

# Energy Efficient Communication Framework for Target Coverage using Trust Concepts



Pooja Chaturvedi, A. K. Daniel

**Abstract:** Target coverage is a challenging task in the field of wireless sensor networks aiming to observe a set of critical targets while considering the limited resources and the network lifetime is improved. The paper proposes an approach to : i) determine the strategy according to which the critical targets can be monitored while satisfying a certain confidence level. ii) determine the maximum and minimum number of nodes which can guarantee the coverage, iii) determine the optimal number of active nodes for various deployment strategies, iv) to determine a routing mechanism using either single hop/multi hop communication based on the reachability of the node to the base station and iv) to develop an aggregation protocol which can reduce the redundancy and number of packet transmissions. The proposed protocol is based on the two level aggregation at the set cover level and at the cluster level using the concept of Locality Sensitive Hashing (LSH) and Jaccard Similarity measure. The efficiency of the proposed aggregation mechanism is determined for various data sets of multiple dimensions. The results obtained through the simulations show the improvement in the network performance with respect to the network longevity, coverage, reliability and of the data transmission as compared to the Boolean coverage model

**Keywords:** Clustering, Network Lifetime, Target Coverage, Fuzzy Inference, Aggregation, Routing, Deployment

## I. INTRODUCTION

Energy conservation and coverage are two major challenges concerning the functionality of the wireless sensor networks. Coverage is usually considered as a quality of service parameter which determines the duration and magnitude of coverage of the given target region, whereas energy conservation approaches aim to devise a mechanism to maximize the network lifetime. These objectives can be met by scheduling the nodes in various set covers such that all the targets are monitored. The sensor nodes are usually deployed in random manner, so nearby sensors tend to sense the redundant or correlated data. In this case the energy conservation is achieved by maintaining the unnecessary nodes in the low energy mode, while other nodes are kept in high energy mode to observe the targets according to the mechanism devised by the base station. A scheduling

approach for target coverage based on coverage probability, node's contribution and trust mechanism is proposed in [1]. The protocol aims to enhance the network lifetime by obtaining several set covers containing the optimal count of nodes. The lifetime of the network is directly proportional to the set covers count.

The work in this paper extends the work of [1] incorporating the deployment mechanisms which can achieve full coverage for the desired region. The minimum and maximum count of nodes to be deployed in a non-overlapping environment is determined such that the desired targets can be monitored. A hybrid routing mechanism for the sensor nodes to transmit the observed data towards the base station as either directly if it's within the sensor node's communication range or using intermediate nodes. In case of the multi hop communication the next forwarding node is selected on the basis of fuzzy inference mechanism which considers the residual energy, link strength and the distance of the nodes. To eradicate the redundancy present in the sensed data, an aggregation mechanism is proposed at the set cover level and the cluster head level considering the Jaccard similarity measure and locality sensitive hashing.

The rest of the paper organization is as follows: state of the art in section 2, proposed deployment, routing and aggregation approaches in section 3, simulation results and discussion in section 4 and section 5 is about the conclusion.

## II. STATE OF THE ART

The target coverage problem in a sensor network consisting of  $n$  sensor nodes as  $S = \{s_1, s_2, \dots, s_n\}$  and  $m$  target nodes as  $T = \{t_1, t_2, \dots, t_m\}$  is the calculation of the set covers to observe all the targets such that the network functionality is prolonged. The set cover is termed as the set of nodes which can observe all the targets with the required coverage level. The set cover are considered to have the lowest number of nodes and is functional till the energy reservoir of the nodes. A network as shown in Fig1 is considered. The nodes as  $S = \{s_1, s_2, s_3\}$  and targets as  $T = \{t_1, t_2, t_3, t_4\}$  and targets are observed as:  $s_1 = \{t_1, t_3\}, s_2 = \{t_2, t_3\}, s_3 = \{t_2, t_3, t_4\}$

Revised Manuscript Received on February 05, 2020.

\* Correspondence Author

**Pooja Chaturvedi\***, Department of Computer Science, School of Management Sciences, Varanasi, India.

E-mail: chaturvedi.pooja03@gmail.com

**A. K. Daniel**, M. M. M. University of Technology, Gorakhpur, India.  
E-mail: danielak@rediffmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

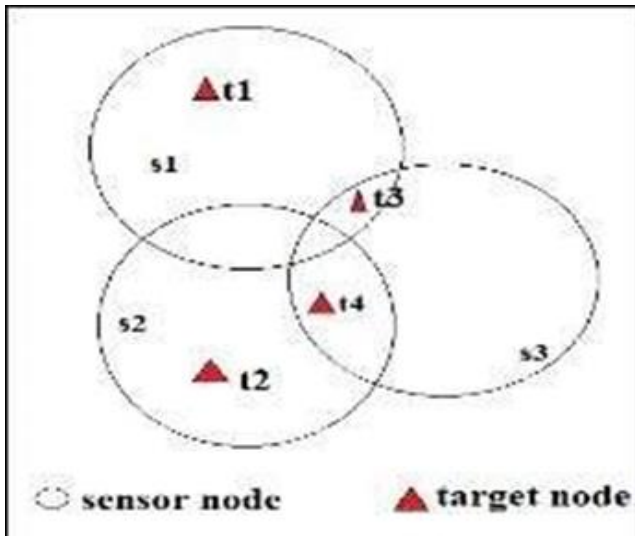


Fig.1. Network of nodes and targets

Consider the targets can be observed by the nodes for 0.2 time units then the network lifetime will be 0.2 time units if all the targets are observed simultaneously by the nodes. The targets are monitored as per the set cover scheduling as:  $C1=\{s1,s2,s3\}$ ,  $C2=\{s1,s3\}$  and suppose the duration of observation of each set covers  $i$  as: 0.4 and 0.6 then the network lifetime can be extended to 1.0 time units.

The target coverage approaches are categorized as: target coverage, energy efficient and connected, target coverage with adjustable sensing range, energy efficient, target coverage under QoS constraints, disjoint and non-disjoint approach, centralized and distributed and [7]-[10]. The nodes in the sensor network are often resource constrained, so energy conservation is desirable at both the hardware and designing levels. The nodes can exist in any one of the four states in the sensor network as idle, transmitting, receiving or sleep. It is established that the nodes consume the most energy in the communication modes, so keeping the nodes in idle/sleep states whenever feasible can lower the energy consumption significantly. Scheduling the nodes in several set covers to be active is one of the solutions to achieve these objectives. The quality of the data sent towards the base station is also crucial to improve the network efficiency. The sensor nodes in the nearby region sense the same environmental features and usually transmit the data to a common sink so some redundancy exists in the network. In such situations data aggregation is considered as effective approach which gather the data from multiple sources by removing the redundancy and reduces the count of message transmissions. These requirements encourage the need to develop data centric mechanism which emphasizes on selecting paths from various originators towards a single end point which ensures the elimination of repeated data within the network rather than the address centric approach which aims to find the optimal routes between the reachable nodes [11] The aggregation of data is emphasized on the transformation of the data using the various techniques. The process of aggregation is as shown in the Fig.2:

