

# A Virtual Grid Based Energy Efficient Data Gathering Scheme for Heterogeneous Sensor Networks

Siddhartha Chauhan, Nitin Kumar Kotania

**Abstract**—Traditional Wireless Sensor Networks (WSNs) generally use static sinks to collect data from the sensor nodes via multiple forwarding. Therefore, network suffers with some problems like long message relay time, bottle neck problem which reduces the performance of the network.

Many approaches have been proposed to prevent this problem with the help of mobile sink to collect the data from the sensor nodes, but these approaches still suffer from the buffer overflow problem due to limited memory size of sensor nodes. This paper proposes an energy efficient scheme for data gathering which overcomes the buffer overflow problem. The proposed scheme creates virtual grid structure of heterogeneous nodes. Scheme has been designed for sensor nodes having variable sensing rate. Every node finds out its buffer overflow time and on the basis of this cluster heads are elected. A controlled traversing approach is used by the proposed scheme in order to transmit data to sink. The effectiveness of the proposed scheme is verified by simulation.

**Keywords**—Buffer overflow problem, Mobile sink, Virtual grid, Wireless sensor networks.

## I. INTRODUCTION

WIRELESS Sensor Network is one of the most important technologies in the twenty first century. It plays an important role in a wide range of applications such as, military areas, environmental monitoring (indoor and outdoor monitoring), human-centric (health Monitoring), in robotics [1]. It consists of large number of tiny sensor nodes can be used as an effective tool for gathering data in various situations. At this time with the available technologies sensor nodes are integrated with small amount of storage space, limited battery powered, limited onboard processing power and wireless communication capabilities.

In the traditional wireless sensor networks data is collected at static sink with the help of the intermediate nodes between source sensor node and the static sink. In this the intermediate nodes forward the source sensor nodes data to static sink. In other words, a node not only transmits the information sensed by it but also relays the data packets generated by others. Due to this nodes near the sinks have a tendency to consume more energy because they may have the overhead of relaying the data of a sensor node which is at a distance from the sink. This causes hot spot problem [2] which may exhaust the nodes nearby the sink. And then whole network goes down because the sink is no more in connection with rest of the network.

Dr. Siddhartha Chauhan is Assistant Professor in National Institute of Technology, Hamirpur, Himachal Pradesh, India (e-mail: siddharthachauhan1@gmail.com).

Nitin Kumar Kotania is with National Institute of Technology, Hamirpur, Himachal Pradesh, INDIA (e-mail: nitin.kotania@gmail.com).

And it is not easy to replace or recharge the battery of a sensor node. A number of researchers have proposed mobility as a solution to this problem of data gathering. For this whole network divided into clusters and there is a cluster head node (rendezvous point) of all the clusters. The rendezvous point gathers the data of all the nodes of its cluster. So the mobile sink comes to this point and collects data from it. This resolves the hot spot problem up to some level because now the hot spot problem is distributed form a static sink to cluster heads of respective cluster heads which results a distributed hot spot problem.

Various types of mobility have been considered for the mobile node (sink). These can be broadly classified as random, predictable or controlled. In the work on data mules [3], the mobile nodes called “data mules” moves randomly.

In most of the cases the sensor nodes are battery constraints which make the problem of energy efficient paramount importance. But in some scenarios where a sensor network is used to sense pollution levels at strategic locations in a large city so it is natural that there will be some area in which the pollution level is different from other areas like, industrial area pollution level is more than the pollution level of the residential areas. For this scenario, the sensing rate of the sensor at different positions will typically need to be different. The sensor node with higher variation in the phenomenon need to sample more frequently.

