Optimization of Node placement and Clustering in Wireless sensor networks using Genetic Algorithm

Pooja A. Chavan\(^1\) Dr. R. D. Kanphade\(^2\) Prof. S. V. Joshi\(^3\)

\(^1\)M.E. Communication Networks, NMIET, Talegaon-Dabhade, India
\(^2\)Professor & Principal, NMIET, Talegaon-Dabhade, India
\(^3\)Assistant Professor, NMIET, Talegaon-Dabhade, India

Abstract—Wireless sensor networks are having many applications in various industries therefore it is essential to improve its currently present parameters and protocols. The WSN steps includes deployment of nodes, clustering, network coverage and aggregation of data. In this paper the concentration is on the nodes placement and clustering of WSN. The simulation and examination of GA based nodes placement and HEAL is conducted for evaluation of no. of movements and dist. travelled using NS2 and The simulations and analysis of the Genetic algorithm based clustering and LEACH based clustering is conducted for evaluation of energy spending and no. of alive nodes using NS2.

Keywords—Wireless sensor network (WSN), Genetic Algorithm (GA), Clustering, Cluster head, Low energy adaptive clustering hierarchy (LEACH), node placement

I. INTRODUCTION

The nodes placement plays an vital role in the optimization of sensor networks. The deployment of nodes on an object field enhances the network performance. Figure 1 shows the nodes placement, for placement of nodes in network, there are three main categories as the grid placement, random placement and stochastic node placement. The nodes additional energy for longer distance transmission, which is not energy effective.

Another main operational step of sensor networks is Clustering of the network. Clustering is method of arrangement of the nodes from the sensor network to form the groups which is termed as division cluster. CH role is to collect information from the nodes located in that cluster and send the information to the sink directly or via. The other clusters heads. The data broadcast to the sink from CH through other cluster head in such cases the optimized clustering drops load amongst sensor nodes effectively which outcomes in more lifetime and energy saving of the sensors in WSNs. Figure 2 demonstrates nodes clustering in network. The enhanced security, less extra data are additional advantages of clustering.

II. RELATED WORK

In the various literature, the necessity of the optimization of the sensor networks is emphasized and GA is used. In [3], author Serbas studied the routing method with least energy consumption in WSN. The routing protocols with high energy saving (Gossiping, PEGASIS, LEACH, Director Diffusion and EESR) are examined. In [6], author S. Hussain has offered the hierarchical clusters to decrease the energy expenditure requirements for networks. For information circulation in WSNs, GA is used to create energy effective clusters. The output show the network lifetime can increase by using hierarchical intelligent technique of clustering for several deployment of network. In [7], authors used the dynamic technique of GA and studied the enrichment of lifespan and energy spending of network to decide the ideal clustering state. They have used offered equation with the simple framework for improvement of lifespan alongside network coverage. Lastly, the projected algorithm is more efficient. In [8], author C. Sergiu analyzed and compared the performance of four ways of nodes placement on a network. Node deployment’s simulation results indicates

---

**Fig. 1:** Nodes placement

**Fig. 2:** Clustering in sensor networks
the enhanced algorithm performance. In [9], authors studied a GA centered technique and apply for a self-organize network. Fitness function is created considering the connectivity, compactness and energy spending of network.

III. NODE PLACEMENT IN WSN

To make sure the network operates with the highest possible performance, the nodes deploy in the network with balanced energy expenditure of all nodes and with preferred node placement protocol.

The Proposed Fitness Function:

Fitness function formed on the extension of area under coverage is presented below to obtain the optimum solution to cover wider area and maintain high efficiency of energy spending. The R is the radius centered at each sensor [10]. The area of union is normalized by total area.

\[
\text{Fitness-function} = \frac{\text{min} - f_i}{(\text{coverage}_i)} \quad (1)
\]

The indicated fitness is taken into consideration for coverage and energy is denoted.

HEAL:

Competently calculate and advance the area covered in WSNs. HEAL is consists of two steps. In the first step the hole identification and border detection is done. The hole detection algorithm is less complex and used for different sizes of holes. The second stage hole healing. It comprises two tasks; find hole heal area and nodes relocation. The only nodes which are near to hole are participating in healing action.