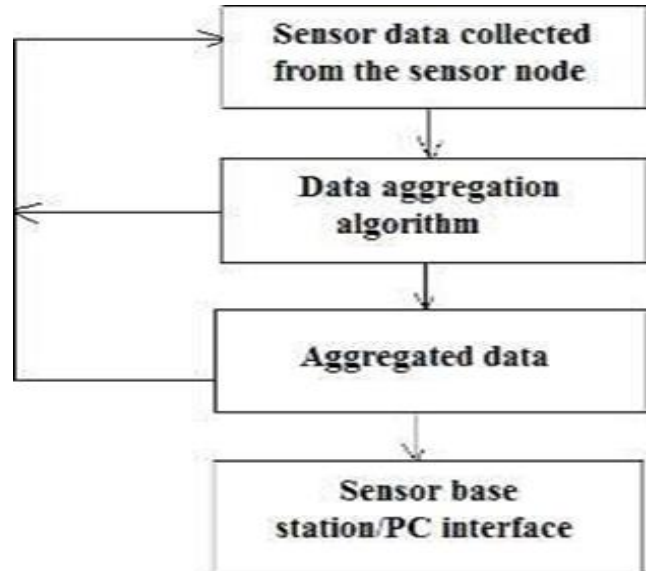


Fig.2. Steps of Data Aggregation

The sensor node collects the data and forwards it to the aggregator, where it is processed to remove the redundancy and duplicate data by using some aggregation algorithm such as min, max, average etc. The data is then forwarded towards the base station. The aggregation based routing schemes can be categorized as two types: address centric and data centric. In address oriented protocol the originator nodes forwards the data independently via the shortest routes towards the sink. In the data oriented approach some sort of aggregation is done on the data which is generated by the multiple sources. The aggregation schemes can be of two types: Optimal aggregation and sub optimal aggregation. In the optimal aggregation the data aggregation tree is constructed rooted at the sink. In this multiple sources sends the data to a sink. The sub optimal aggregation schemes are of three types: center at the nearest source, shortest path tree and greedy incremental tree. In the center at nearest source the aggregation point is the node which is nearest with respect to the central node. In the shortest route tree approach the nodes sends the sensed data along the shortest route to the sink. In the greedy incremental tree based approach the aggregation tree is constructed at each step. In [12] the efficiency of sensor network with and without aggregation. In [13] Dynamic balanced spanning tree approach is proposed which eliminated the problem of hotspot and removed the drawbacks of fixed spanning tree based approaches. The results show that this approach minimized the energy consumption and balanced the traffic load. In [14] authors have proposed an approach to remove the redundancy on the basis of support vector machines. In this paper authors have used the Locality sensitive hashing based approach to detect the outliers in the data. In [15], a cluster based aggregation approach is devised to remove the redundancy in the data by using the inter cluster and intra cluster aggregation mechanisms. In [1], a node scheduling strategy based on coverage probability, node contribution and the node's rust values is devised which aim to determine a number of set covers whose nodes are activated and other nodes are kept in low energy state.

The set covers consist of node which can observe all the targets with the desired satisfaction level. The activation sequence of the set covers is devised by the central processing node. The paper determines the upper and lower bound on the count of nodes to achieve the desired coverage level. The hybrid routing protocol is proposed which determines the next forwarding node considering the distance, residual energy and link strength parameters. To further enhance the network efficiency we have incorporated the locality sensitive hashing based aggregation mechanism at the set cover level and the cluster head level. The proposed extension provides the network improvement in terms of coverage level, overhead, reliability and energy efficiency.

### III. PROPOSED PROTOCOL

Consider a network of  $s$  sensor nodes  $S = \{s_1, s_2, \dots, s_s\}$  and  $t$  targets  $T = \{t_1, t_2, \dots, t_t\}$ , we define the target coverage problem

as which aim to optimize the function  $F = \sum_{b=1}^t A_b$  under

the following constraints:

$$\left\{ \begin{array}{l} x_a \times \sum_{b=1}^t P_{ab}^{obs} > \mathbf{0} \quad \forall t_a \in T \\ P_{c_{vr}}(a) \geq CL \quad \forall t_a \in T \\ x_a \in (0, 1) \quad \forall s_a \in S \end{array} \right.$$

The objective function aims to include the optimal number of nodes in the set cover.  $CL$  represents the coverage level required to monitor a target.  $X_{ab}=1$  if the target  $T_b$  is monitored by the node  $S$  else  $0$ .  $P^{obs}$  represents the observation probability of the target as:

$$P_{ab}^{obs} = Cov(s, t) \times STL \quad (1)$$

$STL$  represents the trust level of the sensor node.

$P_{c_{vr}}(a)$  represents the probability that the target region is covered by any sensor node.

The trust levels of the nodes are calculated using the direct trust, recommendation trust and indirect trust as in [2]. The proposed scheduling protocol works in rounds where each round consists of three phases as: setup phase, sensing phase and transmission phase. In the set up phase the optimal number of nodes is determined which is required for the complete coverage of the target region. The nodes are further divided into various clusters using the energy, distance and node degree parameters. The cluster heads are responsible for the calculation of the trust values of the nodes. The base station schedules the nodes in various set covers on the basis of the coverage probability, trust values and the node contribution in various set covers. The set covers are activated periodically and the active nodes sense the data from the environment. In the transmission phase the data is transmitted to the base station using either the single hop or multi hop communication.

Steps:

#### 1. Setup Phase:

- a. Deployment of the nodes in the regions R1 and R2.
- b. The nodes in region R2 are divided into a number of clusters. The cluster heads are chosen dynamically using

the fuzzy rule based on the residual energy, distance and the node degree.

- c. Aggregation at the cluster head is performed level using the hashing technique to reduce the number of bits required to transmit the message.
- d. Set cover determination by the base station on the basis of coverage probability and trust level. The fuzzy inference mechanism is used to determine the node status as active/sleep.

#### 2. Sensing Phase:

- a. The nodes perform the sensing task according to the schedule determined by the base station.
- b. A leader node is selected within each set cover based on the residual energy and distance parameter. The aggregation of the data is performed at the leader node within each set cover.

#### 3. Data Transmission Phase

- a. If the base station lies within the sensing range of the leader node, it transmits the sensed data directly to it.
- b. Else the next forwarding is determined using the fuzzy logic based on the parameters distance, residual energy and link strength.

Pseudo Code of the Proposed Protocol

Table I. Pseudo code of the proposed protocol

#### Terms

$id$ = id of the sensor node,  $d$ = distance,  $RE$ = residual energy of the node,  $NE$ = node degree of the node,  $(Nei(S_i))$  = one hop neighbour of node  $S_i$ ,  $BS$ = base station,  $CH$ =cluster head,  $CH\_adv\_msg$ = Cluster head advertisement message,  $STL$ = sensor trust level,  $C$ =cluster,  $P_{ij}^{obs}$  = observation probability of node  $i$  for target  $j$ ,  $cover$ = set cover,  $LN$ = leader node,  $FN$ = next forwarding node,  $L$ = network lifetime,  $count$ = number of set covers,  $D$ = distance,  $RE$ = residual energy,  $LS$ = link strength

#### Algorithm

1. Area= $M * M$
2.  $S$ = set of nodes and  $T$ = set of targets
3. Number of Nodes= $n$
4. Number of Targets= $m$
5.  $M$ = fraction of nodes having  $\alpha$  times more energy than the normal nodes.
5. All the nodes are divided in two Regions R1 and R2.
6.  $k$ =number of clusters

