Efficient Technique for Collecting Data in Wireless Sensor Network

Anamika Pandey, Suman Srivastava, Ashu Gupta

Abstract— We are going to focus on the technique for collecting data from sensor nodes that increase the lifetime of wireless sensor networks. In early day's research, every sensor nodes collect raw data from the environment and directly send it to the base station.So more energy are consumed to transfer all these raw data to the sink. In proposed scheme, many different sizes of clusters are created. Each cluster has different number of cluster members and single cluster head. Here number of raw data is collected by each cluster member and then all that cluster members perform local approximate data collection technique on raw data within their cluster. Each cluster member find local approximate data value and send these approximate values to their cluster head .Now cluster head performs global approximate data collection technique on local approximate data values and find global approximate value.Now send this global approximate value to the base station. The approximate data collection approach reduces the redundant data within the networks and makes the wireless sensor networks more energy efficient.

Index Terms— Base station, Cluster , Cluster Member , Cluster Head , Energy ,Global Approximation ,Local Approximation , Network Life Time ,Wireless Sensor Network

——————————— ◆ ———————————

1 INTRODUCTION

A wireless sensor [1] has not only a sensing component but also processing, communication, and storage capabilities. With these enhancements, a sensor node is often not only responsible for data collection but also for correlation, network analysis, and fusion of its own sensor data and data from other sensor nodes. Sensor nodes communicate not only with each other but also with a base station (BS) using their wireless radios, allowing them to disseminate their sensor data to remote processing, analysis, visualization, and storage systems.

Simple sensors may only communicate and collect information about the observed environment; more powerful devices (i.e., devices with large processing, energy, and storage capacities) may also perform extensive processing and aggregation functions.

Wireless sensor networks have many applications. Some of them are futuristic while a large number of them are in practical use. The diversity of applications in the latter category is remarkable – target tracking, environment monitoring, pipeline (like water, oil, gas monitoring), structural health monitoring, agriculture, health care, supply chain
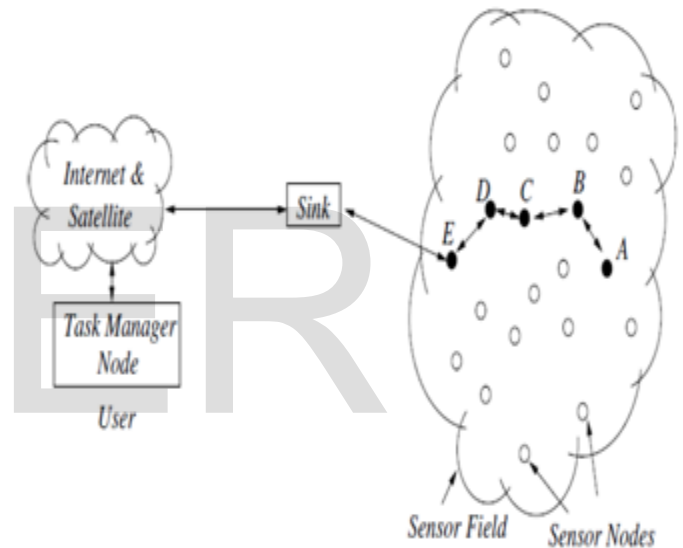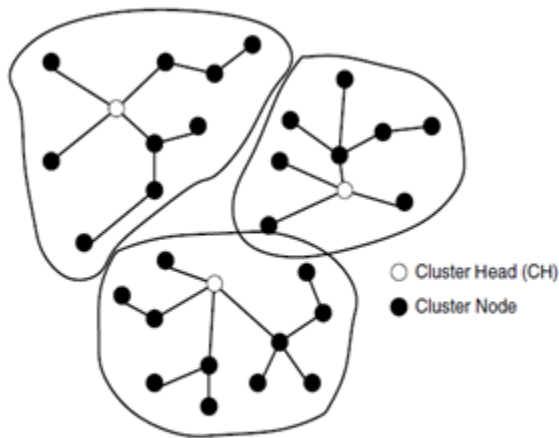
———————————————

