Abstract—Wireless Sensor Network (WSN) is a multi-hop network system in which sensor nodes are deployed in monitoring area to sense some of environmental parameters. WSNs have number of applications in daily life such as environment monitoring, building monitoring, smart phone, smart farming, habitat, and asset management. WSNs have features such as limited resources, multi-hop routing and ad hoc network topology. Clustering is introduced in WSN because it has proven to be an effective technique to achieve scalability, self-organization behaviour and energy saving for large WSNs. To improve scalability and network lifetime this paper proposes new clustering technique based on cluster first approach which uses at the most two-hop intra-cluster communication, which reduces number of messages from member node (Non CH node) to Base Station (BS). This helps to prolong network lifetime.

Key words—Wireless Sensor Networks, clustering, energy efficiency, network lifetime

I. INTRODUCTION

In clustering whole sensor network is partitioned into group of sensor nodes called as cluster with high energy node inside the cluster acting as the cluster head. Each cluster contains Cluster Head (CH) and its members [7]. CHs are nodes that consume more energy than cluster members when they involve in aggregating, processing and routing data. CHs collect data from all the sensors [3].

Figure 1. Clustering in wireless sensor network

Generally there is a single centrally located server called Base Station that collects all the information from the various nodes, examines it according to the needs of the system. These nodes have a disadvantage that they cannot be charged once they are deployed [1]. Thus the amount of energy per sensor node is limited. This situation will lead to drain limited energy from the network. Clustering technique will address these issues [2].

Use of clustering in a hierarchical WSN facilitates efficient utilization of limited energy of sensor nodes and hence extends network lifetime. Although sensor nodes in clusters transmit messages over a short distance such that more energy is drained from CHs due to message transmission over long distances (CHs to the Base Station) compared to normal sensor nodes (member nodes) in the cluster.

Periodic re-selection of CHs within clusters based on their residual energy is carried out to balance the power consumption of each cluster. In addition, clustering increases the efficiency of data transmission by reducing number of sensors that are attempting to transmit data in the WSN, aggregating data at CHs via intra-cluster communication. Two-hop clustering [2] gives more energy efficient results than the 1 hop clustering, so in our proposed model we are allowing two hop clustering.

There are two approaches used in this process, the leader first and the cluster first approach. In the leader first approach the cluster head is selected first and then cluster is formed. In the cluster first approach the cluster is formed first and then the cluster head is selected [1].

In the cluster first approach cluster heads are first selected based on certain parameters, and then they agree on how to assign remaining nodes to different clusters. In Leader-First approaches (e.g. DDAR [4], HEED [6], LEACH [5, 6]), cluster-heads are first selected based on certain parameters, and then they agree on how to assign remaining nodes to different clusters. Based on the order in which cluster formation and leader election are performed, there are two categories of clustering protocols: Leader-First (LF) approaches and Cluster-First (CF) approaches.

II. RELATED WORK

In DDAR [4] protocol, the distance and energy of the node is used as a factor to select CH. Protocol working undergo three phases namely network setup,
routing path construction, and schedule creation. Results shows that DDAR protocol is greater energy efficient as MTE, LEACH, and LEACH-C protocols.

Two Hop Clustering (THC) [2] protocol employs 2-hop clustering method, which reduces the amount of power required by the member node to communicate with the CH nodes. All the 2-hop away member nodes forward their data to the CH nodes via 1-hop nodes; which significantly reduces the amount of energy consumption.

Cluster of cluster heads approach [1] is explained by using three phases namely initialization, setup and steady. Setup phase contains Cluster formation and CH Selection phases. In this approach Cluster First approach is used which increase the lifetime of the network.

III. NETWORK MODEL

A. Sensor Node (SNi)

A sensor node has a number of characteristics such as its unique identity, type, transmission range, state, and energy level. A sensor node for our model is defined as follows:

\[ SNi = \{ID, ST, R, St, E, Mo\} \]

Identifier: (ID) = i, i \in N where N= \{1, 2, .. , N\}

Sensor’s type (ST)\in \{temperature, humidity, sound, pressure, vibration, light, etc...\}

Transmission Range (R) of a node

St represents the state of sensor node: (St)\in \{Active, Inactive\}

Erepresents the battery level at \(tj\) second of lifetime of the node \[E(SNi (tj))\] i.e. remaining capacity of node at time \(tj\)

Mo is the Type of Mobility: (Mo)\in \{Stationary, Mobile\}

B. Cluster (C):

Cluster is defined as follows:

\[ C = \{CID, CHID, N, CC\} \]

Cluster Identifier is represented by (CID) Cluster Head Identifier is represented by (CHID)

Number of Sensor Nodes (N) in a Cluster \in \{fixed, variable\}

Cluster Communication (CC)\in \{Single-hop, Two-hop\}

C. Region (R)

Region is defined as follows:

\[ R = \{RID, CID, NR\} \]

Region Identifier is represented by (RID) Cluster Identifier is represented by (CID)

Number of Sensor Nodes in each Region (NR) is given by,

\[ NR = \sum_{k=1}^{N} N (CID) \]

D. Average Energy Consumption

Let \(E_i\) = the initial energy level of a node, \(E_f\) = the final energy level of a node and 

\[ N = \text{number of nodes} \]

\[ E_a = \frac{\sum_{k=1}^{n} (E_{ik} - E_{fk})}{N} \]

E. Network Lifetime:

This is the time interval from the start of operation of the sensor network until the death of the last alive node.