#### Cluster formation

7.  $r=1$
8.  $S_i \leftarrow \{id_j, d, RE_j, NE_j\} \forall S_j \in Nei(S_i)$
9.  $CH \leftarrow S_i (NE_{max})$
10.  $CH \rightarrow S_j (CH\_adv\_msg)$
11.  $S_j \in C$  such that  $d(S_j, C)_{min}$
12. for  $r=2$  to  $k$
- $CH \leftarrow Fuzzy(D, RE, Ne)$

13. The node which does not receive the  $CH\_adv\_msg$  will be elected as  $CH$ .

**Set Cover Determination**

- 14.  $CH \leftarrow STL \forall S_i \in C$
- 15.  $P_{ia}^{obs} \forall S_i \in S$  and  $\forall T_a \in T$
- 16.  $BS \leftarrow cover(i, a) \forall T_a \in T$

**Data transmission**

- 17.  $\forall cover LN \leftarrow Fuzzy(D, RE, LS)$
- 18.  $LN \rightarrow BS$  if  $d(LN, BS) < RC$
- else
- $LN \rightarrow FN \rightarrow BS$
- 19.  $L \propto count$
- end

- 15.  $P_{ia}^{obs} \forall S_i \in S$  and  $\forall T_a \in T$
- 16.  $BS \leftarrow cover(i, a) \forall T_a \in T$

**Data transmission**

- 17.  $\forall cover LN \leftarrow Fuzzy(D, RE, LS)$
- 18.  $LN \rightarrow BS$  if  $d(LN, BS) < RC$
- else
- $LN \rightarrow FN \rightarrow BS$
- 19.  $L \propto count$
- end

Aggregation at the CH level

**Table II. Pseudo code for the aggregation at the ch level**

<ol style="list-style-type: none"> <li>1. <math>S_i \leftarrow d_j</math></li> <li>2. Select threshold <math>Th_j \forall j \in [1, c]</math></li> <li>3. <math>\forall d_j e_d = \begin{cases} 1 &amp; \text{if } d_i \geq Th_j \\ \text{else} &amp; 0 \end{cases}</math></li> </ol>
---

Aggregation at the Set Cover Level

**Table III. Pseudo code for the aggregation at the set cover level**

<ol style="list-style-type: none"> <li>1. Calculate the <math>sim(d_i, d_j) \forall i, j \in d</math> using the Jacard Similarity as:  <math display="block">sim(d_i, d_j) = \frac{ d_i \cap d_j }{ d_i \cup d_j }</math> </li> <li>2. If <math>sim(d_i, d_j) \geq sim_{th}</math> discard the data pairs.</li> </ol>
---

**Table IV. Pseudo code for the set cover calculation**

<p><b>Input:</b> set of <math>n</math> sensor nodes=<math>\{S\}</math>, set of <math>m</math> target nodes=<math>\{T\}</math>, set of required trust level for every target=<math>RCL \forall t_j \in T, \forall s_i \in S</math> where <math>i=\{1,2...n\}</math> and <math>j=\{1,2...m\}</math></p> <p><b>Terms:</b> <math>RCL</math>= required trust level for every target, <math>STL</math>= trust level of the sensor node, <math>B_i</math>=lifetime of the sensor node <math>i</math>, <math>cov(i, j)</math> = coverage probability of the node <math>i</math> detecting the target <math>j</math>, <math>P_{ij}^{obs}</math> = observation probability of sensor node <math>i</math> for the target <math>j</math>, <math>contr(i)</math>= contribution of the node and is defined as the number of targets it can monitor, <math>TAR</math>= set of targets monitored and <math>cover</math>= cover set which can monitor all the targets and <math>count</math>= number of set covers.</p> <ol style="list-style-type: none"> <li>1. Initialize <math>B_i = 1 \forall s_i \in S, i = \{1,2,..n\}</math></li> <li>2. Initialize <math>count=0</math></li> <li>3. Compute <math>cov(i, j) \forall s \in S, \forall t \in T</math></li> <li>4. Calculate <math>STL \forall s \in S</math></li> <li>5. Compute observation probability <math>P_{ij}^{obs} = cov(i, j) \times STL</math></li> <li>6. Arrange the observation probability in decreasing order.</li> <li>7. Sort the nodes according to the contribution of the nodes.</li> <li>8. Find <math>(i, j) = \arg \max(P_{ij}^{obs})</math></li> <li>9. <math>TAR = \{j\}</math></li> <li>10. If <math>cov_{er}(j) = \{i \mid P_{ij}^{obs} &gt; RCL\} : \arg \max(contr(i))</math></li> <li>11. <math>cov_{er}(j) = cov_{er}(j) \cup i</math></li> <li>12. <math>TAR = TAR \cup j</math></li> <li>13. <math>\forall k</math> such that <math>P_{kj}^{obs} &gt; RCL \mid cov_{er}(k) = \{ \}</math></li> <li>14. <math>TAR = TAR \cup k</math></li> <li>15. <math>count = count + 1</math></li> <li>16. repeat until <math>TAR = T</math></li> <li>17. update <math>B_i = B_i - e_i</math></li> <li>18. return cover</li> </ol>
---

**A. Number of Nodes required for Complete Coverage**

The coverage level in the sensor network is significantly influenced by the deployment mechanism. The nodes can be deployed in either random or deterministic manner in a network. The minimum and maximum number of nodes required for the coverage in a target region is determined for the various deployment scenarios. We have considered non overlapping deployment for the triangular, square and hexagonal deployment and the maximum and minimum coordinates of the target region are as  $(x_{min}, y_{min})$  and  $(x_{max}, y_{max})$  in the  $x$  and  $y$  dimensions respectively.



**Triangular deployment**

The cell length in this case is  $\sqrt{3} R$  where  $R$  is the sensing range. The number of rows ( $N_{row}$ ) and columns ( $N_{col}$ ) in the target region is calculated as:

$$N_{row} = \frac{x_{max} - x_{min}}{R\sqrt{3}}$$

$$N_{col} = \frac{y_{max} - y_{min}}{R\sqrt{3}}$$

$$\text{Minimum number of nodes within a cell}(n_{min}) = \frac{N_{row} \times N_{col}}{(R\sqrt{3})^2}$$

$$\text{Minimum number of nodes in the target region}(N_{min}) = N_{row} \times N_{col} \times n_{min}$$

There can be maximum 3 nodes deployed in each cell for triangular lattice so the maximum number of nodes

$$N_{max} = N_{min} \times 3$$

**Square Deployment**

The cell length in this case is  $\sqrt{2}R$  and the number of rows and columns in the region are calculated as follows:

$$N_{row} = \frac{x_{max} - x_{min}}{R\sqrt{2}}$$

$$N_{col} = \frac{y_{max} - y_{min}}{R\sqrt{2}}$$

$$\text{Minimum number of nodes within a cell}(n_{min}) = \frac{N_{row} \times N_{col}}{(R\sqrt{2})^2}$$

$$\text{Minimum number of nodes in the target region}(N_{min}) = N_{row} \times N_{col} \times n_{min}$$

There can be maximum 4 nodes deployed in each cell for square lattice, so the maximum number of nodes  $N_{max} = N_{min} \times 4$