- *Anamika Pandey currently pursuing masters degree program in Computer Science & Engineering in SIMT,Lucknow affiliated to UPTU,Lucknow, India . E-mail: anamz85@gmail.com*
- *Suman Srivastava is currently pursuing masters degree program (Modular) in Computer Science & Eengineering in SRMGPC ,Lucknow affiliated to UPTU,Lucknow,India. E-mail: suman.srivastava539@gmail.com*
- *Ashu Gupta currently working as Assistant Professor in Computer Science & Engineering department in KNIPSS, Faculty of Engineering,Sultanpur affiliated to UPTU,Lucknow,India. E-mail: ashu.gupta43@gmail.com*

Figure 1 Wireless Sensor Network

management, active volcano monitoring, transportation, human activity monitoring, and undergroud mining, etc.

In figure 1, different types of sensors are there that sense the data and send these data to the base station (sink). The base station collects these data and performs analysis. After analysis on that data, the results are distributed to the different type of user using the data distribution network.

Group of different number of sensor node is known as cluster. In figure 2, There are three Clusters and they have many cluster members and a single cluster head individually. These cluster members communicate to one another and to cluster head also.

Figure 2 Cluster Structure in WSN

Cluster member collects raw data from the environment and send information to their cluster head. Then cluster head collects information from its member node and send that information to the sink or base station. Clustering algorithms [2] limit the communication in a local domain and transmit only necessary information to the rest of the network.

2 RELATED WORKS

There have been many related works on data collection in WSNs.

Usman Raza & et al [3] presented the Model-driven data acquisition techniques which aim at reducing the amount of data reported, and therefore the energy consumed, in wireless sensor networks (WSNs). At each node, a model predicts the sampled data; when the latter deviate from the current model, a new model is generated and sent to the data sink. Evaluation typically focuses solely on the quantity of data reports suppressed at source nodes. the interplay between data modeling and the underlying network protocols is not analyzed. In this paper use the Model-driven data gathering, derivative-based predication. To improve lifetime further, we must revisit network design choices and address the extremely low data rates resulting from data modeling techniques.

D. Chu & et al presented the paper [4]. In this paper Prior database research has proposed to achieve this by pushing data-reducing operators like aggregation and selection down into the network. This approach has proven unpopular with early adopters of sensor network technology, who typically want to extract complete "dumps" of the sensor readings, i.e., to run "SELECT *" queries. Unfortunately, because these queries do no data reduction, they consume significant energy in current sensor network query processors. The conclusion of this paper is Database query processing ideas can play an important role in sensor networks.

Ken, focuses on using probabilistic models to provide approximate answers efficiently. The advantage of Ken is it assures the user of the faithfulness of the approximate answer to the actual sampled value, without assuming model correctness, while significantly reducing the volume of data transmitted.

Apoorva Jindal & et al presented the paper [5]. In this paper, physical phenomena monitored by sensor networks, e.g. forest temperature, water contamination, usually yield sensed data that are strongly correlated in space. With this in mind, researchers have designed a large number of sensor network protocols and algorithms that attempt to exploit such correlations. To carefully study the performance of these algorithms, there is an increasing need to synthetically generate large traces of spatially correlated data representing a wide range of conditions. Further, a mathematical model for generating synthetic traces would provide guidelines for designing more efficient algorithms. These reasons motivate us to obtain a simple and accurate model of spatially correlated sensor network data. The model can capture correlation in data irrespective of the node density, the number of source nodes or the topology and describe a mathematical procedure to extract the model parameters from real traces and generate synthetic traces using these parameters. Then, validate model by statistically comparing synthetic data and experimental data, as well as by comparing the performance of various algorithms whose performance depends on the degree of spatial correlation. Finally, create a tool that can be easily used by researchers to synthetically generate traces of any size and degree of correlation. In this paper use the spatially correlation model and also use the two tools .In this paper, proposed a model to capture the spatial correlation in sensor network data. This model can generate synthetic traces representing a wide range of conditions and exhibiting any degree of correlation and also described a mathematical procedure to extract the parameters of the model from a real data set. These model parameters are then used to generate synthetic traces having similar correlation properties as the real data trace.