GA based node placement:

Hole healing is takes place as the stages like Hole identify and detect border. In hole identification and healing: Every node of the network splits the area into four quadrants by keeping itself as center and checks for neighbors in each region. Node with single neighbor is considered as stuck node. The quad rule identify the stuck node which finds the hole by Hole detection algorithm for healing the found hole [2]

Find HC = \{(X_{\text{min}} + X_{\text{max}}) / 2, (Y_{\text{min}} + Y_{\text{max}}) / 2\} \quad (2)

Where, HC - Hole Center, X_{\text{min}}, X_{\text{max}}, Y_{\text{min}}, Y_{\text{max}} - Minimum and Maximum X and Y coordinates among stuck nodes.

HA = 3.14 \times R_Si \times R_Si \\
SA = 3.14 \times R_Si \times R_Si

Required no. of nodes to cover the HHA = HA/SA

R_Si - Hole radius (Longest dist. bet. two stuck nodes), R_Si - Sensing radius of sensor Si, HA - Hole Area, SA - Sensor Area

Coverage is obtained in terms of dist. bet. nodes and the nodes that have less dist. indicates high overlapping area when compared to other nodes and it outcomes in less fitness nodes. By only moving less fitness nodes appropriate coverage and reduced energy expenditure can be achieved [1].

Fitness= \frac{\text{Avg.Dist.}}{(\text{Rc} \times \text{Avg.dist})} \quad (3)

Avg.dist is the length from moving node to centre of hole. Rc is the node range communication.

Simulation and evaluation results:

Software Requirements:

OS: Ubuntu 15.04 LTS 64bit
Simulator: NS 2.35
Language: C++, TCL and AWK script

Simulation Model:

<table>
<thead>
<tr>
<th>Simulator</th>
<th>Network Simulator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of nodes</td>
<td>200,250,300,350</td>
</tr>
<tr>
<td>Area</td>
<td>200m x 200m</td>
</tr>
<tr>
<td>Communication Range</td>
<td>24m</td>
</tr>
<tr>
<td>Sensing Range</td>
<td>12m</td>
</tr>
<tr>
<td>Interface Type</td>
<td>Phy/ WirelessPhy</td>
</tr>
<tr>
<td>Mac Type</td>
<td>IEEE 802.11</td>
</tr>
<tr>
<td>Queue Type</td>
<td>Droptail/Priority Queue</td>
</tr>
<tr>
<td>Queue Length</td>
<td>50 Packets</td>
</tr>
<tr>
<td>Antenna Type</td>
<td>Omni Antenna</td>
</tr>
<tr>
<td>Propagation Type</td>
<td>TwoRayGround</td>
</tr>
<tr>
<td>Transport Agent</td>
<td>UDP</td>
</tr>
<tr>
<td>Application Agent</td>
<td>CBR</td>
</tr>
<tr>
<td>Simulation Time</td>
<td>50 seconds</td>
</tr>
</tbody>
</table>

Hole identification and healing:

Initially sensor nodes are randomly deployed. To advance the coverage initial random node placement is adjusted through hole detection process. In the HEAL, the only nodes which are near to hole are participating in action of healing hole. However, in the GA approach the nodes having the low fitness will be moved to heal the hole.

Fig.3: NAM- Hole identification and healing
A node which is not able to communicate at least any one direction in network then it is known as stuck nodes. Nodes located in the border of network are eliminated from stuck nodes. Those nodes are colored in blue.

Stuck nodes exchanges HD packet and lowest ID node is identified as Hole Manager. Hole Manager: 60

Nodes in dodger blue color belongs to HHA.

Hole Center:  (80.0, 75.0)

Sensor Area:  452.15 m²

HHA = 20096.0 m²

Required no. of nodes to cover the HHA = 44

Available Nodes in HHA: 125

A population (chromosome) is produced using nodes located inside the area of hole healing. Every chromosome is created with no. of genes. Every gene denotes the node inside area of hole healing. Every gene is randomly initialized as 0 and 1. If the gene is initialized as 0 then it signifies it has to move to heal the hole.