**Hexagonal Deployment**

The cell length in this case is  $R$  and the number of rows and columns in the region are calculated as follows:

$$N_{row} = \frac{x_{max} - x_{min}}{R}$$

$$N_{col} = \frac{y_{max} - y_{min}}{R}$$

$$\text{Minimum number of nodes within a cell}(n_{min}) = \frac{N_{row} \times N_{col}}{R^2}$$

$$\text{Minimum number of nodes within the target region}(N_{min}) = N_{row} \times N_{col} \times n_{min}$$

There can be maximum of 4 nodes deployed for the hexagonal lattice, so the maximum number of nodes in the target region

$$N_{max} = N_{min} \times 7$$

**B. Determination of Number of Active Nodes for Target Cov,**

In this section we aim to determine the optimal deployment of nodes in terms of number of nodes and the deployment strategy. We have evaluated the number of active nodes required for various coverage levels ( $k= 1,2$  and  $4$ ) by considering the node deployment in various grid sizes as  $3*3, 4*4$  and  $5*5$ . The number of active nodes required for various coverage requirements depends on the target region, sensing range of the nodes and the grid pattern in which the nodes are deployed. The number of active nodes required for 1, 2 and 4 coverage requires 1, 2 and 4 nodes per grid. 1-coverage ensures that each point in the target region is monitored by at least one sensor node, 2-coverage requires that each point in the target region is monitored by at least 2 sensor nodes and 4-coverage requires that each point in the target region is monitored by 4 sensor nodes. The optimal number of active nodes is proportional to the number of targets. The optimal

node deployment can be achieved by activating the minimum number of nodes in each grid so that the energy consumption is minimized and the network lifetime can be improved.

**C. Routing**

The sensor nodes sense the environment and collect the data from the target points according to the set cover obtained in the setup phase. The node within the set covers forward its data to the node which is nearest to the base station. The data is further transmitted to the base station directly if it is reachable from the node otherwise the next forwarding node is determined based on the distance of the node from the base station, residual energy and the link strength parameters. The link strength is defined in terms of the estimated transmission count (ETX) and is defined in terms of average number of transmissions of data frames and acknowledgement frames necessary for the successful transmission of a packet. The aim of the efficient network protocol is to minimize the number of transmissions required for the transmission of number of packets.

$$\text{Link strength} \propto 1 / \text{Estimated transmission count}$$

$$\text{Estimated transmission count (ETX)} = \frac{1}{p_d \times p_{ack}}$$

where  $p_d$  represents the probability of the successful transmission of the data packets

and  $p_{ack}$  represents the successful reception of the acknowledgement.

We have used the fuzzy inference for the selection of the relay nodes. The membership function for the input parameters distance, residual energy and link strength is defined as follows:

$$\mu_A(E) = \begin{cases} 0 & \text{if } e \leq TH_1 \\ (e - TH_1) / (TH_1 - TH_2) & \text{if } TH_1 < e < TH_2 \\ 1 & \text{if } e \geq TH_2 \end{cases}$$

$$\mu_B(LS) = \begin{cases} 0 & \text{if } ls \leq TH_3 \\ (ls - TH_3) / (TH_3 - TH_4) & \text{if } TH_3 < ls < TH_4 \\ 1 & \text{if } ls \geq TH_4 \end{cases}$$

$$\mu_C(D) = \begin{cases} 1 & \text{if } d \leq TH_5 \\ (TH_5 - d) / (TH_5 - TH_6) & \text{if } TH_5 < d < TH_6 \\ 0 & \text{if } d \geq TH_6 \end{cases}$$

where  $TH_1, TH_2, TH_3, TH_4, TH_5$  and  $TH_6$  are the thresholds which determine the activation region of the fuzzy inference process. The fuzzy rule set for the next forwarding node is shown in Table V:



Table V. Rule set for next forwarding node

Rule No	Distance	Residual Energy	Link Strength	Next Forwarding Node Probability
R1	Low	Low	Less	Very Poor
R2	Low	Low	Medium	Poor
R3	Low	Low	High	Considerable
R4	Low	Medium	Less	Considerable
R5	Low	Medium	Medium	High
R6	Low	Medium	High	High
R7	Low	High	Less	Considerable
R8	Low	High	Medium	High
R9	Low	High	High	Very High
R10	Medium	Low	Less	Poor
R11	Medium	Low	Medium	Considerable
R12	Medium	Low	High	High
R13	Medium	Medium	Less	Considerable
R14	Medium	Medium	Medium	Considerable
R15	Medium	Medium	High	High
R16	Medium	High	Less	High
R17	Medium	High	Medium	High
R18	Medium	High	High	Very High
R19	High	Low	Less	Poor
R20	High	Low	Medium	Considerable
R21	High	Low	High	High
R22	High	Medium	Less	Considerable
R23	High	Medium	Medium	High
R24	High	Medium	High	Very High
R25	High	High	Less	High
R26	High	High	Medium	High
R27	High	High	High	Considerable

**D. Aggregation**

To further improve the network efficiency we have incorporated the aggregation mechanism based on the locality sensitive hashing. We have considered the aggregation at two levels in the proposed protocol:

1. At the cluster head level and
2. At the set cover level.

The objective of the aggregation at the cluster head level is to reduce the number of bits required to transmit the data and at the set cover level the number of packet transmissions is reduced by using the aggregation.

*Locality Sensitive Hashing (LSH)*

LSH represents the similarities between the objects using the probability distribution over the hash function. Hash collision is used to determine the object similarity. The basic concept of LSH is to project the data into low dimensional binary (Hamming) space; such that each data is mapped to a *b* bit vector called hash key. Each hash function must satisfy the LSH property  $\Pr[(h(x_i) = h(x_j))] = sim(x_i, x_j)$ , where  $sim(x_i, x_j) \in [0,1]$  is the similarity function of interest

[16].

The hashing function to generate the Hash code is:

$$h_r(x) = \begin{cases} 1, & \text{if } r^T x \geq 0 \\ 0, & \text{otherwise} \end{cases} \text{ where } r^T(x) \geq 0 \text{ is a}$$

hyperplane separating the space.

After generating the Hash codes the similarity measure between the every data pair is determined using the Jacard similarity metric. The Jacard similarity between two data sets *S* and *T* is defined as:

$$sim(S, T) = \frac{S \cap T}{S \cup T}$$

**E. Performance Metrics**

We have considered several metrics to evaluate the performance of the proposed protocol as:

*i. Number of Set Covers*

It defines the number of the set covers consisting of the optimal number of nodes which can monitor all the targets. The network lifetime is proportional to the number of set covers obtained. The higher the number of set covers higher will be the network lifetime.

*ii. Number of Iterations*

The number of iterations defined the number of rounds such that all the targets are monitored with the required confidence level.

*iii. Number of Un Used Nodes*

This metric represents the number of nodes which do not contribute in the set cover. The lower the number of un used nodes in every iteration, higher will be the network efficiency.

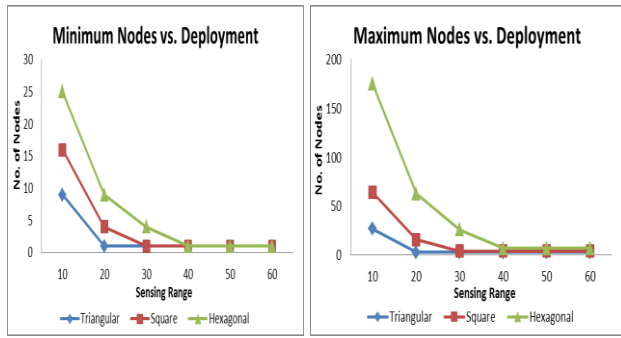
*iv. Node Contribution*

Node contribution defines the number of targets a node can monitor. The higher the number of targets a node can monitor with required coverage level, higher will be its contribution. In the set cover formation, nodes with higher contribution are considered first so as to achieve the efficient energy utilization and maximum lifetime.