3 PROBLEM DEFINITION

Earlier all the nodes in the sensor network send raw data to the sink (base station). So directly sending a large amount of raw data to the sink can lead to several undesired problems –

- The data quality may be deteriorated by packet losses due to the limited bandwidth of sensor nodes.
- The intensive data integration can lead to excessive energy consumption.
- Nodes may send redundant data to the sink

So aim of this paper is to carefully design the data collection strategy that reduce energy consumption on sensor nodes and increase the lifetime of the network and reduce the network traffic.

In proposed scheme, we use the approximate data collection approach. In this approach, a number of clusters are created in network and performs local approximate data collection on each cluster member. Now cluster head performs global approximate data collection according to model parameters.

4 PROPOSED SYSTEM

4.1 LOCAL APPROXIMATE DATA

In local data approximation, each cluster member sends the approximate value to the cluster head rather than raw data. So local approximation function is applied to all cluster members within their cluster.

Let each cluster members have a set of sensed raw data value T_1, T_2, T_3, T4…….. T_j, with particular time interval. Local approximate data value for each cluster member are calculated by using minimum function as in equation (1)-

$$LAD_i = min \{T_1, T_2, T_3, T_4, \ldots T_j\} \qquad (1)$$

In equation 1, where LAD is the local approximate data that is calculated on each cluster members within their cluster and then send these approximate values to the cluster head. Here "j" is the number of raw data values and "i" is the number of cluster members in a cluster.

4.2 GLOBAL APPROXIMATE DATA

In global data approximation, the cluster head collects the local approximate data values from all cluster members within cluster .So global approximation function is applied to cluster head .Global approximate data value is calculated by using minimum function as in equation (2)-

$$GAD = min \{LAD_1, LAD_2, LAD_3, LAD_4 \ldots LAD_i\} \qquad (2)$$

In equation (2) where $LAD_1,LAD_2,LAD_3\ldots\ldots.LAD_i$, are the local approximate data of "i" cluster member within their cluster and using equation (2) we will find out global approximate data (GAD) on cluster head and send this global approximate data value to the base station.

In figure 3, there are total number of five sensor nodes and one base station.In five sensor nodes, one is cluster head and remaining four are cluster members. Each cluster member CM1,CM2,CM3, and CM4 are collecting three
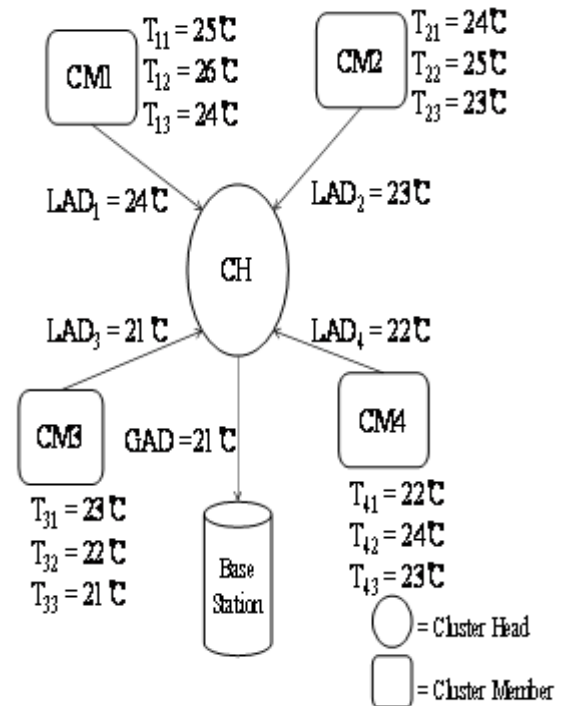

Figure 3 Approximate data collection

different temperature value T_1, T_2, T_3 and calculate minimum local approximate values LAD_1, LAD_2, LAD_3, LAD_4 respectively using equation (1).Then send these minimum local approximate values to their cluster head CH. Now cluster head CH calculate minimum global approximate value GAD from the values given by cluster members using equation (2) then send the minimum global approximate value GAD to the base station.

