Multi Parameters Based Heterogeneous Clustering Algorithm for Energy Optimization in WSN

Nehal Agrawal¹  
Avadesh Kumar Dixit²  
Sudhir Srivastava³

Abstract—Wireless Sensor Network (WSN) consists of tiny sensor nodes that are able to sense data from vicinity and send this data to the sink or Base Station (BS). The main issue during data collection is the non-rechargeable energy resource or battery. So the data collection should be energy efficient for long use of WSN. As we know that clustering is one of the most efficient technique for saving energy in WSN. Purpose of the paper is to find out multi-factor based clustering process having four conflicting factors: a) delay based on remaining energy, maximum power transmission etc., b) average lifetime of normal sensor nodes, c) lifetime of CHs, and d) maximum power consumption by any normal node. The algorithm is developed with an economical approach for finding CHs using a fitness function. Optimal path is also found utilizing the distance. Results are far better than the traditional algorithms.

Index Terms—Wireless sensor network (WSN); Clustering; Delay; Re-cluster; Lifetime of Nodes; Lifetime of CH; Maximum Power Consumption

I. INTRODUCTION

Wireless sensor networks (WSNs) consist of spatially distributed autonomous sensor nodes [1], which cooperatively collect the information from the surrounding process this information and send the information to the sink. Recently Micro-Electro-Mechanical-System (MEMS) emerged as an advanced technology, that increases the application of WSN [17], like: in disaster management, health and military etc [2]. But due to the implementation in remote areas, it is impossible to replace or charge the battery, that means energy is the main constraint in WSN. So the use of available power should be in effective way. Many researcher provided solutions [3] [7] [8] [13] [14] - [16] for this issue. Routing [3] - [6] and clustering [8] - [12] are the most effective techniques to reduce power consumption. Clustering algorithms divides the whole network into some clusters, Every cluster contains a Cluster Head (CH) that is responsible for collecting the sensed information and convey this information to the sink or Base Station (BS). This information transmits directly or hop by hop. BS assumes to be easily accessed by the user to collect the sensed information.

The base algorithm [18] divided the area into clusters and used only delay to select the CHs. Here, we have proposed clustering preprocessing strategy to select the CHs which considers several factors: a) delay based on remaining energy, maximum power transmission etc., b) average lifetime of normal sensor nodes, c) lifetime of CHs, and d) maximum power consumption by any normal node. Re-clustering is done to restructure the clusters that also helps in reduction of energy consumption. Route identification technique is used to find an optimal path from source to destination. Remaining paper is illustrated as: Literature review is presented in section 2. Section 3 contains the proposed approach in detail. Comparison among proposed algorithm, LEACH and base algorithm is described in section 4. Conclusion is explained in section 5.

II. RELATED WORK

WSN is an important field of research from last few years due to the digitalization phase. Heinzelman et al. [19] proposed one of the most important heuristic clustering algorithms Low Energy Adaptive Clustering Hierarchy (LEACH) that contains two phases in each round, first round is CH selection and another is for data transmission. CHs are selected based on probabilistic approach based on threshold value in first phase and selected CHs broadcast a CH selection message to all its neighbor and normal sensors join the CHs on the basis of signal strength that demonstrate the distance indirectly. A Time Division Multiple Access (TDMA) schedule is provided to each node for sending the sensed data. CHs are responsible to convey the information to the sink. After aggregation the data CHs forward the collected information to the BS. To reduce the load among all the nodes, CHs selection phase is repeated. Randomness is a big disadvantage of LEACH algorithm that makes the CHs based on
probability and did not consider other factors like residual power etc. Many researchers [21], [22] provided central control to the BS using LEACH. In it, BS is having all the information about the location of nodes and residual energy of nodes. It selects the CHs based on different factors that helps in enhancing the lifetime of network but make the process complex for large network, so it is a fruitful solution for small and medium networks only.