This paper tries to address the energy consumption problem of data gathering and reduce the buffer overflow problem with the help of the sharing the memory of the 1-hop neighbors of the Cluster Head of the cluster. As we know traditional sensor network suffers from hot spot problem. To solve this many researchers used mobile sink to collect the data but that is a time consuming process because in this mobile node moves to each sensor node to collect the data. So now researchers use clustering to avoid this time consuming process i.e. clusters. In this each cluster has a cluster head which gather the whole data to the cluster and transfer it to the mobile sink but this leads to another problem that we call distributed hot spot problem because the cluster head is fixed and every time the same node is performing the data gathering and forwarding it to mobile sink. In this paper the cluster head selection is dynamic and the mobile sink traverse path is also dynamic which is decided on the basis of buffer overflow time. In this the cluster head selection and traverse path will be select on basis of buffer overflow metric. So here traversing is also a big issue. In this, mobile sink have to collect data from the Cluster Heads before their deadline time. It is quite similar to the TSP (Traveling Sales Man Problem) [4] because both are

traversing the nodes on the basis of the given deadlines. But it is important to clearly outline the difference between these two. In TSP, the goal is to find a minimum cost tour that visits each node exactly once. However in this paper approach, a node may need to be visited multiple times before all other nodes are visited depending on the strictness of its deadline i.e. sampling rate ( buffer overflow time).

## II.RELATED WORK

There are many researchers worked over this area to make an energy efficient scheme for a wireless sensor network. There are two kinds of clustering schemes. The clustering algorithms applied in homogeneous networks are called homogeneous schemes, and the clustering algorithms applied in heterogeneous networks are called heterogeneous clustering schemes. Thus most of the clustering algorithms are homogeneous schemes, such as LEACH and PEGASIS.

The LEACH [5] protocol selects cluster heads periodically and drains energy uniformly by role rotation. Each node decides itself whether or not a cluster heads distributed by a probability. And under the homogeneous network it performs well but in heterogeneous networks its performance is not that much good. And in PEGASIS [6], nodes will be organized to form a chain, which can be computed by each node or by the base station. But implementation of this protocol required a global knowledge of the network which is difficult.

In NeBUST(Protocol for Neighbor Buffering for congested Sensor Data Transmission) [7] is an approach to overcome the buffer overflow problem. In this the node never pass forward to the next hop node in the multi hop transmission but it select the node which has the sufficient message buffer nearer to the sink. But it does not provide the permanent solution for the buffer overflow problem.

But the value of  $k$  restricted up to a constant value. If the  $k$  is large so Somasundara et al. [8] propose an approach in which the MS calculates a weighted value and uses the Minimum Weighted Sum Value First (MWSF) algorithm to determine the next destination. The weighted value consists of two factors, cost and deadline, in which the cost is the distance between the MS and the potential destination node and the deadline is the time to buffer overflow for the potential destination node.

Ma and Yang [9] introduced an energy-efficient data gathering scheme with an MS called SenCar. In this paper researcher are traversing the network with the help of BISECTION method. The straight line traversing is not an energy efficient traversing. So this paper allow sink to take turn in the middle of the path. Due to the more turning points may reduce the number of hops for data forwarding, which in turn reduces energy consumption but it will increase the traverse distance and now the Mobile Sink will required more time to traverse the whole ROI, which may produce the buffer overflow problem.

In TTDD (Two Tier Data Dissemination) [11], a source node creates an entire grid. When the event of the source is triggered, the source node will send this information from its nearest dissemination node, which is called the immediate

dissemination node, and thus the other ordinary nodes which locates near the dissemination area will continue to relay this message until the entire network is aware of this incident, so that all the dissemination nodes of the grid could record the information of this event sent from the source node. To locate the sink with in a grid or boundary of any two grids, there must be a dissemination node around the sink. When a sink wants data, it only needs to do the flooding. A reverse transmission path is built back to the sinks based on data trace after the sink sent queries, so the overhead for controlling the transmission and keeping track of the sink is limited to the local grid. TTDD carry most traffic that fix on the virtual grid and caused intensive energy consumption on the dissemination path.

This paper considers heterogeneous WSNs, in which sensor nodes have different data sampling rates and, thus, different levels of severity of the buffer overflow problem.

