Dynamic Data Aggregation in Wireless Sensor Network

Karamjot Kaur
Student
Yadavindra College of
Engineering, Punjabi
University, Talwandi Sabo,
Punjab (India)

Sukhwinder Singh Sran
Assistant Professor
Yadavindra College of
Engineering, Punjabi
University, Talwandi Sabo,
Punjab (India)

Navjot Kaur
Assistant Professor
College of Engineering and
Management, Neighbourhood
Campus Punjabi University,
Rampura Phul, Punjab (India)

ABSTRACT

The wireless sensor networks are the networks used for the data collection and also used to sense various physiological parameters like humidity, temperature, atmospheric pressure, etc. The wireless sensors are equipped with limited energy which increases the requirement of energy efficient methods in order to run the WSNs for the longer periods. The data traffic aggregation is the important energy efficiency factor. The data traffic aggregation methods are used to produce the cumulative traffic streams to reduce the routing overheads. In our dissertation, we have proposed an algorithm based on the average values for the sensing networks. The proposed model works in the binary model based average values, where every cluster is connected to the two clusters in the next connectivity level. The cluster heads collect the data from the other nodes in the cluster and calculate the average value which is further forwarded to the next level cluster head towards the sink node. The average value calculated from the gathered data obtained from the CH's of different levels is forwarded to the sink node. The experimental results give the throughput two times more than previous method and reduction in the value of transmission delay and load.

Keywords

Data Aggregation, Wireless Sensor Networks, Clustering.

1. INTRODUCTION

A wireless sensor network consist a number of sensor nodes which co-operatively sense data and transmit data to base station. The size of wireless sensor node is very small. These nodes have limited processing capability very low battery power. Each of these nodes collected data and its purpose is to route this information back to a sink. All the sensor nodes are interacting with each other or intermediate sensor nodes. In wireless sensor network source node is the node which will send the requested data. Similarly sink node is the node which will send the request for the data. A sensor node that gathers data, based on its sensing system observations it processes data and route data packet to the base stations. The base station is serving as a gateway between end users and sensor nodes. The base station will also used for aggregation of data between the nodes.

This process needs more energy to transmit data when base station is located very far away from sensor nodes so that better technique is have fewer nodes sends data to base station. These nodes are aggregated nodes and a process knows as data aggregation in WSN. Data aggregation is the procedure of summarizing the data which comes from neighboring nodes and then transmitting compressed data to their destination. The main purpose of data aggregation algorithms is to collect and aggregate data in an energy efficient aspect in order that lifetime of network is increased. In aggregation, when one is lost may cause the loss of whole information aggregated from a number of nodes so packet loss become more destructive. So robustness is crucial issue for data aggregation. The major issue of data aggregation for event detection lies in dynamics of data traffic. Dynamic aggregation to handle events triggered irregular traffic. In some particular location the event appears at some time instance, data aggregation is necessary only when event occurs and only at nodes engaged in sensing and transmitting the results. In the contiguous link scheduling for data aggregation, links circumstance to one node are listed together to attain the time slots. By this node can start only one time to monitor the channel in a scheduled period. Due to broadcast nature of wireless sensor network the packets send away a node may be acquired by all the nodes with in transmission range. So there is interference between links. Two types of algorithms are used in contiguous link scheduling, centralized algorithms and distributed algorithms.

For the data aggregation, if there is tree topology, each intermediate node needs to start up only two times in duration, once for accepting data from its children nodes and once for imparting data to its parents. The objective of contiguous link scheduling problem in wireless sensor networks is to minimize the number of time slots in given period by reusing the coexistence transmissions to avoid collision and traffics. In this research, the new data aggregation model has been proposed using the binary average factor based approach for the sensor networks. The primary objective of the proposed work is to maximize the energy efficiency factor with the minimization of the probability of the error in the average value calculations. To implement our work, we have used Network simulator ns2.