LEACH-B [23] algorithm is based on the reduction of CH number in succeeding rounds by examining the remaining power of nodes. This approach fixes the CHs for each round by taking care of residual energy and balance the consumption of energy to enhance the network lifetime.

Halder et al. [24] used archimedes spiral model is used to make uniformity of distance among the nodes. It starts from the center and process to outside area. LEACH-C [25] is an advance version of LEACH where nodes that have higher energy than average of total network are only able to be the CHs. H-LEACH (Hybrid-LEACH) [26] used a partition approach for clustering and data collection. But efficient clustering in partition is provided by Aditya et al. [27] using re-cluster. Efficient clustering is also provided to provide information about any abnormal activity in pipelines [29]. Node placement [31], data aggregation [30] are also very helpful techniques for reducing the energy consumption. Multi-objective problem solutions [28] are also used to in various clustering approaches.

Here, we have proposed a clustering approach by considering that the network is a heterogeneous network where all nodes are of different characteristics and the proposed clustering considers a fitness function that is based on many factors.

III. PROPOSED WORK
Here we have described our proposed approach considering various factors and the advantage of choosing these factors. Network model-
Here we have assumed that many sensors of different characteristics are distributed in a region and are static in nature. BS has storage and a battery having capability to perform infinite processing. BS knows the location of nodes, residual energy of nodes and also having the ability to perform computations. BS is also static in nature means its location is fixed. All nodes having information regarding the location of BS.

A. Clustering-
Clustering is the technique to divide the whole network into clusters so that the energy consumption could be minimized. The clusters should also be optimized to make energy consumption optimized.

Here we have proposed the clustering algorithm based on four conflicting factors. Fitness function is designed to make coordination among these factors and also to select the best CHs. The factors are as follows:
1. Delay considering total transmission power required, residual energy etc.,
2. Lifetime of CHs,
3. Average lifetime of nodes, and
4. Maximum power consumption by any node.

Coordination among factors shows the solution will have lower delay with high lifetime of nodes and CHs, and also perform load balance by taking care of maximum power consumption by any node. This combination will lead to the higher lifetime of whole WSN. Clustering phase performs following steps:

1) Cluster Formation: We divided the whole WSN into optimal number of clusters same as in base paper [18]. CHs are chosen in next phase.
2) CH Determination: After creation of clusters, we need to identify a CH in each cluster. Here, we describe the CHs selection approach to select the optimal CHs using four conflicting factors. It compiles following steps:

a) Step 1: Delay Calculation:
Delay is inversely proportion to the remaining power of nodes. Its minimum value is required and it confirms higher remaining power of sensors. So lower delay will help in enhancing the lifetime of WSN.

$$\text{Delay} = \left( \frac{E_{\text{initial}} - E_{\text{residual}}}{E_{\text{initial}}} + r \right) \times \text{RTD}$$

where $r$ - random number (0<r<1), RTD - round trip delay.

b) Step 2: Computation of CH Lifetime:
CH lifetime determines the number of rounds till the node can work as a CH before being out of energy. The higher value of CH lifetime shows that the node is a good option to choose as a CH since it will increase the total lifetime of WSN.
c) **Step 3: Computation of Average of Nodes Lifetime:**

It shows the average lifetime of member nodes in the cluster if a node is chosen as a CH. Its higher value shows how much time the normal nodes will send the data if the node becomes a CH. It is calculate as

\[
\text{Node}_i\text{-Life} = \frac{E_{\text{res}\_Node}_i}{\text{Transmission\_Energy\_Needed\_Node\_to\_CH}}
\]

d) **Step 4: Maximum Power Required by Node:**

It shows maximum transmission power required by a particular node in the cluster to transmit data to the respective CH. Its lower value is required to restrict the overhead on a single node and confirms load balancing during the transmission of data to CH. It is calculated as:

\[
\text{Max\_Power\_Node} = \max(E_T{\_NodesInCluster\_to\_CH})
\]

where \(E_T\) - Transmission energy required.