## III.PROPOSED SCHEME

There are many ways to reduce the communication energy consumption. Mobile Sink is one of them that is most efficient option and most popular in the researchers. Our proposed scheme namely Location Based Clustering algorithm (LBC) divides the whole network in a small (fixed size) sections and selects a dissemination node (cluster head) in every cluster through which the mobile sink collects the data of whole cluster. Mobile sink collects the data from the dissemination nodes in such a way that buffer overflow deadline of the dissemination nodes are met.

The proposed scheme has been designed for the sensor nodes that have different sampling rates; therefore they have different buffer overflow deadlines. It is assumed that nodes are location aware, through GPS system. All the nodes know their cluster size, which is predefined for all nodes during initialization. The Mobile sink will traverse the cluster heads according to Minimum Weighted Sum First Algorithm (MWSF) [8]. The mobile sink will traverse the dissemination nodes only according to [8], in order to gather the data.

## IV.LOCATION BASED CLUSTERING

This Location Based Clustering (LBC) algorithm run by each node deployed over a pre-defined area. There are some assumptions for this algorithm is as follows:

1. Every node and Mobile Sink knows the size of the whole network in terms of length and breadth.
2. Every node and Mobile Sink knows the size of the cluster that is fixed. Variables used for the size of the cluster are  $X_s$  and  $Y_s$ .
3. Mobile Sink knows the dedicated region for the first cluster head of every cluster.

### A. ALGORITHM: LBC

In this algorithm in time  $T_{Selection}$  cluster head will be selected by the nodes only. And for time  $t_1$  nodes get to know their location and cluster\_id and draw a region A.

1. Every node first find out its location with the help of GPS (Global Positioning System) device.
2. Then every node find its Cluster\_id.

$$X_{id} = \left[ \begin{array}{c} X \\ X_s \end{array} \right]$$

$$Y_{id} = \left[ \begin{array}{c} Y \\ Y_s \end{array} \right]$$

3. Find Region\_A ( )

If the node present inside the region A then the node will be the cluster head.

4. If

$$(((X \geq X_1) \&\&(X \leq Y_1)) \&\&((Y \geq Y_1) \&\&(Y \leq Y_1)))$$

Then node i is in cluster head selection race. The node which is closer to the center  $(X_c, Y_c)$  will be the cluster head.

Else

Wait for a time  $t_2$

5. Set  $T=t_2$

While  $T \neq 0$  do

T--

End while

If (node[i] which is present in area A receive no declaration message form any node of area A)

{  
then declare itself as a cluster head and broadcast a declaration message to rest of the node of its cluster.  
}

Else

{  
if( $node\_d\_to\_c[i] < node\_d\_to\_c[j]$ )  
{  
then Reply a message with its distance to node j and node j withdraw, broadcast the node i's message.  
}

}

Reply with a withdraw message tonode j.  
}

6. After this Cluster Head broadcast a Head Declaration message to its cluster nodes.

7. Now after  $t_3$  Mobile Sink start to collect data from the Cluster Heads.

8. And at the end this whole process in time

$$T_{selection} = t_1 + t_2 + t_3$$

After this process the network is virtually divided into clusters and every cluster knows its cluster head for the first round. So in the algorithm LBC first every node virtually set its region or make a virtual cluster. Every node first finds their

*cluster\_id* with the help of two variables which are initialized in every node at the time of deployment. So whenever a node transmits message for anything it append its *cluster\_id* in message. With the help of the *cluster\_id* the virtual grid formation is possible the cluster formation is shown in Fig. 1.

Let consider 5 node which can be anywhere in the network and node\_1, node\_2, node\_3, node\_4, node\_4 and there coordinates are (10, 15), (15, 15), (10, 20), (20, 20), (20, 15) respectively. And the  $X_s$  and  $Y_s$  values are 25, 25 respectively. The size of the network is  $100 * 100$ .

Now all five nodes find their locations (X, Y) with the help of GPS device and then with the help step 2 as shown in Algorithm LBC they find their Cluster\_id. So here in this case all the nodes have the same cluster id (1, 1). With these results every node has the same Cluster\_id. Rests of the nodes perform the same scenario and make virtual clusters. And in the similar way Mobile Node use the location of the Cluster Head to identify the Cluster\_id.

