Data Aggregation and its Impact on Overall QoS of Lossy Wireless Sensor Network: A Survey

Harshada Kupade
Department of Computer Engineering
JSPM's JSCOE Hadapsar-28
Pune, India
harshkupade.47@gmail.com

Madhav Ingle
Department of Computer Engineering
JSPM's JSCOE Hadapsar-28
Pune, India
ingle.madhav@gmail.com

Abstract— In Wireless sensor network, packet loss due to buffer overflow of sensor nodes and the delay caused due to retransmission of that lost packet is a major problem. So a big challenge is to improve the delay performance along with reliability of WSN. The goal of the proposed system is to broadcast sensed data to focus on delay performance. Also congestion at collector node due to duplicate sensed data in network is reduced by data aggregation, for this we compute a function at the collector node from the information gathered by spatially distributed sensor nodes. Because wireless sensor network applications require various levels of communication reliability (CR), the data transmission should satisfy the desired CR of the applications. Here, we propose a flexible loss recovery mechanism (called Active Caching) for sensor network applications with various CRs. The proposed scheme caches data packets at intermediate nodes over routing paths computed by CR to retransmit lost packets during multi-hop transmissions.

Keywords- *WSN, CR, CH, Data Aggregation, TDMA*

*****

I. INTRODUCTION

Wireless sensor network (WSN) is nowadays emerged as a promising technology in the recent studies in electronics, networking, and information technologies. The data which is transferred over the wireless sensor network has great impact on the link load. The handling of this data load against the congestion, its overall reliability, and loss recovery is very tough task. Data aggregation is the process of collection of useful data and aggregation of it for further transmission In WSN, data aggregation is an effective way which saves the limited resources. Data coming from multiple sensor nodes is first collected and then aggregated as if they are about the same attribute of the phenomenon when they reach the same designated routing node on the way back to the sink node. This task of collection of information at intermediate forwarding node can substantially increase network efficiency by reducing the total number of transmissions.

On the other side, data aggregation also increases the amount of the information contained in a single IP packet which makes the system vulnerable to packet loss. Instead of retransmitting each lost packet, which incurs additional delay in transmission, wireless broadcasting is an effective strategy which will increase delay performance while satisfying reliability constraint.

Due to the requirement of low packet loss rate in wireless sensor networks, more reliable data transmission is desirable. Because wireless sensor network applications require various levels of communication reliability (CR), here we propose a flexible packet loss recovery for sensor network applications with various CRs. This scheme caches data packets at intermediate forwarding nodes over routing paths computed by CR to retransmit these lost data packets during multi-hop transmissions. The rest of this paper is organized as follows: Section II presents literature review about WSN and data aggregation, Section III describes the proposed model. Section IV shows the implementation results and analysis. Finally, the conclusion is presented in Section V.

II. RELATED WORK

First, a survey of challenges in wireless sensor network is done. Since the Wireless Sensor Network consists of tremendous sensors and have limited power source, they face the problem of energy depletion. Apart from this Wireless Sensor Network faces the problem of Congestion, Unbalanced Energy Distribution, Routing, Security, Node Redundancy and Data-logging [2],[11].

Activating the hot spot problem gives the unbalanced energy distribution. In WSN, there unbalanced power consumption exists, that is, nodes near the data sink or base station , called hot spots, have a high probability of forwarding a high amount of packets and die early. These hot spots consume more power in WSN. Here further study includes sensor nodes that need to be deployed very densely and in a random fashion. They should be able to operate without human intervention. Clustering is a technique employed to increase the various capabilities of a sensor network [3].

Cluster-based communication has been addressed for these networks for various reasons such as scalability and energy efficiency. The problem of adding security to cluster based communication protocols for homogeneous wireless sensor networks consisting of sensor nodes with severely limited resources [4].

Energy Efficient Clustering Scheme for Data Aggregation proposes a comprehensive energy consumption model for multi-tier clustered sensor networks in which all the energy consumptions not only in the phase of data transmissions but also in the phase of cluster head rotations are taken into account. But mobility of nodes is also a considerable issue [5].

Another one technique of data aggregation called Redundancy Elimination for Accurate Data Aggregation (READA) has been proposed. By exploiting the rang of spatial correlations of data in the network, READA applies a grouping and compression mechanism to remove duplicate data in the aggregated set of data to be sent to the base station

without largely losing the accuracy of the final aggregated data. But even after compression of data, data duplication at some level exists [6].