e) **Step 5: Computation of Fitness Function:**

Fitness function helps in making a cooperation among the considered conflicting factors to optimized the energy consumption by selecting the effective CHs. Here we computes the fitness function as:

\[
\text{Fitness Value} = \alpha \ast F_1 - \beta \ast F_2 - \gamma \ast F_3 + \eta \ast F_4
\]

where, factor \(F_1=\text{Delay, F}_2=\text{CH Life, F}_3=\text{Avg Nodes Life, and F}_4=\text{Max Power Node, } \alpha=0.25; \beta=0.3; \gamma=0.3; \text{and } \eta=0.15.\) Algorithm is explain in algorithm 1 using all.

f) **Step 6:**

Fitness value is calculated for all nodes in each cluster and node having lower fitness value will be chosen as a CH for that particular cluster and rest nodes will work as normal nodes.

3) **Re-Cluster:** Since the CHs are selected from each cluster, now there is a chance that the member sensor node is far from the selected CH in that cluster and nearer to a CH selected in other cluster. So the re-cluster phase is included to reduce some extra energy consumption. In this phase clusters are formed again according to the selected CHs and each member sensor node join the nearest CH.

**Algorithm 1: CH selection approach to find optimal CHs.**

Data- Transmission energy required \(E(T), \text{BS and Nodes Location, Initial Power(i)}\)

Result: CH in each cluster

1 start
2 for each node k do
3 for every round do
4 Compute \(E_{\text{res}}\) of k;
5 Compute \(\text{Delay for k};\)
6 Compute \(\text{Avg\_Nodes\_Life};\)
7 Compute \(\text{Max\_Power\_Node};\)
8 End
9 End
10 Let \(\text{Sa(fuz\_Value)}, \text{Sb(fuz\_Value)}\) as the fuzzy values of sensors \(\text{Sa and Sb};\)
11 \(E_{\text{minimum}}\) is computed as the minimum power required to collect information and transmit to the sink:// Process for all sensors in the cluster;
12 for every sensor in the cluster; 13 if((\(\text{Sa(Fit\_Value)}<\text{Sb(Fit\_Value)}\) && \(E_{\text{res}}(k)>E_{\text{minimum}}+\text{threshold})\)
14 { now \text{Sa will be a sensor with optimal fuzzy value;}
15 \text{Sa announced as the selected CH}
16 17 }
18 else
19 { \(\text{Sa will work as a member sensor and sense the data from its vicinity}
20 21 }
22 end
23 End
B Route Selection Approach:
Now there is a need to select the most favorable route to transfer the sensed information to the sink. So here we have where, factor F1=Delay, F2=CH Life, F3=Avg Nodes Life, and F4=Max Power Node, \(a=0.25\); \(\beta=0.3\); \(\gamma=0.3\); and \(\eta=0.15\). Algorithm is explain in algorithm 1 using all.

g) **Step 6:** Fitness value is calculated for all nodes in each cluster and node having lower fitness value will be chosen as a CH for that particular cluster and rest nodes will work as normal nodes.

3) **Re-Cluster:** Since the CHs are selected from each cluster, now there is a chance that the member sensor node is far from the selected CH in that cluster and nearer to a CH selected in other cluster. So the re-cluster phase is included to reduce some extra energy consumption. In this phase clusters are formed again according to the selected CHs and each member sensor node join the nearest CH updated the route selection process provided in the base paper [18] by considering the \(d_0\) distance as this distance plays an important role in energy consumption. Algorithm 2 describes the process in detail [32].