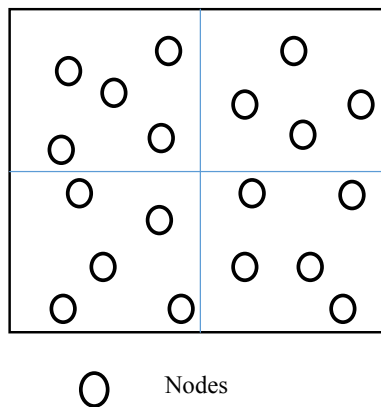


Fig. 1 Clustering with help of algorithm LBC

There is one function *Region\_A* mention in algorithm LBC is a region form where the first Cluster Head will be select. With the help of *Region\_A* function, this paper restricts the nodes which contribute for the Cluster Head selection for the first round. Due to this few numbers of nodes use their transmitter for selection so it saves some sort of energy of some nodes. And it will be helpful or the mobile sink to find out the first Cluster Head. And the function *Region\_A* as shown below and also showed in Fig. 2.

Function: *Region\_A* ( )

1. Draw a virtual circle from the center of the cluster of radius  $R_{ms}$  of Mobile Sink.
2. Find the center coordinates of the cluster.

$$X_c = (X_{id} \times X_s) - \frac{X_s}{2}$$

$$Y_c = (Y_{id} \times Y_s) - \frac{Y_s}{2}$$

3. First find the variable d, EB, AE, BC and AD values.

4. Find the region A coordinates A, B, C, and D.

$$A = (X_1, Y_1)$$

$$B = (X_2, Y_1)$$

$$C = (X_2, Y_2)$$

$$D = (X_1, Y_2)$$

$$\text{find}(r)$$

$$r = \frac{R_{ms}}{\sqrt{2}}$$

$$X_1 = X_c - r$$

$$X_2 = X_c + r$$

$$Y_1 = Y_c$$

$$Y_2 = Y_c - r$$

As shown in Fig. 2, first we find the center of the cluster, denoted as  $E$  and from the center draw a circle with a radius equal to the range of the Mobile Sink  $R_{ms}$ . In Fig. 2 paper considers a cluster with a constant size (25, 25). Then draw a rectangle (ABCD) in which one edge of rectangle passes through the center of cluster as shown in Fig. 2. And the length and the breadth of the rectangle are  $l$  and  $b$  respectively. Then find the values of the some variable  $l$  and  $r$ . With the help these variables values find the coordinates for A, B, C and D.

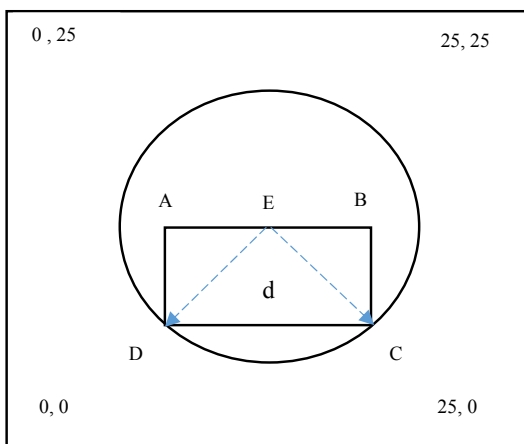


Fig. 2 Cluster head selection in a Region

As shown in Fig. 2, DEC forms a right angle triangle.

$$EC = ED = R_{ms}$$

$$DC = d$$

$$AE = BE = AD = BC = r$$

$$r = \frac{d}{2}$$

By Pythagoras theorem in right angle  $\triangle DCE$

$$d^2 = (R_{ms})^2 + (R_{ms})^2$$

$$d = \sqrt{2}R_{ms}$$

$$r = \frac{d}{2} = \frac{R_{ms}}{\sqrt{2}}$$

Now the node which is closer to the center will be the cluster head according to the *Location Based Clustering* algorithm. And it broadcast its presence to the rest of the nodes of the cluster with a broadcasting declaration message which contains some fields such as, Cluster\_id, Coordinates, distance to the center. Then after the  $T_{move}$  time the mobile sink starts and traverse all Cluster Heads and collect the useful information. And the current Cluster Head also inform about the Cluster Head for the next round.