The less fitness chromosome qualified for movement towards hole center. Fitness computation is performed in iterative manner till max iteration is reached. In each iteration best fitness value and its corresponding chromosome is found. New generation of chromosome is made in preceding iterations using the cross over and mutation process. These processes are performed by best chromosome found in previous iteration. Fitness is computed for newly generated chromosome and so on. Nodes found in the finally identified chromosome is moved towards hole center.

Last three low fitness nodes are moved to the exact hole center. Last three low fitness nodes (38, 4, 115) are the nodes those leads to high overlapping. By moving those nodes, node placement is optimized in terms of reduced overlap and improved coverage.

Fitness for population after 6th generation:

Total Fitness for population 0=3.4683147548011761
Best Fitness=3.354653584670825 and population=1

Total Fitness for population 1=3.354653584670825
Best Fitness=3.354653584670825 and population=1

Total Fitness for population 2=3.692619586277659
Best Fitness=3.354653584670825 and population=1

Total Fitness for population 3=4.0704410618019118
Best Fitness=3.354653584670825 and population=1

Generation:6 Overall best Fitness=3.54653584670825 and population=1

low fitness nodes = 38, 4, 115

After the healing process of existing hole, rediscovery of the stuck nodes is done. No node is stuck in the network at this time.

Performance metrics:

Total Distance Traveled: The total dist. traveled by the sensor nodes to optimize the node placement

No. of movements: It is the no. of movements carried out by a node to optimize the node placement

Performance Evaluation and results:

Comparison is made amongst the GA approach proposed here and existing HEAL protocol. Total Dist. traveled and No. of movements are assessed against no. of nodes. After the healing process of existing hole, rediscovery of stuck nodes is done. No node is stuck in the network at this time. Based on output using the graph tool in NS2 the result graphs are plotted.
IV. CLUSTERING IN WSN

To enhance the network’s lifetime with least-energy expenditure can be achieved by optimized clustering. The clustering focused applications are mainly for monitoring with requirement of continuous data transmission from sensors, therefore the cost reduction is essential for timely data message sending by routing protocols.

The Proposed Fitness Function:

The formula [1] for getting the best energy spending by using the more coverage of the sensor is given below. In that, \((e_\text{i}^*T)\times(e_\text{j}^*T)\) denotes for the whole energy spending and \(((D_a \times \text{nodes})\times (D_b \times \text{CHs}))\) denotes for entire length bet. sensors and nodes and every cluster are multiplied by total length among CHs[1].

\[
F(i) = (e_i^*T/Da^*\text{Nodes})*(e_j^*T/Db^*\text{CHs})
\]  

(4)

The intelligent suitability function \(F(i)\) is competent for ranking chromosome. The ideal chromosome choice is done considering passing generation to acquire the ideal solution[1].

LEACH based clustering:

The hierarchical approach is used in the LEACH protocol for clustering the network. The CH deals in each cluster. The collecting data sent by nodes, aggregation of data after collecting, and transmission of aggregated data to the BS are the duties of the CH. Generate TDMA schedule to allot a time slot to every cluster which is used for transmission of information is the key activity. The cluster participant node-
sacquire the schedule when the cluster head circulates it.

GA based clustering:

For the lifespan enhancement of the network and to lessen energy spending, the parameters of the GA were fixed as per the software services. If we reduce the length bet. participant nodes and relevant CH for this case every cluster might have one or more CHs, which is costly considering the energy spending. If we chose more clusters we can avoid longer distances. Considering this, to accomplish average energy expenditure by every node, a ratio of entire energy consumed to the total distances of nodes was indicated. The intelligent suitability function \(F(i)\) is competent of ranking chromosome which is found by total energy consumption, distance among nodes and CHs.

\[
\text{Fitness} = (1/D\text{CHN})*(1/D\text{CHBS})
\]  

(5)

Where, DCHN denotes the dist. between sensors and CH and DCHBS denotes the dist. bet. CH and BS.