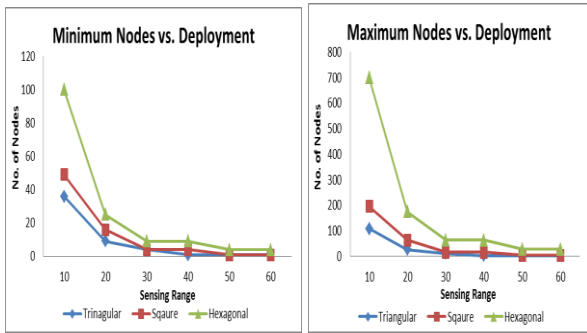
**IV. RESULTS AND ANALYSIS**

We have obtained the minimum and maximum number of nodes for the target regions of the dimensions 50\*50, 100\*100, 200\*200, 300\*300. We have obtained the results for the triangular, square and hexagonal deployment by varying the sensing range of the nodes from 10 to 60 with the increment of 10. The results are shown in Fig.3. The results show the exponential decay in the number of nodes with the increase in the sensing range. It can be observed from the results that there is large number of nodes for the hexagonal deployment; hence the effective area covered is increased with the increase in the data acquisition process. There is maximum number of nodes within a cell which are connected in the logical hexagon, so direct trust will have the higher impact for the hexagonal lattice in comparison to the triangular and square lattices.

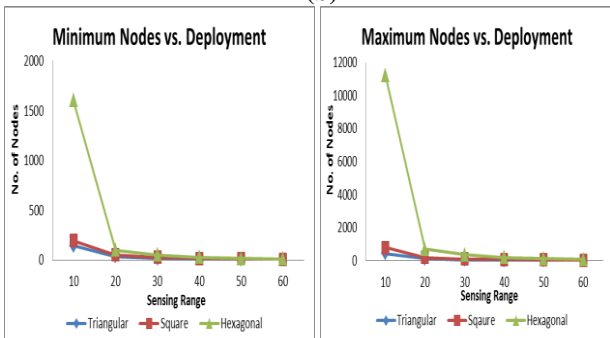




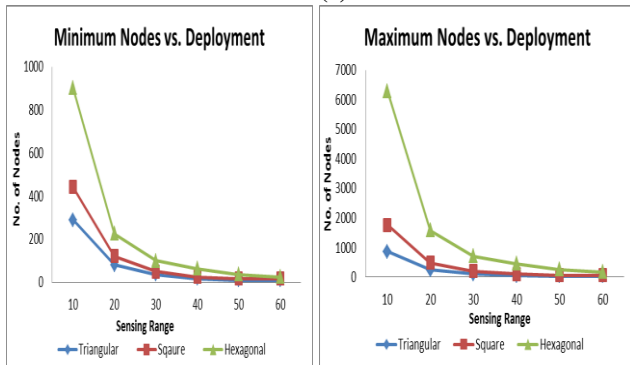
(a)



(b)



(c)

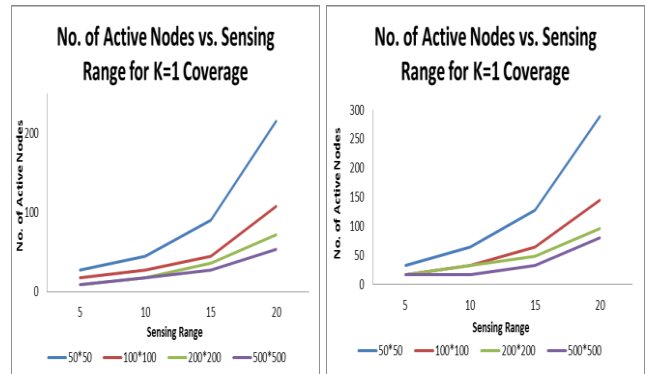


(d)

**Fig.3. Minimum and Maximum nodes obtained for the triangular, square and hexagonal non-overlapping deployment for the region of the dimensions: a. 50\*50, b. 100\*100, c. 200\*200 d. 300\*300.**

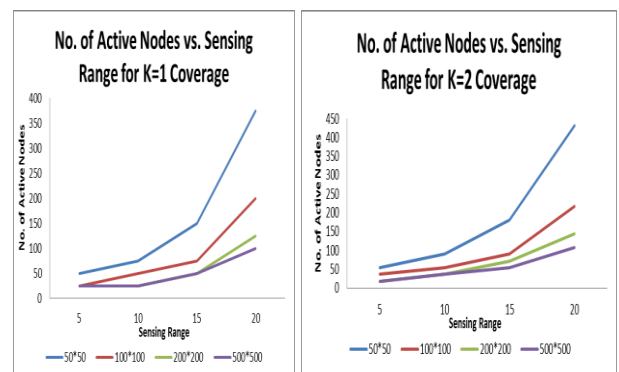
We have determined the number of active nodes required for target coverage for various target regions of dimensions 50\*500, 100\*100, 200\*200 and 500\*500. In the first experiment we have considered the deployment strategy of the varying grid sizes 3\*3, 4\*4 and 5\*5 for 1 coverage, 2-coverage and 4 coverage. The results are obtained for the different sensing ranges varying from 5 to 20 with an increment of 5 as shown in the Fig 4. , where a,b and c shows the results for the grid size 3\*3, 4\*4 and 5\*5. The results

shown that the number of active nodes is optimal for the grid size 3\*3 as compared to 4\*4 and 5\*5. The results shows that the number of active nodes required increases linearly with the increase in the sensing range.



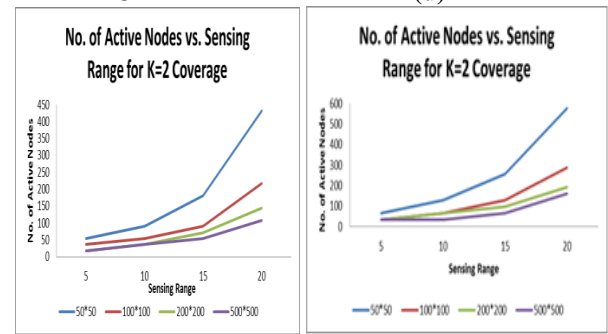
(a)

(b)



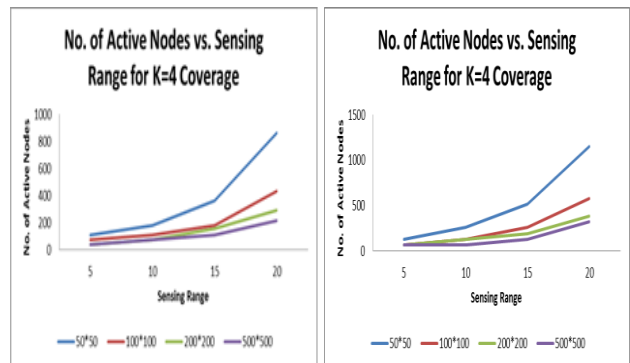
(c)