The sensor node: A sensor node is a tiny device which has a less power, less memory, less cost and self organizing capability. These sensor nodes are used to sensing the pressure, vibration, temperature, humidity, sound, motion etc. in the wireless sensor network, these nodes can perform the function like computation of data, storage of data, communication of data and sensing of data. A sensor node consist four main components which are a sensing unit, a processing unit, a transceiver unit and a power unit and may have some additional components such as a location finding system, a power generator and a mobilizer. Sensing unit has two subunits: sensors and analog to digital converters (ADCs). The sensors have three categories Passive Omni-Directional, Passive Narrow-beam and Active sensors. When the sensor nodes sense the analog signal it is converted into digital signals by the ADC and then fed into the processing unit. On the processer sensed data is processed and stored into the memory. A transceiver unit is communication medium of node and it connects the node to the network. Power unit is the very important component of the sensor unit because it decides the lifespan of the sensor network. Some protocols are

used to increase the lifetime of the network which will keep some nodes at sleep mode. We can use the sensor nodes for continuous sensing, event detection, even ID, location sensing.

2. LITERATURE REVIEW

A literature review goes beyond the search for information and includes the identification and articulation of relationships between the literature and our field of research. While the form of the literature review may vary with different types of studies, the basic purpose remains constant:

Junchao Ma et al. [9] determinate the contiguous link scheduling problem in wireless sensor networks. In their work, during scheduling cycle a node wakes up once to receive data from its neighbours. In asset of the contiguous link scheduling, state transition can be reduced. In this paper they certify that the contiguous link scheduling complication in wireless sensor networks to be NP-complete. They propose centralized and distributed algorithms with formularized execution bounds in heterogeneous and homogeneous networks. **Virmani et al. [8]** explained that Data aggregation mechanism is one of the possible solutions to prolong the life time of sensor nodes. The Dynamic Data Aggregation Tree Algorithm (DDAT) creates aggregation tree. In DDAT the node having maximum available energy is used as parent node/ aggregator node. Final tree produced using the proposed algorithm proves to be best for enhancing the network lifetime. They present the best possible aggregation tree to minimizing energy utilization, minimizing cost and hence maximizing network lifetime. **Diaz and Leung [3]** proposed the distributed protocol for wireless sensor network such as Fast aggregation Tree (FAT) method. It is build an aggregation tree to report data generated by sporadic events which located randomly in monitored area. It is a cross-layer because the routing decision considers the sleep periods of the nodes with regard to stimulates the tree construction process. If the nodes sleep period is long, when compared to centralized protocols this protocol divides the tree construction time by number of hops. For this maximum tolerable delay, this allows us to increase the sleep periods and thus save energy. **Zhang and Cao [1]** they explained techniques to detect and track a mobile target. They introduce the concept of dynamic covey tree-based collaboration and consider the real constraints of a sensor network. For expansion and pruning, they explained two schemes: the conservative scheme and the prediction-based scheme. The prediction based scheme outperforms the conservative scheme and it can achieve a high coverage and low energy consumption. If the same tree expansion and pruning scheme is used, the localized reconfiguration shows better results when the node density is high and when the node density is low, the trend is revered. **Weigang et al. [7]** states that a approach with cross-layer design to recognize dynamic and robust data aggregation by use of direct support from MAC layer. A new MAC protocol, DA-MAC gives the channel content information. Based on this information a node can dynamically check when and where aggregation is performed. To tolerate packet losses, used a virtual overlay rings to forward one packet to multiple nodes. With low cost, this approach can handle both traffic dynamics and loss of packets.