IV. RESULT AND EXPERIMENTAL SIMULATION
We have used MATLAB to model our WSN for experimental purpose. Since MATLAB provides a realistic and object oriented support so it shows the actual and clear energy dissemination and information collection process using our proposed algorithm, base approach (EEHC) and LEACH approach. So MATLAB provides a clear comparison among the algorithms. We have taken many cases for the variable number of nodes and BS position. We have used two parameter number of dead nodes after each rounds and total remaining energy of nodes after each round. Dead nodes comparison is shown

Algorithm 2: Path selection approach for information collection using WSN

**Data:** Network after Re-cluster phase

**Result:** Optimal Path between CH and Sink

1. **for every node** \(k\) **do**
   2. **if (Sensor acts as a member) then**
      3. transmit the information to the nearest CH;
   4. **end**
   5. **if (Sensor acts as a CH) then**
      6. **if (Sink Distance <\(d_0\)) then**
         7. transmit the information to sink directly;
      8. **else**
         9. **if \(E_{res_k} > (Transmission\ Power\ to\ d_0\ distance\ +\ Threshold)\) then**
            10. transmit the information to a CH that has a distance near and less than \(d_0\), if there is no such CH then transmit information to the nearest CH;
      11. **end**
      12. **else**
         13. transmit the information to the nearest CH;
      14. **end**
      15. **end**
   17. **end**
   18. **if \(s\ is\ an\ obstacle\) then**
      19. Process Shortest Path finding approach on vertex V1;
      20. Process Dijkstra (G, Source)
   21. **for every vertex** \(u\ in G\) **do**
      22. distance cost \((u)=\infty\)
      23. previous vertex \((u)=\)undefined;
      24. distanceTo(Source)=0;
      25. Set=all vertex in G;
      26. **while Set is not empty do**
      27. \(v=\) vertex in Set having the lowest distance\(u\);
      28. remove vertex \(v\ from Set;\)
      29. **for every nearest vertex** \(u\ of v\ do**
      30. new distance=distance_cost\(u\)+distance_cost\(v,u\);
31if
   (new distance < distance cost(u))
   then
   32distance_cost(u)=new distance;
   33previous_vertex(u)=v;
   34return_previous_vertex();
   35end
   36end
   37end
   38end
   40else
41Transmit the information to nearest CH or BS.
42else
43Identify sensor k.
44end

by 1, 3, 5, and 7 figures and remaining power comparison is shown by 2, 4, and 6 figures. Results verifies that proposed algorithm is the most favourable approach among all three compared algorithms.

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of Network</td>
<td>100m x 100m</td>
</tr>
<tr>
<td>Position of Sensors</td>
<td>Randomly deployed</td>
</tr>
<tr>
<td>Primary Power</td>
<td>1J</td>
</tr>
<tr>
<td>Deployed sensors count</td>
<td>varying</td>
</tr>
<tr>
<td>Transferral Energy</td>
<td>100W</td>
</tr>
<tr>
<td>Rounds</td>
<td>varying</td>
</tr>
</tbody>
</table>
CASE 1: When position of BS = (0, 0), Total sensors = 500

Fig. 1: Comparison graph for dead sensors till $R_{max}=10$

Fig. 2: Comparison graph for remaining energy of network till $R_{max}=100$

CASE 2: When position of BS = (50,50), Total sensors = 100

Fig. 3: Comparison graph for dead sensors till $R_{max}=100$

Fig. 4: Comparison graph for remaining energy of network till $R_{max}=100$

CASE 3: When position of BS = (200,200), Total sensors = 10

Fig. 5: Comparison graph for dead sensors till $R_{max}=500$
CASE 4: When position of BS = (0,0), Total sensors = 100

V. CONCLUSION

Clustering approach helps in power utilization during information gathering process. CH selection using single factor is not a fruitful process because there are many other conflicting factors that strike the network lifetime. The proposed method provides a clustering process that makes cooperation among four conflicting factors. So the identified CHs using proposed process are more favorable than other algorithms because the selected CHs are of higher lifetime and also having high remaining power. The member sensors are also have higher lifetime and no member node consumes large amount of energy. Re-cluster phase also reduces energy consumption and provides lower distance to CH. Path identification technique is also provide utilization of power during packet forwarding. Results verify the superiority of proposed algorithm over base and LEACH algorithm.

REFERENCES