(d)



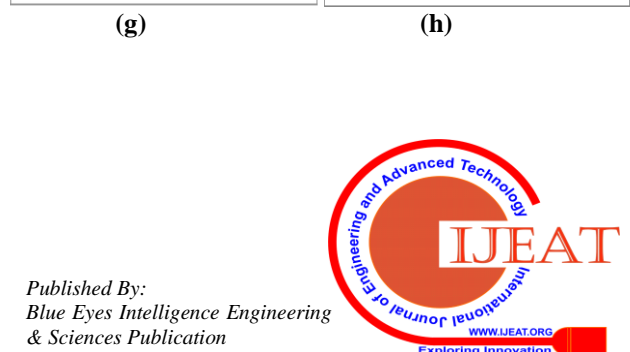
(e)

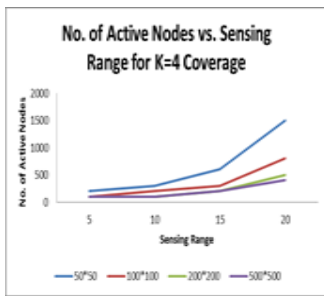
(f)



(g)

(h)

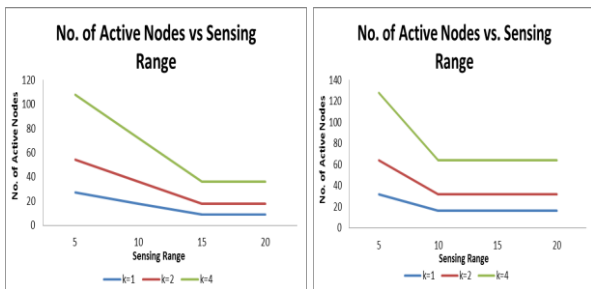




(i)

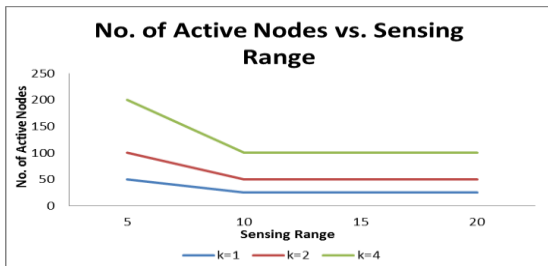
Fig.4. Number of active nodes for various coverage levels.

The number of active nodes for the various grid sizes are shown in the Fig.5. It can be observed from the results that the number of active nodes increases for the higher coverage levels.



(a)

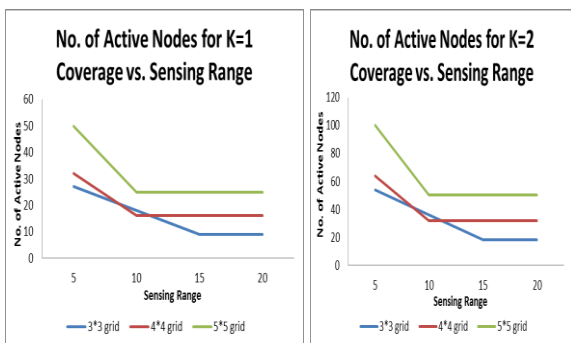
(b)



(c)

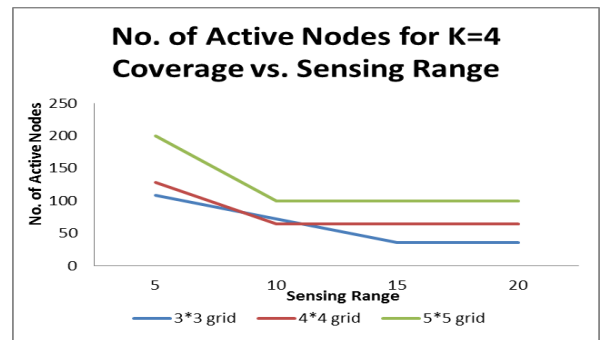
Fig. 5. Number of active nodes vs. sensing range for various coverage levels for 50\*50 region: a.3\*3 grid, b. 4\*4 grid, c.5\*5 grid

In the next experiment we have determined the optimal number of active nodes for various grid sizes and target region for the coverage level as k=1,2 and 4 as shown in the Fig.6. The results shows that the 3\*3 grid size requires the minimum number of active nodes for coverage level 1,2 and 4. The results shows that the number of active nodes increases linearly with the increase in the sensing range.



(a)

(b)



(c)

Fig.6. Number of active nodes vs. sensing range for different grid sizes: a. K=1 coverage, b. K=2 coverage, c. K=4 coverage.

For a network of 30 sensor nodes and 5 targets, we have obtained the 20 set covers as follows:

Table VI. Set covers obtained for the considered network of 30 nodes and 5 targets

Set Cover	Node Set	Set Cover	Node Set
C1	{10,14,19,23}	C11	{11,18,23,26}
C2	{10,14,23,26}	C12	{11,18,23,29}
C3	{10,14,23,29}	C13	{7,14,23,26}
C4	{10,18,19,23}	C14	{7,18,23,26}
C5	{10,18,23,26}	C15	{10,14,20,23}
C6	{10,18,23,29}	C16	{10,18,20,23}
C7	{11,14,19,23}	C17	{11,14,20,23}
C8	{11,14,23,26}	C18	{11,18,20,23}
C9	{11,14,23,29}	C19	{7,14,23,26}
C10	{11,18,19,23}	C20	{7,18,23,26}

The number of targets monitored by each node is defined as node contribution. The contribution of each node for target coverage is shown in Fig.7. as follows:

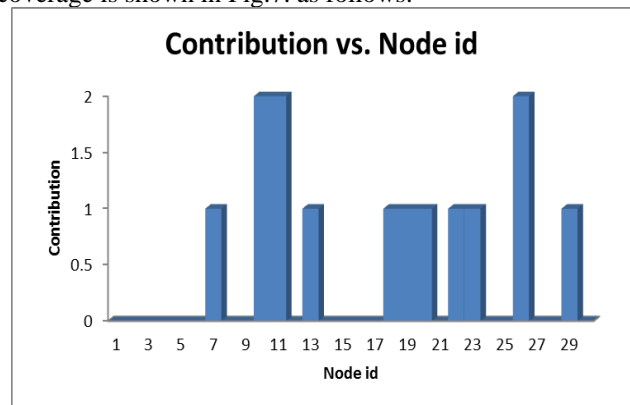


Fig.7. Contribution of each node for the target coverage

Consider the set cover C10={10,14,19,23} node 10 collects all the data within the set cover as it is the nearest node to the base station. Node 10 has neighbors as {9,22,24}. The various threshold values are as:  $TH_1 = 0.1, TH_2 = 0.7, TH_3 = 20, TH_4 = 30, TH_5 = 0.4, TH_6 = 0.7$





To determine the next forwarding node we apply the fuzzy inference process as follows:

**Table VIII. Readings sensed by the nodes**

S. No.	A	B	C
1	15	40	20
2	20	45	25
3	23	47	30
4	30	50	45
5	18	56	70
6	28	60	55
7	26	75	60
8	48	82	80
9	50	86	70
10	44	85	35