After election of the cluster head for the first round proposed approach find the one hope neighboring nodes to share the memory resource. With the help of this paper tries to reduce the buffer overflow problem. The node with the least buffer overflow time is elected as a cluster head for the next cluster head selection rounds. Current cluster head has the responsibility to select the new cluster head for the next round and this calculated on the basis of the buffer overflow time. The node which are in the cluster other than cluster head send their buffer overflow time to the cluster and now cluster head arrange them in an order and select first k one hope neighboring nodes for sharing the memory. The larger k allows more neighbors to share their memory at the cost of greater data management and communication overhead. It causes more energy consumption. Therefore, due to the sensor nodes limited communication and computation resources, k cannot be too large in practical applications. The cluster head passes the new cluster head location with the Mobile Sink and the rest of the node of the cluster.

### V.SIMULATION AND RESULTS

Effectiveness and performance of our proposed scheme was analyzed with the help of simulations. The energy computation model has used for calculations is similar to the one proposed in [8].

#### A. Simulation Environment

The network is considered as a heterogeneous network. The parameters for the simulation environment are as follows:

1. Grid Size: The Region of Interest (ROI) is  $200 \times 200$ .
2. Node Deployment: 50 to 300 randomly deployed over the ROI.
3. Data sampling rate is randomly assigned a sampling rate between 2 to 4 bytes per unit of time.
4. Initial energy of each node is 8 Joules.

The medium access control (MAC) protocol applied in our simulation is CSMA. The transmission rate for each node is

19.2 Kb per unit of time. The simulation duration is 100000 units of time.

### B. Lifetime and Energy Consumption

This section shows the comparison of energy consumption performance for the various algorithms. Simulations were performed in Region of Interest considering 50-200 sensor nodes randomly deployed. Each sensor node has a data sampling rate of either 2 or 4 bytes per unit of time. The proposed scheme is compared with Random Waypoint [3], [10] and Static algorithm [6] and the Bisector approach [9].

In Fig. 3, compares the network lifetime (in unit of time) against the number of nodes for various data-gathering schemes. Simulation results show that our proposed scheme improves the network lifetime as mobile sink only collects data from the cluster heads and traverses the path to the cluster heads according to buffer overflow limits.

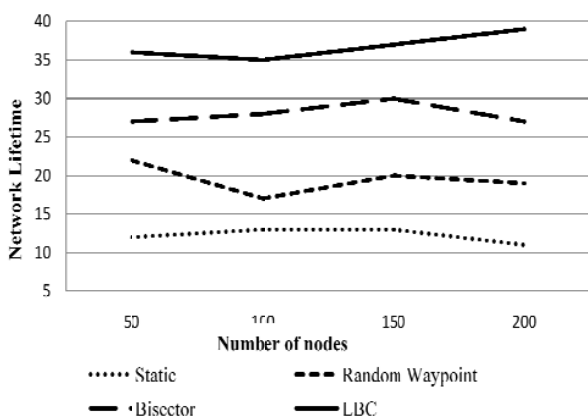


Fig. 3 Network Lifetime (in unit of time) for each data-gathering approach

Fig. 4 shows the average energy consumption against the number of nodes for the various data-gathering algorithms. The results show that our proposed scheme is better in conserving energy as there are fewer transmissions as the data is gathered from the cluster heads.