3. METHODOLOGY

The proposed model has been designed for the sensing applications working on the environmental factor sensing. The humidity sensing is the task to sense the humidity in the environment in the given area. We will propose a scenario which will work on balanced tree-based formation of clusters using Binary Average factor. The lower level dominating nodes will aggregate data from sensor nodes within their own cluster and then forward average value of data towards higher level Cluster Head. Each higher level CH receive data from its neighbouring two lower level Cluster Heads, along with data from non-dominating nodes within its range. Then higher level cluster head calculate average value using weighted average factor method to minimize effects of extremely low and high readings produced due to the faulty or misplaced sensors and forward towards higher level Cluster Head. This Cluster Head forwards data towards Sink node. Thus whole network forms a layered hierarchical formation in the WSNs. The proposed model has been implemented using the network simulator 2 (NS-2). The various steps required to simulate the countermeasure for algorithm based on the average factor method are shown in figure:
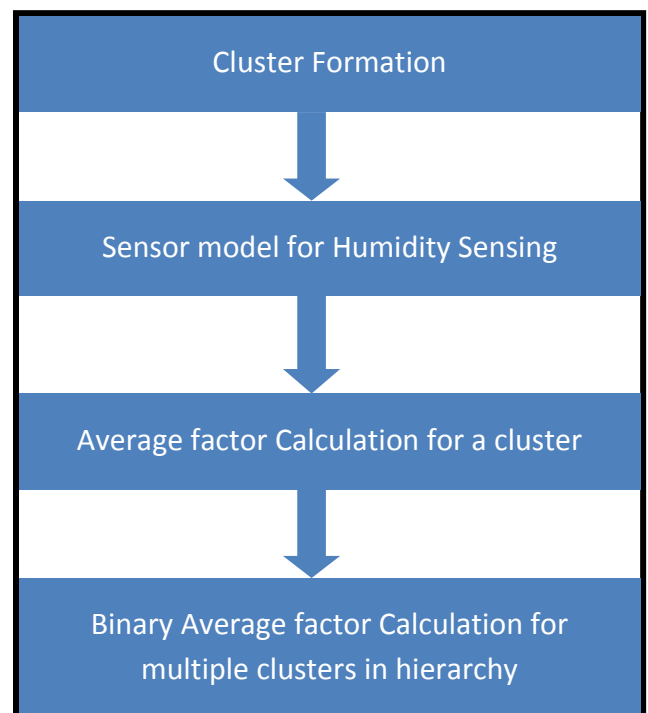

Fig 1 various steps to simulate countermeasure

Cluster formation: The cluster formation process leads to a two-level hierarchy where CH nodes from the higher level and the cluster-member nodes from the lower-level. The sensor nodes periodically send their data to the corresponding CH nodes. The CH nodes aggregate the data and forward them to base station through the intermediate communication with other CH nodes. Clusters are arranged in a tree-based structure for the hierarchical wireless sensor networks.

Sensor model for environment sensing: The environment sensing applications are deployed in the many applications, like humidity sensing, temperature sensing, weather prediction, crop science etc. The humidity applications depend upon the average value of the obtained data from multiple sources in order to get the overall humidity spread over the given area. The humidity sensing applications usually transmit the all sensed data towards the sink node, which adds the extra load on the sensor node path and the nodes lying on those paths.

Average factor Calculation for a cluster: The average factor is the method of calculating the average value on dominating point (cluster head) in the layered hierarchical formation in the WSNs. The proposed model offers the weighted average factor for the average calculation of the values on the cluster heads of the edge clusters. The cluster heads perform the average value calculation on the data obtained from the local nodes as well as the data received from the other cluster heads in the lower hierarchy. The data is undergoing the average values calculation method till it reaches the sink node.

Algorithm 1: Working of Averaging Factor Calculation (Weighted)

1. Input the value from each node.
2. Calculate the number of nodes N.
3. Find the extremely low values.
4. Find the extremely high values.
5. Count the number of extremely value sensors.
6. Sum all values other than extreme.
7. Divide the sum by (Total number of sensors – Extreme value sensors)
8. Return averaging factor.

Binary Average factor Calculation for multiple clusters in hierarchy: The binary connectivity model offers the connectivity of two nodes with the node in their higher level hierarchy. The binary connectivity model has been infused with the weighted average factor for the sensor networks for the humidity sensing.

As the result, the sink gets the perfect values which have undergone several level weighted average value calculations using the binary average factorization for the layered WSN. The proposed model offers the use of the mostly used wireless sensor formation in the layered fashion. The layered formation is the primary structure of the higher percentage of the sensor networks so we have designed our proposed model over such topological formation. The algorithmic structure has been describes in their functional sequence given below:

Algorithm 2: Working of Model

1. Power up the nodes.
2. Formation of the clusters.
3. Nodes sense the humidity rate.
4. Nodes forward the sensing data to clusters head.
5. Clusters head run the averaging factor calculation method (AVG).
6. Average factor calculation method returns the humidity average (HA).
7. Humidity average is forwarded to sink.

