the number of hops from sensors to the sink. In such networks, however, the more energy is consumed by the cluster head (CH) and one hop neighbor nodes of CH. This is because more computing and communication loads are assigned to the CHs. This non-uniformity of energy consumption among nodes results in some nodes dying earlier than others [6]. This imbalance of energy among the nodes has motivated us to design the proposed re-clustering algorithm which concentrates mainly on efficiently balancing the energy among sensor nodes in a cluster.

1.1 Clustering in wireless sensor networks
The clustered WSNs are mainly focused on data gathering applications which involves in the continuous delivery of sensory data over multihop routes. This creates network congestion, especially at locations close to a sink [3]. Many researchers have proposed novel protocols for cluster formation and CH selection [10]. In-network data fusion and clustering have been proven to be effective techniques in reducing energy consumption[2] in WSNs. Data communication in clustered wireless sensor network has issues such as cluster formation, cluster based sensor organization, network management[11]. Managing the network time effectively during clustering is also an issue of major concern. The extra energy and time are consumed to reform clusters at the setup phase of every round. This side effect is worse as the number of clusters increases[6]. The clustering time, CH selection period. CH tenure adds up to the sensor network time. Managing the time effectively includes allotting the network time effectively among all the activities of a cluster. The sensor nodes listen, sense, communicate, transmit and receive during its tenure. More energy is dissipated during transition, listening and transmitting. Negligible amount of energy is dissipated during computation and receiving. CH selection in clustering is of high focus which enables balanced and effective energy utilization thereby prolonging the network lifetime.

In a sensor node’s tenure as CH, it performs multiple tasks like receiving sensed data, updating node’s neighborhood tables, data fusion, relaying received data to the sink, etc... In this process, CH node is drained of more energy and may lead to fatal results such as the early death of a node. The death of a node creates holes in the network. This will result in nodes tracing another route which may be lengthier compared to the earlier one. CH rotation can avoid the premature death of a sensor node and thus re-clustering is an emerging technology which emphasizes on energy preserving and energy balancing among the clusters. The nodes nearer to the sink can be
avoided to be selected as CH as these nodes are termed as “hot- spots” and retaining their energy levels is so crucial. Also the border nodes which are often engaged in sensing must also be spared as they are depleted of more energy while sensing. In the proposed algorithm, role of the CH node is rotated based on the residual energy and preference is given to those nodes which are having surplus energy to reduce the imbalance among the nodes in the cluster. And one hop nodes around the sink are not considered for CH role to balance the energy level of the network.

The proposed algorithm explains how some of the drawbacks of the earlier CH selection algorithms are addressed. It can achieve load balancing in cluster head selection. Simulation results show lower energy consumption as CH selection is not triggered unnecessarily for every constant period of time there by achieving prolonged network lifetime.

2. RELATED WORK
There has been extensive work undertaken in the field of re-clustering and CH selection. The LEACH is a dynamic hierarchical clustering method to balance the energy in WSN. The operation of LEACH is broken up into rounds where each round begins with a set-up phase. When the clusters are organized, followed by a steady-state phase, then data transfer to the base station occurs[13]. The main idea of LEACH is to reform clusters once every period of time, called a round, in order to rotate the role of the CH among members in a cluster. Many researchers have been doing studies in the past many years exploring the LEACH protocol to improve performance. However, there has been no work to address energy consumption during the cluster-reforming process [6]. LEACH with Deterministic cluster-head selection is one of LEACH’s variations, in which cluster heads are elected by reducing the threshold.

An adaptive, energy-efficient clustering protocol (AEEC) aimed to re-select the cluster head for each round to balance the energy[15] because CH dissipates more energy compared to its neighbours. The AEEC divide its operation into three phases: cluster formation, routing tree construction, and data sending for each round. The cluster formation involves in broadcasting multiple control messages such as CH_msg, join_msg and re_clustering_msg which leads to more energy consumption for each round. Another overhead of this protocol is construction of the routing tree for each round of clustering. In AEEC, re-clustering is triggered based on the round timer and energy threshold. When a new round starts, cluster head named ‘A’ checks its energy consumption. If it is equal to or greater than threshold energy, ‘A’ sends a message to its members and triggers re-clustering process. But in the event driven sensor network, more energy is consumed, when an event is triggered, as CH involves in receiving the data from event affected neighbours and aggregate the data before sending it to the base station. There are possibilities where CH energy consumption reaches the threshold energy level before it reaches the round timer which causes energy imbalance in the network. In an event driven sensor network application, when an event occurs, all the nodes which have detected the event start sending an event data to CH which leads to data redundancy and more energy consumption.