F. Total Communication cost:

\[ TCC = \sum_{k=1}^{N} \text{no. of messages sent by node } k \]

IV. PROPOSED WORK

A. Assumptions:

In our proposed work following are assumptions for achieving efficient data routing to improve network lifetime. The assumptions are given as follows:

- The network area is fixed
- All sensor nodes are homogeneous in the network
- All sensor nodes and BS are stationary
- BS knows location of all nodes
- The intra-cluster communication is two-hop
- Sensors can compute the approximate distance to other sensors, based on received signal strength and transmitting power
- Sensor nodes and BS are stationary after deployment
- Communication is symmetric between any two sensors, meaning that the sensor node SNi can listen to the sensor node SNj, SNj can listen to SNi

B. Problem Definition:

To design a clustering approach which will increase the energy saving and improve the network lifetime of WSN. The proposed method uses two-hop connectivity for cluster formation, which will help to reduce number of CH than the 1-hop clustering. This helps to increase energy saving in WSN. It also reduces energy requirement for sensor nodes within a cluster.

C. Proposed Solution:

One of the major constraints of WSN is energy. In this research work, this problem is kept in mind and a solution for limited energy source has been proposed. To maintain high energy efficiency, some cluster-based algorithms utilize clustering methods that adjust the size of the cluster by itself or sometimes reform the clusters.
This global decision is taken by the network entities itself through local interactions among themselves. Hence sensor network is forced to behave as self-organized.

The proposed solution to group the sensor network into clusters which uses Cluster First approach consists of following four steps:

Step 1. Splits whole network:
Split network area into fixed number of rectangular cells. Statically we divide sensor area into rectangular cells.

Step 2. Formation of Cluster:
One cluster is formed within each rectangular cell, in which we are allowing at the most two-hop connectivity which cover more elements as compared to single hop clustering.

Step 3. CH Selection and Connectivity with member nodes:
CH selection is based on energy of node and node degree. In this step CH selection is taken place according to highest energy and advertisement of that CH to remaining nodes in particular region. After certain time new CH selection for next round taken place when its energy goes up to threshold value. Here single hop member is formed within the transmission range of the cluster node. We use shortest distance to join single hop nodes and next hop (i.e. two-hop).

Step 4. Actual Data Transmission: In this phase, the CHs create a time slots, gather information from the member nodes and finally transmit the aggregated data to the base station. CHs creates time slot for each node telling when it can transmit the data. This helps in data collision from the different nodes. The CH nodes collect the sensed data from all the member nodes during their time slot. The node which is 2-hop away sends its data to the 1-hop neighbour. The 1-hop neighbour adds its data to the received one and finally transmits it to the CH. the CH nodes aggregate the data received from the member nodes. This reduces the number of transmissions to the base station and also energy consumption. The CH nodes send their data to the base station using multi-hop communication. The far away CHs transmit their data to the nearest CHs in the direction of base station. These CHs add their data and transmit the same to the next level CHs.

V. EVALUATION PARAMETERS
Simulation is performed with network area of size 250m x 250m in which 100 sensor nodes are deployed. BS location is fixed. In cluster, sensor nodes communicate with CH and CH with the BS. We define parameters and their values at the time of initialization. The characteristics of simulation parameter are summarized into table 1 as follows:
TABLE 1. SIMULATION PARAMETERS

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Area</td>
<td>250 m x 250 m</td>
</tr>
<tr>
<td>Number of Sensor nodes</td>
<td>100</td>
</tr>
<tr>
<td>Base station location</td>
<td>Fixed (X, Y)</td>
</tr>
<tr>
<td>Initial energy of node</td>
<td>10 J</td>
</tr>
<tr>
<td>Transmission Range</td>
<td>50 m</td>
</tr>
<tr>
<td>Threshold energy</td>
<td>0.5 J</td>
</tr>
<tr>
<td>Packet size</td>
<td>4000 bits</td>
</tr>
</tbody>
</table>

VI. EXPECTED PERFORMANCE BEHAVIOUR

In this section, we have done theoretical calculation for residual energy and based on that performance of our proposed clustering technique is done. To calculate behaviour we assume some parameters as follows: initial energy = 10J, Number of Nodes = 100, consumed energy by SN =0.5J. Following table 2 shows amount of residual energy after each round.

TABLE 2. EXPECTED RESULTS

<table>
<thead>
<tr>
<th>Round</th>
<th>Number of Nodes</th>
<th>Total Energy</th>
<th>Residual Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>1000</td>
<td>950</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>950</td>
<td>900</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>900</td>
<td>850</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>850</td>
<td>800</td>
</tr>
</tbody>
</table>

VII. CONCLUSION

Clustering is most suitable to reduce communication overhead for large scale wireless sensor networks. The life time of a sensor network is one of the most important parameters in WSNs, which depends on energy level of battery powered sensor nodes and once deployed; the network is left unattended to perform monitoring and reporting functions. Clustering technique shows the self-organized behaviour of WSNs. Energy wastage of sensors in WSN due to direct transmission between sensor nodes and a base station can be avoided by clustering the nodes in WSN. Clustering further enhances scalability of WSN in real world applications while conserving energy. Proposed work gives new clustering technique which aims to minimize the number of clusters and energy consumption per cluster.

REFERENCES