4. COMPARATIVE ANALYSIS

Transmission Delay: Network delay is an important characteristic of the network. The transmission delay is the delay time taken by a packet during its propagation from the source towards the sink node.

As shown in the figure x-axis shows time in seconds and y-axis shows delay in milliseconds. In the simulation, the delay has been computed for the simulation length of 9.5 seconds.
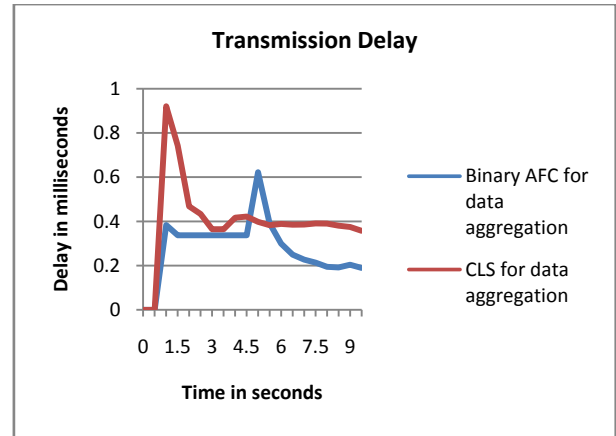

Fig. 2 Transmission delay Binary AFC v/s CLS for data aggregation

The recoding interval has been set to 0.5 seconds for each transmission. In the Binary average factor calculation method, the minimum delay has been recorded at 0 milliseconds and the maximum delay has been recorded at 0.622 milliseconds. The delay average for all 20 transaction has been calcualted at 0.276 milliseconds. In the contiguous link scheduling for data aggregation, the the minimum delay has been recorded at 0 milliseconds, the maximum delay has been recorded at 0.921 milliseconds and delay average has been calcualted at 0.398 milliseconds. Delay is decreased in binary average factor calculation method and hence energy of the network is saved and network lifetime is increased. So the binary AFC method offers the better performance on the basis of the transmission delay than the CLS for data aggregation.

Throughput: Throughput is the number of bits transfer per second. The amount of data transferred from one place to another or processed in a specified amount of time. It is typically measured in Kbps, Mbps and Gbps. Network throughput is the rate of successful message delivery over a communication channel.
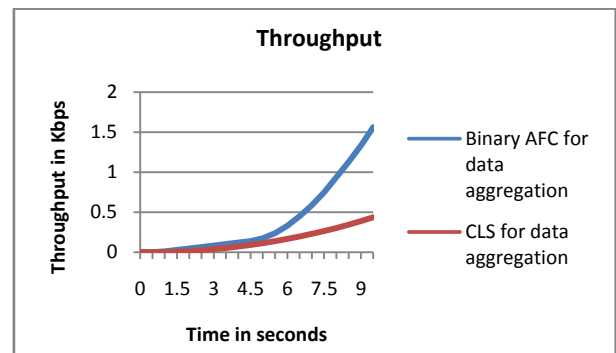

Fig. 3 Throughput Binary AFC v/s CLS for data aggregation

As shown in the figure x-axis shows time in seconds and y-axis shows throughput in kbps. The throghput has been computed for the simulation length of 9.5 seconds in the simulation. The recoding interval has been set to 0.5 seconds for each transmission. The minimum throughput has been recorded at 0 kbps and the maximum throghput has been recorded at 1.564 kbps in binary average factor calculation method. In contiguous link scheduling, minimum and maximum throughput values are registered 0 kbps and 0.436 kbps respectively. Average throuput values are registered 0.405 kbps in binary AFC method and 0.144 kbps in

contiguous link scheduling for data aggregation. So we can see that the throughput of the binary AFC method is higher than the CLS for data aggregation, so the objective of the dissertation is fulfilled.