In [12], authors have proposed to send the sensed data to neighbour node for aggregation when the CH is in multi hop. So that communication distance reduces along with data compression which leads to increase the life time of sensor as well as whole network. The authors of [4], proposed an algorithm to select the appropriate node with more residual energy among the event detector nodes to send the event data to CH. The selection of the node to communicate the event data is based on residual energy of the cluster member and the cluster head and setting a timing interval to wait. This algorithm may work well when all sensors nodes having different residual energy, but if any two nodes are having same highest residual energy then more than one node starts sending the event data which will lead to collision at the CH.

3. SYSTEM MODEL

3.1 Challenges in Clustering based data collection
- **Redundancy:** Many sensors detect the same event, each tends to transmit the data to their respective CH. This leads to too many data packets with similar content sent to CH unnecessarily.
- **CH selection:** While selecting CH node, care has to be taken that the node is having least hops from the cluster nodes. It should possess more residual energy from the rest. CH role should be rotated so as to establish uniform load distribution [9]. Least data correlation should also be taken care.
- **Energy conservation:** As sensors are battery enabled devices, in multi-hop communication, the large quantity of data transfer may quickly drain the limited resource which results in topology change or disconnect the network. In single hop network, energy easily gets drained off in hot spot nodes [5], so hot spot nodes energy consumption must be minimal. Therefore these nodes should not be used for sensing but just for relaying.

3.2 Energy Efficient Clustering and Routing
In this section, proposed approach is discussed for managing the sensor network with a main objective of load balancing and extending the life time of the sensors in a particular cluster. In LEACH and AEEC, extra energy and time are consumed to reform clusters at the setup phase of every round. This process gives rise to many complications which are dealt in the following section. In the proposed algorithm, the focus is mainly on residual energy and constant threshold energy $E_t$. We consider a simple wireless sensor network with following assumptions and characteristics:
- Sensor nodes are organized into clusters including a cluster head and some member nodes as in Fig 1.
- The clustering is performed frequently to distribute the load overall the network.
- The nodes are homogenous and the network is reformed at the end of each iteration.
- Every member node can communicate and broadcast using a communication radius $r$.
- The CH selection computation is performed in all nodes(decentralised and adaptive).
- Every node is aware of distance to its immediate neighbours and the base station.
The main objective is to optimize the time required to select the CH by using threshold energy for the clustering task instead of waiting time. The clustering mechanism which is an effective method to deal with the hot spots problem and it can prevent the premature creation of energy holes in wireless sensor networks. The CH selection process is carried out in-network and each sensor nodes in the cluster involves in cluster head selection process. Every node maintains a table to store neighbors information. The table includes: Nodeld, Residual Energy, and Proximity of base station.

Nodeld: Every node is given an identification number. Residual Energy: Amount of energy left with the node at that particular time. For instance consider the Fig 2., which is in accordance with Table 1., which is the neighborhood table of node 3. As shown in Table 1., node 6, 18 have same residual energy and both may broadcast CHmsg to its neighbors. This induces CHmsg collisions. To avoid collision one more parameter, proximity of bases station is introduced. The proximity of base station: Based on the distance between base station and the node, the proximity of Base Station is calculated. It can be either far or nearer. This parameter is considered when the residual energy of more than one node is same. The node which is far away from the Base Station is preferred for CH, as it is not a hot spot. Considering proximity is necessary even though all the nodes are at a known distance from the base station. In this case, node 18 would be selected as its proximity is far.

In the initial stage of clustering, every node broadcasts the BeaconMsg using the communication radius \( r \), which includes information of all the above table content. On receiving, every other node updates its neighbor table. Now let \( E_{t} \) be the minimum threshold energy, which is the amount of energy to be spent by the CH during its tenure. The \( E_{t,ch} \) is the initial energy of the CH. \( E_{t,ch} \) will be the remaining residual energy left with CH after its tenure which is calculated as in equation 1.

\[
E_{t,ch} = E_{t,ch} - E_{t} \quad - - - - - - (1)
\]

When a CH depletes \( E_{t} \) amount of energy, then its current residual energy reaches \( E_{t,ch} \) then the cluster head selection will be triggered. This indicates that the current CH has completed its tenure and new CH selection process has to be triggered.

<table>
<thead>
<tr>
<th>Table1. Neighborhood table of node 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodeld</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>22</td>
</tr>
</tbody>
</table>

Compelling all the sensor nodes to wait for a particular period for the CH selection causes every node to negotiate with certain drawbacks which has to be resolved.