In above figure, raw data of cluster member CM1 are $T_{11} = 25℃$, $T_{12} = 26℃$, $T_{13} = 24℃$. Using equation 1, CM1 calculates $LAD_1 = 24℃$.Like CM1, all cluster member CM2 , CM3 , CM4 calcultate $LAD_2 = 23℃$, $LAD_3 = 21℃$, $LAD_4 = 22℃$ respectively using equation 1.Using equation 2, cluster head CH calculates GAD = 21℃.

Now CH send the value of GAD to the base station.So redundant value will be removed and there is no wastage of energy to transfer duplicate value.

5 SIMULATION RESULT

Here we are using NS2 (Network Simulator Version 2) simulator to implement the result and ubuntu operating system to perform simulation tool.

In figure 4, a cluster member send its approximate data to its cluster head.This is known as local collection of data.

In figure 5, a cluster head send its approximate data to the base station.This is know as global collection of data.
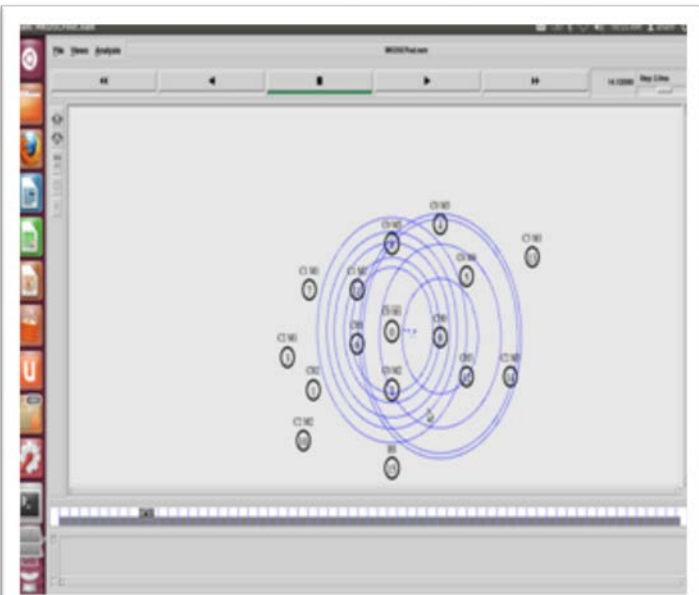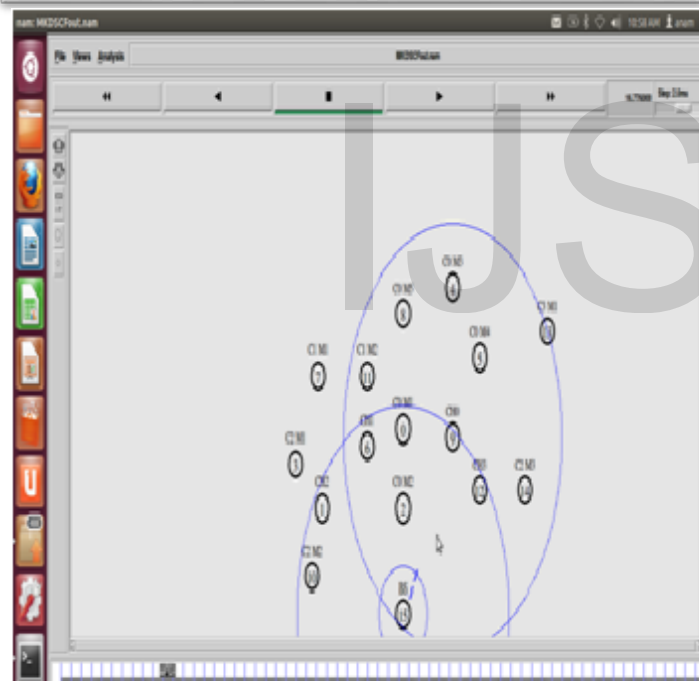
Figure 4 Local Approximation