Table 1. Comparison based on parameters

Parameters	Binary AFC			CLS		
	Minimum	Maximum	Average	Minimum	Maximum	Average
Transmission Delay	0.0	0.62	0.27	0.0	0.92	0.39
Throughput	0	1.56	0.41	0	0.43	0.14
Network Load	0	0.09	0.03	0	0.19	0.09
PDR	0	1	0.27	0	1	0.41
Data Lost	0	72	34.2	0	110.8	28.1

Load: Load is the amount of data (traffic) being carried by the network at a particular time. The network load varies from time to time. It is typically measured in Kbps. Load depends on the whether the source node has the required data or not.
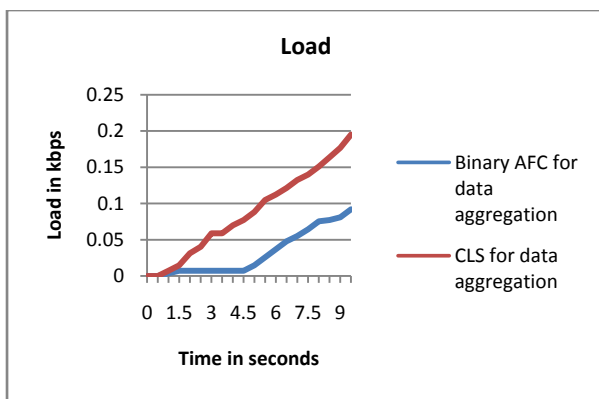

Fig. 4 Load Binary AFC v/s CLS for data aggregation

Figure present that simulation time is given along with x-axis and y-axis shows load in kpbs. The registered interval has been set to 0.5 seconds for each transmission. In binary average factor method, minimum and maximum load are registered 0 kbps and 0.092 kbps respectively. The minimum load has been recorded at 0 kbps and the maximum load has

been recorded at 0.195 kbps in contiguous link scheduling. Average load for AFC method and CLS is calculated 0.031 kbps and 0.087 kbps respectively from all the transactions. There is the less load in binary average factor method as compared to contiguous link scheduling. So the objective is gained.

Packet Delivery Ratio: This is the ratio of number of packets received by the node to the number of packets transmitted by the node. If this is greater in value it means protocol or network is good in performance. By this we can measure how efficiently our system is working. It is measured in packets sent. If the PDR for a system is low then it is not a good network and some problems can occur in that due to delay etc.
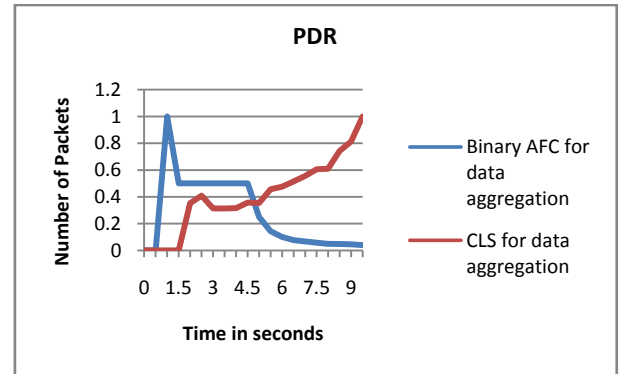

Fig. 5 PDR Binary AFC v/s CLS for data aggregation

In the figure x-axis shows time in seconds and y-axis shows number of packets. The number of packets has been computed for the simulation length of 9.5 seconds simulation. The recoding period has been set to 0.5 seconds for each transmission. In both binary AFC method and cotiguous link scheduling the minimum number of packets is registered at 0 and the maximum number of packets registered at 1. For all 20 rotations the average number of packets has been recorded at 0.269 in binary average factor method and 0.409 in contiguous link scheduling for data aggregation.

Data Lost: Packet loss is the breakdown of single or more delivered packets to reach their destination. This incident can cause visible effects in all digital infrastructures.
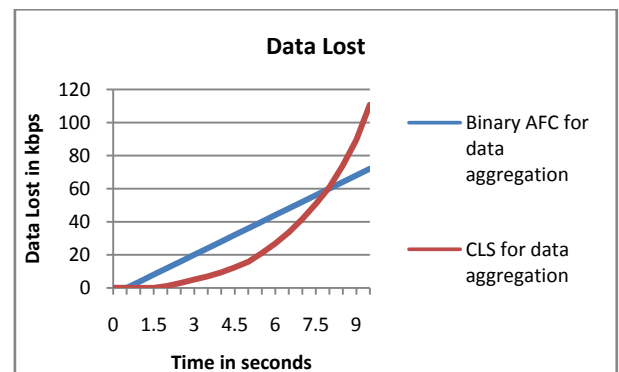

Fig. 6 Data Lost Binary AFC v/s CLS for data aggregation

The causes of packet loss comprise of insufficient signal power at the destination, usual or human made obstruction, extreme system sound, hardware crash, overwhelmed network nodes.