**Table IX. Four dimensional reading sensed by the nodes**

S. No.	A	B	C	D
1	15	40	20	30
2	20	45	25	35
3	23	47	30	40
4	30	50	45	55
5	18	56	70	80
6	28	60	55	60
7	26	75	60	56
8	48	82	80	38
9	50	86	70	42
10	44	85	35	37

**Table X. Five dimensional readings sensed by the nodes**

S. No.	A	B	C	D	E
1	15	40	20	30	50
2	20	45	25	35	38
3	23	47	30	40	60
4	30	50	45	55	40
5	18	56	70	80	65
6	28	60	55	60	47
7	26	75	60	56	32
8	48	82	80	38	45
9	50	86	70	42	70
10	44	85	35	37	59

After the projection of high dimension with respect to the hyperplane space the encoded data set are as:

**Table XI. Encoded data sets**

S. No.	3 dimensional data set	4dimensional data set	5dimensional data set
1	000	0000	00000
2	000	0000	00000
3	000	0000	00001
4	000	0001	00010
5	001	0011	00111
6	001	0011	00110
7	011	0111	01110
8	111	1110	11100
9	111	1110	11101
10	110	1100	11000

On performing the pair wise comparison on each data set the similarity measure compared are as follows:

**Table XII. Similarity Matrix For 3 Dimensional Data**

Data sets	000	000	000	000	001	001	011	111	111	110
000	1	1	1	1	2/3	1/3	1/3	0	0	1/3
000	1	1	1	1	2/3	1/3	1/3	0	0	1/3
000	1	1	1	1	2/3	1/3	1/3	0	0	1/3
000	1	1	1	1	2/3	1/3	1/3	0	0	1/3
001	2/3	2/3	2/3	2/3	1	1	2/3	1/3	1/3	0
001	2/3	2/3	2/3	2/3	1	1	2/3	1/3	1/3	0
011	1/3	1/3	1/3	1/3	2/3	2/3	1	2/3	2/3	1/3
111	0	0	0	0	1/3	1/3	1/3	1	1	2/3
111	0	0	0	0	1/3	1/3	1/3	1	1	2/3
110	1/3	1/3	1/3	1/3	0	0	1/3	1/3	1/3	1

We have considered the similarity threshold as 0.5. So we discard the data sets having the higher similarity measures. The number of redundant transmissions is 46. The number of bits required to transmit the information for the 3 dimensional data set is 480 bits but after applying the aggregation techniques the number of bits required is 162 bits. Similarly we have calculated the number of bits required and the number of message transmissions before and after the aggregation for the 4 and 5 dimensional data sets as shown in the tables below:

**Table XIII. The number of bits required before and after the aggregation for each data set**

	3 Dimensional Dataset	4 Dimensional Dataset	5 Dimensional Dataset
Before Aggregation	480	640	800
After Aggregation	162	164	135

**Table XIV. The number of message transmission before and after the aggregation for each data set**

	3 Dimensional Dataset	4 Dimensional Dataset	5 Dimensional Dataset
Before Aggregation	100	100	100
After Aggregation	54	41	45



The results of the aggregation approaches are shown in the Fig.8. The result shows that the number of bits required to transmit the data and the number of transmissions required to transmit the data packets is significantly reduced by employing the proposed aggregation approach.

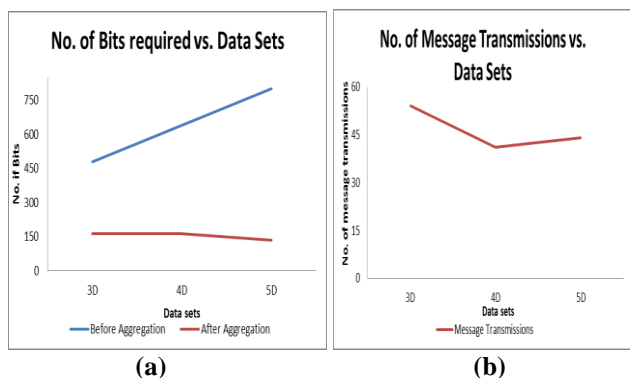


Fig.8. Performance of the aggregation approach: a. No. of bits required to transmit the information. b. No. of bits required for the message transmissions.

We have compared the performance of the proposed protocol with the Boolean coverage model for target coverage on the basis of several performance metrics by varying the sensing range as shown in the Fig.9. The results show that the proposed protocol achieves an improvement in the network lifetime. As the sensing range increases, the number of set covers formed is higher for the proposed protocol as compared to the Boolean coverage model. The incorporation of the probabilistic model and trust concept improves the number of set covers obtained. The results shows that the proposed protocol executes for the larger number of iterations hence can extend the network lifetime. The node utilization for the proposed protocol is better than the Boolean coverage model.

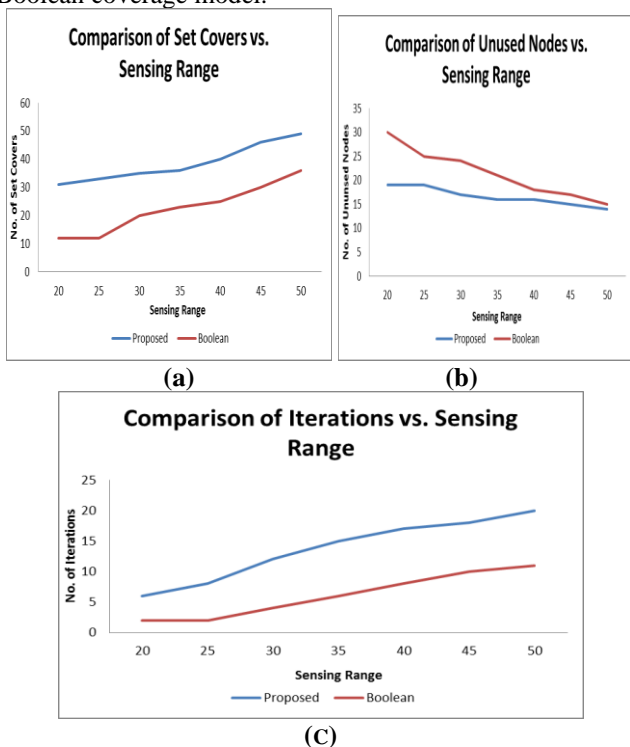


Fig.9. comparison of the proposed protocol with Boolean coverage model: a. no. of set covers, b. no. of un used nodes, c. no. of iterations

## V. CONCLUSION AND FUTURE SCOPE

The proposed protocol is trust based energy efficient node scheduling protocol for target coverage which considers the energy scarcity constraint. The proposed protocol emphasized on maximization of the network lifetime by reducing the redundancy in the sensed data and reducing the number of transmissions by using the LSH based Jaccard similarity measure. We have determined the minimum and maximum number of nodes in the various deployment scenarios. The result shows that the hexagonal deployment utilizes node efficiently as compared to the triangular and square deployment, whereas the number of active node is least in the case of triangular deployment. The hybrid routing is proposed which aims to provide the reliable data transmission towards the base station. The results show that the proposed protocol improves the network efficiency in terms of coverage, reliability, overhead and network lifetime as compared to the Boolean coverage model.

The network lifetime can be further extended by using the mobility and redeployment of the unused nodes. To study the incorporation of mobility strategy for the redeployment of the unused nodes is our future work.