Figure 5 Global Approximation

In figure 6, we are taking time in millisecond at x-axis and energy consumed in mile-joule at y-axis. When time increases, more energy will be consumed in normal data collection approach. But in approximate data collection approach energy consumed graph going to low with respect to time. So the graph shows how the energy will be consumed using normal data collection approach and approximate data collection approach. We show here that in approximate data collection approach using min function, low energy will be consumed and using this approach the life time of networks will increase.
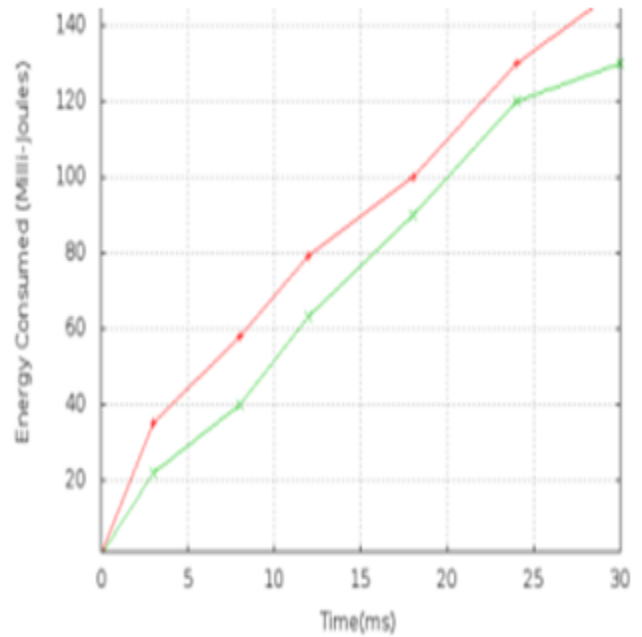
Figure 6 Energy Consumption Graph

When raw data will send directly by all sensor nodes to the base station then this approach is known as **Normal data collection technique**. But when nodes are in cluster form and each cluster member perform local approximation on raw data and cluster head perform global approximation on data given by all cluster member then this technique is known as **Approximate data collection technique**.

If a sensor node receives all temperature value instead of minimum temperature value, the redundant data may send to the base station and more energy will be consumed to transfer all these data. So we use **"min"** function to reduce the redundancy and to make energy efficient network.

6 CONCLUSION

We have studied that how a cluster member send a data value to its cluster head using local approximation and how a cluster head send a data value to the base station (sink) using global approximation. This technique is useful when large amount of sensor nodes are in network. As sensor nodes are battery driven, an efficient utilization of power is essential in order to use network for long duration. So we are using the approximate data collection approach that makes the wireless sensor network more energy efficient and also reduce the data traffic inside the networks. The approximate data collection is also enhancing the life time of the wireless sensor network.

REFERENCES

[1] Dargie and Poellabauer , 2010 , "Fundamentals of Wire less Sensor Networks" , John Wiley and Sons Ltd,27-28

[2] Akyildiz and Vuran August 2010, "Wireless Sensor Net Works", 978-0470-03601-3, John Wiley and Sons Ltd,333-334.

[3] Usman Raza, Alessandro Camerray, Amy .L Murphy, Themis Palpanasy and Gian Pietro Picco — "What does Model –driven data acquisition really achieve in wireless sensor network?".

[4] D. Chu, A. Deshpande, J.M. Hellerstein, and W.Hong,2006, "Approximate Data Gathering in Sensor Networks Using Probabilistic Models," Proc. 22nd Int'l Conf. Data Eng. (ICDE '06).

[5]A. Jindal and K. Psounis, "Modeling Spatially- Correlated Sensor Network Data", ACM Trans. Sensor Networks, vol. 2, no. 4, pp. 466- 499,2006

[6] Ali Dabirmoghaddam , Majid Ghaderi and Carey Wil liamson "Cluster-based correlated data gathering in wireless sensor networks".

[7] Sixia Chen, Matthew Coolbeth, Hieu Dinh, Yoo-Ah Kim and Bing Wang "data colletion with multiple sinks in wireless sensor networks".

[8] Ameer Ahmed Abbasi , Mohamed Younis "A survey on clustering algorithms for wireless sensor networks".

[9] Ian F. Akyildiz, Weilian Su, Yogesh Sankarasu bra maniam, and Erdal Cayirci "A Survey on Sensor Networks".

[10] K.Ramesh and Dr. K.Somasundaram "A comparative study of cluster head Selection algorithm in wireless sensor network".

[11] Sohraby, Minoli, and Znati , 2007 , "Wireless sensor networks" , John Wiley and Sons Ltd.