3.2.2 Proposed Re-clustering Algorithm

The main objective is to optimize the time required to select the CH by using threshold energy for the clustering task...
This - at 1 - sin - illness leads the multi hop path to take alternative path -

When new CH selection process triggers then each sensor nodes in the group start the CH selection process using their neighborhood table, which includes residual energy and node's proximity to base station. Transmission cost is more expensive than computation cost[1]. Therefore CH selection computation is carried out in every node rather than computing in one central node. This would have required all the member nodes to transmit their information to central node and vice versa. The procedure identifies the node having highest residual energy by scanning the recently updated neighborhood table and declares it as CH for the next cycle. Initially all nodes possess same energy levels. This will lead to CH message collisions. And also in later stages the possibility of more than one node having same highest residual energy may arise. In such cases, along with residual energy, proximity of the base station is also considered as shown in the Fig 2. In each group of nodes, the node which is far away from the base station and within the c neighborhood - includes residual energy and node's proximity to base station - is used to select CH. It need not wait for any period of time. All remaining nodes pick the same shared medium to transmit at the same time - multiple nodes pick the same shared medium to transmit at the same time - or two or more nodes having same residual energy - two or more nodes having same residual energy - or two or more nodes having same energy which is equal to the threshold energy E_r - two or more nodes having same energy which is equal to the threshold energy E_r - are declared as CH contenders. As depicted in the equation (2). Therefore for an undirected graph G(V,E), where i= 1 to η, η is number of nodes in that cluster.

\[ d(v_{self}, v_{sink}) > d(v_i, v_{sink}) \] - equation (2)

The nodes which are nearer to base station are the vital nodes on the multi hop path. If these nodes are selected as CH then these nodes get depleted of more energy and sometimes they may die early which leads to creating a hole in the multi hop path. This leads the multi hop path to take alternative path which may be long and time consuming. Therefore it is an efficient way to select nodes which are at a far distance from the base station. And if that node dies, the multi hop path will not be affected badly. The pseudo code of proposed re-clustering is shown in Fig. 3.

Cluster_head_selection():

1. Initialize the state of node as contender for CH except the one hop nodes to sink - one hop nodes to sink - the node having lowest residual energy.
2. broadcast Beacon_Msg using radius r
3. receive Beacon_Msg , update the node’s Neighbourhood table NHT[]
   it includes
   Nt: Node identification number
   R: Residual Energy
   d: distance between node i,j.
   P: proximity of base station (far or near)
4. vi: randomly selected node
   E_ch ← GetCurResidualEnergy(vi)
5. if(E_ch > E_i)
   E_ch = E_ch - E_i
   else
   go to step 4.
6. endif
7. while(true)
8. while(Er_ch >= GetCurResidualEnergy (vi))
   //get the count of highest residual energy nodes
   v ← GetHighestResEngNodes()
9. if(v.size = 1 & & v.id = vi & & v.state = contender)
   v.state = CH
   else
   v.state = normal
10. endif
11. if(v.size > 1)
12. for i=1 to v.size
13. if (vi.id = v.self.id) & & (v.self.proximity = far) & &
   d(v.self, vi) > d(vi, v.self)
   v.self.state = CH
   else
   v.self.state = normal
14. endif
15. endfor
16. endif
17. if(v.self.state = CH)
   E_ch = V_vf
   E_ch ← GetCurResidualEnergy(v_i)
18. if(E_ch > E_i)
   E_ch = E_ch - E_i
   broadcast CH msg
   break
19. endif
20. endif
21. endwhile
22. endwhile

The proposed system uses shortest path routing algorithm and stores the available routing table. It uses the two best shortest paths to route the control message and payload transmission. If the first shortest route is not feasible due to active or sleeps duty cycle then it selects the second best shortest path until the first best route is available [7]. In an event driven WSN, if multiple nodes pick the same shared medium to transmit at the same time then it leads to network congestion, collision at receiver end therefore the previously proposed algorithm [7] is used for sender node selection. As depicted the steps in pseudo code for cluster head selection mainly uses the neighborhood table to select the highest residual energy node. If a node possesses highest residual energy it declares itself as CH. It need not wait for any period of time. All remaining node having lowest residual energy can join the newly elected CH. There is no waiting time and the computation time for CH selection is very negligible. Thus if an event occurs during the CH selection process then the possibility of sensed data becoming stale is very less.

A sensor node continues as a CH node till it depletes the energy which is equal to the threshold energy E_t. After that node which has the next highest residual energy will be selected. The current CH node may be re-selected if it still has highest residual energy among all the member nodes. This procedure enables the member nodes of the cluster to emphasize the balanced energy levels among them. A CH node has to deplete E_t amount of energy during its tenure and
after that the next CH selection process will be kick-started. So if no event occurs during a CH node’s tenure, there is no need to trigger new CH selection until current CH depletes $E_i$ amount of energy. The residual energy of sensor nodes and their proximity to the base station are the parameters considered for the CH selection. Thus number of neighbor nodes will not affect the CH selection.

4. The base station is fixed in the middle of the top at (50,98) location.
5. Nodes communicate with CH via single-hop or multi-hop
6. The wireless transmitter power is configurable and parameters are shown in Table 2.