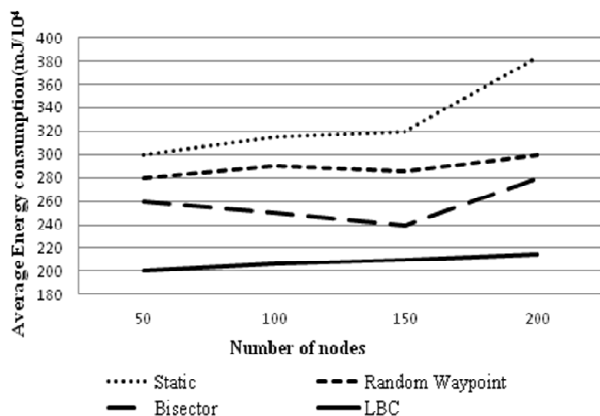


Fig. 4 Average energy consumption for each data-gathering approach

### VI. CONCLUSION AND FUTURE WORK

This paper presents a scheme to reduce the energy consumption in WSN. This paper used clustering to overcome the hot spot problem and with the dynamic cluster head selection, paper alleviated the distributed hot spot problem. In this sensor nodes collect samples at different rates. Our proposed scheme reduces the buffer overflow problem up to some level with the help of sharing the resource (memory) of the nodes.

There are many directions in which this work may be pursued further. We can model this work with the multiple mobile sink.

### REFERENCES

- [1] Th. Arampatzis, J. Lygeros, and S. Manesis, "A Survey of Applications of Wireless Sensors and Wireless Sensor Networks", in *Proceeding of the 13<sup>th</sup> Mediterranean Conference on Control and Automation*, 27-29 June 2005.
- [2] Jaichandran, R. Irudhayaraj, A.A. Raja and J.E., "Effective strategies and optimal solutions for Hot Spot Problem in wireless sensor networks (WSN)", in *10<sup>th</sup> International Conference of Information Sciences Signal Processing and their Applications (ISSPA)*, 10-13 May 2010.
- [3] R. C. Shah, S. Roy, S. Jain, and W. Brunette, "Data mules: Modeling a three-tier architecture for sparse sensor networks", In *IEEE Workshop on Sensor Network Protocols and Applications (SNPA)*, 2003.
- [4] E. L. Lawler, J. K. Lenstra, R.-K. A. H. G., and D. B. Shmoys, "Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization", John Wiley & Sons, 1990.
- [5] Heinzelman, W.R. Chandrakasan, A.; Balakrishnan, H. "Energy-efficient communication protocol for wireless microsensor networks", in *Proceedings of the Hawaii International Conference on System Sciences*, Maui, Hawaii, Jan 2000.
- [6] Lindsey, S, Raghavendra, C.S., "PEGASIS: Power-efficient gathering in sensor information systems" in *Aerospace Conference Proceeding, IEEE*, volume 3, 2002.
- [7] Yuki Kaneko and Hiroaki Higaki. "Ad-Hoc Buffering in Neighbor Nodes for Burst Data transmissions in Wireless Sensor Networks," in *12<sup>th</sup> International Conference on Computing on Computational Science and its Applications*, 2012.
- [8] A. A. Somasundara, A. Ramamoorthy, and M. B. Srivastava, "Mobile element scheduling for efficient data collection in wireless sensor networks with dynamic deadlines," in *Proceedings of the 25th IEEE International Real-Time Systems Symposium*, Washington, pp. 296-305, IEEE Computer Society, Washington, DC, USA, 2004.
- [9] M. Ma and Y. Yang, "SenCar: an energy-efficient data gathering mechanism for large-scale multihop sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 18, no. 10, pp. 1476-1488, 2007.
- [10] S. Jain, R.C. Shah, W. Brunette, G. Borriello and S. Roy, "Exploiting mobility for energy efficient data collection in wireless sensor networks", in *Mobile Networks and Applications*, vol. 11, no. 3, pp. 327-339, 2006.
- [11] Haiyun Luo, Fan Ye, Jerry Cheng, Songwu Lu, Lixia Z Hang "Two-Tier Data Dissemination in Large-Scale Wireless Sensor Networks" *Wireless Networks* 11, 161-175, 2005.