As Shown in figure, simulation time is given along with x-axis and y-Axis gives amount of lost data packets in kbps. In this amount of lost data has been computed for simulation

period of 9.5 seconds. After every 0.5 second, amount of lost data is calculated.The minimum value is registered at 0 kbps and maximum value is recorded at 72 kbps for binary average factor method. In the contiguous link scheduling minimum and maximum values are calculated 0 and 110.8 respectively.

5. CONCLUSION

The binary average factor model has been proposed to calculate the average of data collected at various layers. The proposed model has been designed for the energy efficient wireless sensor applications. This technique reduces the data volume at each level after calculating the weighted average. Each CH detects the extremely high and low values in gathered data and eliminates these extreme values in order to prevent the miscalculations for calculating the average. The results of the proposed scheme have been evaluated on the basis of transmission delay, throughput, network load, packet delivery ratio and packet loss. It has been found that proposed scheme performs better than the existing scheme. The Throughput of the techniques is 2 times more than the existing scheme and the transmission delay is reduced in contrast existing scheme. Network load is 2 times less in binary average factor method as compared to contiguous link scheduling.

6. FUTURE SCOPE

In future, the proposed model can be enhanced by introducing multi-factor average calculation for the higher level of precision. The same scheme can be replicated on the upper layer or levels.

7. REFERENCES

[1] W. Zhang and G. Cao, "DCTC: Dynamic Convey Tree-Based Collaboration for Target Tracking in Sensor Network", IEEE Transactions on Wireless Communication, Volume 3, pp. 1689-1701, 2004.

[2] K. W. Fan, S. Liu, P. Sinha, "Structure-Free Data Aggregation in Sensor networks", IEEE Transaction on Mobile computing, Volume 6, Number 8, pp. 929- 942, 2007.

[3] M. O. Daiz and K. K. Leung, "Dynamic Data Aggregation and Transport in Wireless Sensor Network", IEEE Transactions on Wireless Communication, 2008.

[4] N. S. Patil, P.R. Patil, "Data Aggregation in Wireless Sensor Network", in proceedings of IEEE International Conference on computational Intelligence and Computing Research, 2010.

[5] K. Maraiya, K. Kant, N. Gupta, "Architectural Based Data Aggregation Techniques in Wireless Sensor Network: A Comparative Study", in proceedings of International Journal of Computer Science and Engineer, Volume 3, Number 3, pp. 1131-1138, 2011.

[6] S. Kwon, J. Hoon Ko, J. Kim, C. Kim, "Dynamic timeout for data aggregation in wireless sensor networks", in proceedings of Computer networks, pp. 650-664, 2011.

[7] W. Wu, J. Cao, H. Wu, J. Li, "Robust and Dynamic data aggregation in wireless sensor network: A cross-layer approach" in proceedings of IEEE 9th International Conference on UIC-ATC, pp. 306-313, 2012.

[8] D. Virmani, G. Mandal, N. Beniwal, S. Talwar, "Dynamic Data Aggregation Tree for Data Gathering in Wireless Sensor Network", in proceedings of International Journal of Innovative Technology and Exploring Engineering(IJITEE), Volume 2, Number 3, pp. 226-230, 2013.

[9] J. Ma, W. Lou, X. Y. Li, "Contiguous Link Scheduling for Data Aggregation in Wireless Sensor Networks", IEEE Transactions on parallel and distributed systems, Volume 25, Number 7, pp. 1691-1701, 2014.

[10] A. Kumar, R. Baksh, R. K. Thakur, A. P. Singh, "Data Aggregation in Wireless Sensor Networks", in proceedings of International Journal of Science and Research (IJSR), Volume 3, Number 3, pp. 249-251, 2014.

[11] P. Kaur and M. Katiyar, "Improved Chain Based Routing Protocol for WSN", in proceedings of International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Number 11, pp. 115-121, 2012.