Spatial and temporal multiple aggregation (STMA) scheme has been proposed to minimize energy consumption and traffic load when a single or multiple users gather state based sensor data from various sub-areas through multi-hop paths. The performance of STMA has been evaluated in terms of energy consumption and area-to-sink delay [7].

A flexible loss recovery mechanism called Active Caching, caches data packets at intermediate nodes over routing paths computed by CR to retransmit lost packets during multi-hop transmissions. Because the Active caching presents a tradeoff between end-to-end delays and memory requirements dependent on CR, it can be used flexibly in various sensor network applications. This mechanism can overcome the problem of data loss due to buffer overflow at each node; we are going to do implement it in proposed system [8].

III. PROPOSED SYSTEM

In general in hierarchical wireless sensor network each sensor node not only generates its own data, but also relays others data to the sink via multi hop wireless communications. The wireless channel is assumed to be lossy. A packet loss can happen due to buffer overflow of sensor node and this lost packet can be restored by retransmitting it, which however, results in additional delay. Since many applications have both reliability and delay constraints, the relationship between the reliability and the delay performance is important.

The proposed system consists of the hierarchical cluster based network where each cluster having some sensor nodes which sense the event and transmit the data packet to the respective cluster head, cluster head will gather all the information and send it to the special node called sink. The proposed work is to focus on the delay performance of in-network aggregation in lossy wireless sensor networks. The packet loss in the network and the delay caused in retransmitting it is examined, and the main focus is to decrease this delay in retransmission. In data aggregation, since the information contained in a single packet is highly intensified after several in-network computations, a packet loss can significantly impact the computation result, and thus a higher level of protection is required for each packet transmission. We show that aggregation with wireless broadcast can substantially reduce the delay while satisfying the reliability constraint.

Also another focus of proposed research is on data loss due to buffer overflow at any node in network. To reduce this loss first buffer monitoring of overall network is done. The system caches data packets at each node when its buffer reaches a certain value called threshold. In the cache this data is stored with high priority and this priority is used to retransmit lost packets during next transmissions.

A. System Design

First step is to initialize all the nodes in the network. Then design and implement the hierarchical topology. In this, sensor nodes are divided into clusters, and for each cluster, one representative node, which called cluster head (CH), aggregates all the data within the cluster and sends that data to other collector node close to the sink.

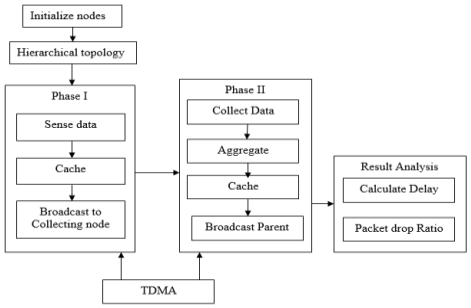

Figure 1. Proposed System Design

1) Phase I: In first phase, Sensor node senses the data from environment and sends it to collecting node through wireless broadcast. It uses TDMA for broadcasting.

If the buffer of sensor node is full and the data is going to be lost, then there is a caching mechanism with which we are going to cache this data and after some time transmit it within the network with higher priority.

2) Phase II: In second phase, Collector node collects all the data from different sensor nodes. Aggregation function is computed to remove the duplicate data. Again this aggregated packet is broadcasted to send it to sink node. It also uses TDMA for broadcasting.

At collector node also caching mechanism is attached to handle the buffer overflow.

3) Phase III (Result Analysis): Delay calculation and packet drop ratio are calculated with the help of trace file data. The information in the trace file includes all data about every packet which will be further used to generate graphs of comparison.

B. Objectives of Proposed System:

- Primary objective of proposed research is to reduce the traffic by data aggregation in wireless sensor network.
- Other objective of proposed research is to work on buffer overflow problem in WSN and to decrease the data loss due to it.
- Another one objective of proposed research is to improve the QoS parameters of network.

C. Algorithm With Tired Structure:

In this section we propose new network architecture with wireless broadcast to improve the delay performance while achieving the same level of reliability. Each node (at depth *d*) has at least *x(n)* parents (at depth *d-1*), and transmits a packet through the wireless broadcast channel to all parents. At the root, we assume that the sink has *x(n)* antennas and it can process signals from multiple antennas. In this architecture, we say that a node successfully transmits a packet if the broadcasted packet is successfully received by one of *x(n)* parents. The intuition can be better described using Fig. 2.