## REFERENCES

1. Pooja Chaturvedi, A. K. Daniel, "Recovery of Holes Problem in Wireless Sensor Networks", IEEE 4th International Conference on Information Communication & Embedded Systems (ICICES-2014), DOI: 10.1109/ICICES.2014.7034172, 2014.
2. Pooja Chaturvedi, A. K. Daniel, "An Energy Efficient Node Scheduling Protocol for Target Coverage in Wireless Sensor Networks" in the 5<sup>th</sup> International Conference on Communication System and Network Technologies (CSNT-2015), April 2015.
3. Pooja Chaturvedi, A. K. Daniel "Trust Based Node Scheduling Protocol for Target Coverage in Wireless Sensor Networks", 3<sup>rd</sup> International Conference on Emerging Research in Computing, Information, Communication & Applications (ERCICA-2015), Bangalore, Springer India 2015, N.R. Shetty et al. (eds.), Emerging Research in Computing, Information, Communication and Applications, DOI 10.1007/978-81-322-2550-8\_16.
4. Pooja Chaturvedi, A. K. Daniel, "Trust Based Target Coverage Protocol for Wireless Sensor Networks Using Fuzzy Logic", in the 12<sup>th</sup> International Conference on Distributed Computing & Internet Technology (ICDCIT-2016), organized by KIIT University, Bhubaneswar, Odisha, India, during the period 15<sup>th</sup> Jan 2016-18<sup>th</sup> Jan 2016.
5. Pooja Chaturvedi, A. K. Daniel, "Lifetime Optimization for Target Coverage in Wireless Sensor Networks" in 8<sup>th</sup> Annual ACM India Conference Compute 2015, organized by Academy of Business & Engineering Sciences Engineering College (ABESEC), Ghaziabad during the period 29<sup>th</sup> Oct 2015-31<sup>st</sup> Oct 2015.
6. Pooja Chaturvedi, A. K. Daniel, "Trust Based Energy Efficient Coverage Preserving Protocol for Wireless Sensor Networks", in the International Conference on Green Computing & Internet of Things (ICGIOT-2015).
7. N. Jaggi, and A. A. Abouzeid, Energy-Efficient Connected Coverage in Wireless Sensor Networks, Proc. 4th Asian International Mobile Computing Conference (AMOC), pp. 77-86, January 2006 .
8. XiaochunXu and SartajSahni, Approximation Algorithms For Wireless Sensor Deployment, April 21, 2006.
9. Z. Zhou, S. Das and H. Gupta. Connected k-coverage problem in sensor networks. In Proc. of International Conference on Computer Communications and Networks (ICCCN'04), Chicago, IL, October 2004, pp. 373-378.
10. Wang, J.; Zhong, N. Efficient point coverage in wireless sensor networks. J. Comb. Optim. 2006, 11, 291-304.



11. Bhaskar Krishnamachari, Deborah Estrin, Stephen Wicker, "Impact of Data Aggregation in Wireless Sensor Networks", DEBS, 2002.
12. Nandini. S. Patil, Prof. P. R. Patil, "Data aggregation in Wireless Sensor Network," IEEE International Conference on Computational Intelligence and Computing Research.
13. Avid Avokh, Ghasem Mirjaliliy Patil, "Dynamic Balanced Spanning Tree(DBST) for Data Aggregation in Wireless Sensor Networks", 5th International Symposium on Telecommunications (IST2010), 978-1-4244-8185-9/10, 2010.
14. Prakash Goud Patil, Umakant Kulkarni, "SVM based Data Redundancy Elimination for Data Aggregation in Wireless Sensor Networks, 2013.
15. Basavaraj S. Mathapati, Siddarama. R. Patil, "Energy Efficient Reliable Data Aggregation Technique for Wireless Sensor Network", International Conference on Computing Sciences, 2012.
16. M. Datar, N. Immorlica, P. Indyk, and V. Mirrokni. Locality-Sensitive Hashing Scheme Based on p-Stable Distributions. In *SOCG, 2004*.
17. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless Sensor Networks: A Survey," *Computer Networks*, Vol. 38, No. 4, 2002, pp. 393-422.
18. J. Yick, B. Mukherjee, D. Ghoshal, *Wireless Sensor Networks Survey*, *Computer Networks* 52 (2008) 2292–2330.
19. Raymond Mulligan, Habib M. Ammari, *Coverage in Wireless Sensor Networks: A Survey*, *Network Protocols and Algorithms*, ISSN 1943-3581, 2010, Vol. 2, No. 2.
20. Q. Gao, K. J. Blow, D. J. Holding, I. W. Marshall, and X. H. Peng, "Radio range adjustment for energy efficient wireless sensor networks," *Ad Hoc Networks*, vol. 4, no. 1, pp. 75–82, 2006.
21. V. Mhatre and C. Rosenberg, "Design guidelines for wireless sensor networks: communication, clustering and aggregation," *Ad Hoc Networks*, vol. 2, no. 1, pp. 45–63, 2004.
22. S. R. T. Zheng and V. Sarangan, "PMAC: an adaptive energy efficient MAC protocol for wireless sensor networks," in *Proceedings of the Parallel and Distributed Processing Symposium*, pp. 65–72, April 2005.
23. Z. Zhou, S. Das and H. Gupta. Connected k-coverage problem in sensor networks. In *Proc. Of International Conference on Computer Communications and Networks (ICCCN'04)*, Chicago, IL, October 2004, pp. 373–378.
24. Wang, J.; Zhong, N. Efficient point coverage in wireless sensor networks. *J. Comb. Optim.* 2006, 11, 291–304
25. Hai Liu, Pengjun Wan, XiaohuaJia, Maximal lifetime scheduling for K to 1 sensor–target surveillance networks, *Computer Networks* 50 (2006) 2839–2854.
26. S. Mini, Siba K. Udgate, and Samrat L. Sabat, *Sensor Deployment and Scheduling for Target Coverage Problem in Wireless Sensor Networks*, *IEEE SENSORS JOURNAL*, VOL. 14, NO. 3, MARCH 2014.
27. My T. Thai Feng Wang Ding-Zhu Du, *Coverage Problems in Wireless Sensor Networks: Designs and Analysis*, *International Journal of Sensor Networks*, Vol. 3, Issue 3, May 2008, pp. 191-200.
28. Zahra Taghikhaki, Nirvana Meratnia, Paul J.M. Havinga, "A Trust-based Probabilistic Coverage Algorithm for Wireless Sensor Networks", 2013 International Workshop on Communications and Sensor Networks (ComSense-2013), *Procedia, Computer Science* 21 (2013), 455 – 464.
29. Jinfang Jiang, Guangjie Han, Feng Wang, Lei Shu, Mohsen Guizani, "An Efficient Distributed Trust Model for Wireless Sensor Networks", *IEEE Transactions on Parallel and Distributed Systems*, 2014.
30. Hyo Sang Lim, Yang Sae Moon, Elisa Bertino, *Provenance based Trustworthiness Assessment in Sensor Networks*, *DMSN'10*, September 13, 2010, Singapore.
31. Saaty, T. L., 1990, How to make a decision. *The Analytical Hierarchy Process*. *European Journal of Operational Research*, 48, 9–26.
32. Heinzelman, W. R., Chandrakasan, A. & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks, *HICSS '00: Proceedings of the 33<sup>rd</sup> Hawaii International Conference on System Sciences-Volume 8*, IEEE Computer Society, Washington, DC, USA, p. 8020.