<table>
<thead>
<tr>
<th>Table 2: Simulation parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symbols</strong></td>
</tr>
<tr>
<td>$E_{tx}$</td>
</tr>
<tr>
<td>$E_{rcv}$</td>
</tr>
<tr>
<td>$\Delta$</td>
</tr>
<tr>
<td>$\delta_{mp}$</td>
</tr>
<tr>
<td>$E_{DA}$</td>
</tr>
<tr>
<td>$K_{bits}$</td>
</tr>
<tr>
<td>$E_i$</td>
</tr>
<tr>
<td>MxN</td>
</tr>
<tr>
<td>maxNodes</td>
</tr>
<tr>
<td>Deployment</td>
</tr>
<tr>
<td>Sink node location</td>
</tr>
</tbody>
</table>

4.2 Wireless communication model
The proposed algorithm uses the wireless communication model for data transmission. Each sensor consumes energy during various functions such as sensing, receiving, transmitting and data processing (aggregation). To calculate the energy consumption by the nodes to transmit ‘$K_{bits}$’ of data packet to distance $d_{ij}$, the algorithm uses the simple following energy consumption model [16] as in equation 3 and 4 to evaluate the energy consumption for transmitting and receiving respectively.

$$E_{tx}(i,j) = K_{bits} (E_{dpb} + \delta d_{ij}^\sigma)$$

$$E_{rcv}(i,j) = K_{bits} * E_{dpb}$$

Where $E_{dpb}$ is the energy consumed by a sending node to send each bit of data, or a receiving node to receive each bit of data. $\delta$ is the energy consumption in the amplification circuit for forwarding each bit of data. $d_{ij}$ is the distance between node $i$ and node $j$. $K_{bits}$ is the data amount transmitted from node $i$ to $j$. Generally, $\sigma$ is considered to be two for small distances and four for large distances. Here let $\sigma = 2$ in case of intra-cluster and $\sigma = 4$ for inter-cluster communication. The energy for receiving $E_{rcv}$, is almost independent of the
distance between the transmitter and receiver and depends on the electronic parts of the receiver.

4.3 Energy consumption
The total energy consumption curve of sensor nodes is shown in Fig. 5. The proposed algorithm reduces the energy consumption of cluster heads as it performs the in-network computation during re-clustering which will reduce the communication of required information to central node for CH selection. In proposed algorithm re-clustering triggers only after dissipation of threshold energy $E_M$ which reduces the risk of the frequent cluster formation when there are no events. This balances the energy consumption of the whole network, extends the lifetime of cluster heads which may die earlier and optimizes the performance of the network thereby reduces the total energy consumption of the effective lifecycle. From the analysis of Fig. 5, it shows that in the whole lifetime of the network, the energy consumption of the proposed algorithm is much lower than that of LEACH protocol at the same round of simulation. For 1000 rounds, for the sum of energy 501, the proposed algorithm has saved around 9% of energy when compared to LEACH algorithm. As the number of rounds increase the amount of energy saved will also increase. These results are consistent with the different design parameters. The results are energy efficient if the value of the parameters is increased, like number of rounds and energy of each sensor. If the number of nodes and their energies are increased then more saving of energy can be achieved.

4.4 Life time of the network
Network lifetime is perhaps the most important metric for the evaluation of sensor networks. The commonly used network lifetime definition in literature is the network lifetime ends as soon as the first node fails in the network which results in the first topology change after the deployment. The network will be stable till the first node fails to perform its operations or dies. As soon as some nodes begin to die then network operation may become unstable and there will be an impact on data transfer. Therefore, the longer the stable period is, the better the performance of the network. In LEACH Protocol, cluster heads are responsible for data aggregation and communicating with the base station. Randomly distributing the nodes and randomly selecting the cluster heads causes some cluster heads die earlier because of the low energy which results in creating the hole in the network. The hole in the network may isolate some of the nodes from the network.

In the proposed algorithm, energy is efficiently balanced among all nodes such that the chance of node dying earlier is effectively controlled. By analyzing the Fig. 6, for 1000 rounds the first node dies at 653rd round in LEACH algorithm. But in the proposed algorithm even at 1000th round no node has died. The first node has died at 1224th round. This drastic difference clearly shows the network longevity and stability of the proposed algorithm.

5. CONCLUSION
Many clustering algorithms including LEACH start the cluster head selection based on threshold time interval and random selection of cluster head. Sometimes these cluster heads may have less energy which leads to heavy energy burden and they will soon die. One hop nodes around the sink node are also considered as normal nodes which results in creation of holes in network. For these issues, the proposed Energy-efficient re-clustering algorithm aims to balance energy consumption of the whole network and extend the network lifetime by balanced and effective use of cluster heads energy. The new proposed algorithm is simulated by MATLAB platform, the simulation results clearly shows that the energy efficiency and the lifetime of network fare better than that of LEACH Protocol.

6. REFERENCES