Simulation and evaluation results:

Software Requirements:

- OS: Ubuntu 15.04 LTS 64bit
- Simulator: NS 2.35
- Language: C++, TCL and AWK script

Simulation Model:

<table>
<thead>
<tr>
<th>Simulator</th>
<th>Network Simulator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>50</td>
</tr>
<tr>
<td>Area</td>
<td>600m x 600m</td>
</tr>
<tr>
<td>Communication Range</td>
<td>200m</td>
</tr>
<tr>
<td>Interface Type</td>
<td>Phy/WirelessPhy</td>
</tr>
<tr>
<td>Mac Type</td>
<td>IEEE 802.11</td>
</tr>
<tr>
<td>Queue Type</td>
<td>Droptail/Priority Queue</td>
</tr>
<tr>
<td>Queue Length</td>
<td>50 Packets</td>
</tr>
<tr>
<td>Antenna Type</td>
<td>Omni Antenna</td>
</tr>
<tr>
<td>Propagation Type</td>
<td>TwoRayGround</td>
</tr>
<tr>
<td>Routing Protocol</td>
<td>LEACH, GA_Clustering</td>
</tr>
<tr>
<td>Transport Agent</td>
<td>UDP</td>
</tr>
<tr>
<td>Application Agent</td>
<td>LeachApp</td>
</tr>
<tr>
<td>Initial Energy</td>
<td>2 Joules</td>
</tr>
<tr>
<td>Simulation Time</td>
<td>300seconds</td>
</tr>
</tbody>
</table>

Cluster head selection:

In the LEACH the cluster heads are chosen based on all the sensors that generates the random no. less than the threshold. However, in the generic algorithm process for selecting the CH the process of crossover and mutation are applied. If no. of CHs are more or less then it result in extra energy spending.
Genetic algorithm is applied on the chromosomes. The best fitness chromosome is utilized in the clustering process. Each chromosome comprises of genes. Chromosome replicates the nodes in the network. Each gene in the chromosome replicates the state of the node. The optimized no. of CH can produce increased fitness.

Fitness for candidates in each generation:
Fitness=0.140888 generation=0 num_cluster=10 dchn=13539.540048 dchbs=5242.320978 Fitness=0.197813 generation=0 num_cluster=7 dchn=14474.075481 dchbs=3492.639065 Fitness=0.192114 generation=1 num_cluster=7 dchn=15506.475750 dchbs=3356.825089 Fitness=0.278206 generation=1 num_cluster=6 dchn=12044.474444 dchbs=2984.318051 Fitness=0.197492 generation=2 num_cluster=4 dchn=14529.934438 dchbs=3484.865549 Fitness=0.547502 generation=2 num_cluster=4 dchn=9205.368031 dchbs=1984.143102
Best fitness = 0.547502

Best fitness after 2nd generation:

The state of all the other nodes are 0. The node state is CH and non-CH corresponding to the value 1 and 0 respectively. Therefore, the nodes 6, 14, 16, 32 having state 1 are selected as CH as per Genetic algorithm.

Performance Metrics
Energy consumption: The total energy expenditure by sensor nodes.
No. of alive nodes: The total nodes remain alive in network.

Performance Evaluation and results:
Comparison is made in the proposed GA based approach and LEACH cluster-based protocol. Both are simulated based on simulation model and examined. Clustering and data transfer is repeated for each round and no. of alive nodes and energy consumed is measured in each round. Using the graph tool in NS2 the results are plotted.

Fig. 7: NAM- Data transmission in the network

Advertise by cluster head 32
Join Request:
27 sending JOIN_REQ to 32, distance = 146.000000, at time 13.263694 mac=32
22 sending JOIN_REQ to 32, distance = 112.000000, at time 13.323019 mac=32

Fig. 8: No. of alive nodes vs. No. of rounds
V. CONCLUSION

In Node placement distance travelled and total no. of movements for optimized node placement are reduced in proposed GA approach when compared to HEAL, nodes deployment is carried out in optimized manner based on fitness of the node that depends on the node dist. to hole centre and node coverage area whereas in the existing HEAL, the neighbour nodes are participate in hole healing and move towards hole centre. The reduced no. of movements and distance travelled result in less energy expenditure in network this in turn improves network lifetime. In Clustering, the proposed GA based clustering achieves reduced energy expenditure and increased no. of alive nodes when compared to LEACH, through the optimized no. of CH selection. The optimized no. of CH can produce increased fitness. Choosing high fitness chromosome reduces energy spending and increases no. of alive nodes in GA based clustering.

REFERENCES