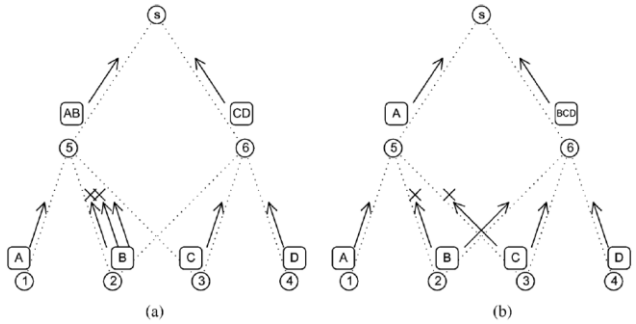

Figure 2. (a) Aggregation with unicast (b) Aggregation with broadcast

In above figure, transmissions over lossy wireless links are assumed. Each transmission is denoted by an arrow, and a failed transmission is denoted by a cross at the end of the arrow. Under aggregation with unicast, it needs four transmissions for information B to be successfully delivered to the sink, while it needs two transmissions under aggregation with broadcast.

Now, we assume that *n* wireless sensor nodes are uniformly deployed over a disk of radius 1. The network is a time-slotted TDMA system. At the beginning of each time-slot, each sensor node generates a packet with the sensed information. A time-slot is further divided into mini slots, and in each mini-slot, a single packet can be transmitted.

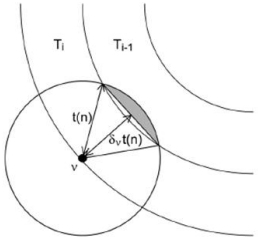

Figure 3. Parents (in T_{i-1}) of node *v* (in T_i) are located in shaded area

Routing is simplified using the tiered structure. Every node *u* in is a parent of node *v* in T_{i+1} if its distance is no greater than *t(n)*. Transmissions are scheduled from the outermost tier to the sink tier by tier one at a time, so that nodes in T_i can transmit only after all nodes in T_{i+1} finish their transmissions. We group nodes in each tier into mutually exclusive subsets such that all nodes in a subset can transmit simultaneously. Let *H(i,j)* denote the j^{th} subset in T_i, and let h_i denote the total number of subsets in each tier such that all nodes in T_i can finish a single transmission in h_i mini-slots.

Algorithm:

for $i = 1/\delta t(n)$ to 1 **do**

for $j = 1$ to h_i **do**

Each node μ in $H(i,j)$ broadcasts its (aggregated)

information $(1 + r_b(n))$ times.

if node $\mu \in T_{i-1}$ receives the packet **then**

Node μ does aggregation and updates its

information.

end if

end for

end for

In Algorithm, each node has at least $nt(n)^2$ parents and the maximum hop distance from a node to the sink is $1/t(n)$. Suppose that node v is located in T_i as shown in Fig.3. The number of parents of node *v* in T_{i-1} is no smaller than the number of nodes in the shaded area. Since nodes are uniformly distributed, it can be easily shown that the number of nodes in the shaded area is bounded by $nt(n)^2$.

IV. RESULT & ANALYSIS

In this section we will concentrate on performance analysis and measurable parameters of proposed system. Aim of proposed system is to increase the reliability and minimize the delay in retransmission of lost data packet. Hence performance optimization can be achieved in WSN using simulations; we developed an evaluation environment using ns-3.20. To run simulation, topology set up is done.

The scenario is such that, there is a hierarchical topology of 50 nodes Here all common nodes in the cluster will generates the events according to the frequency set for packet generation (inverse of time interval set) and forward these packet to their cluster heads. We can observe that if the incoming packet flow rate is more than processing and forwarding speed of the intermediate nodes Buffer Overflow may occur in the network and packet may get lost. Ultimately we may lose the important information which is carried by that packets as well as energy loss (for processing of that packet till it reaches to that node). Ultimately end to end delay will be more, Reliability, Performance of the network will degrade and hence efficiency too. So in this scenario we set the threshold value for the node buffer, so when node buffer reaches that threshold value incoming data is cached to cache. Again data aggregation at collecting node will reduce data duplication in the network and congestion caused due to it. Ultimately data broadcasting will overcome the delay in retransmission of lost packet.

When we run this scenario (script) with ns-3.20, we will get all the related readings from which we will plot the graphs showing effects of this system on different parameters. Parameters we are going to use for this performance analysis are:

- Packet Delivery Ratio
- End-to-End Delay
- Packets Generated
- Packets Received

The expected results from the proposed system includes the reduction in end-to-end delay and reduction in packet loss due to buffer overflow problem, also improved reliability along with network lifetime, efficiency and overall QoS of the network.

V. CONCLUSION

In cluster based wireless sensor network common sensor nodes within the cluster sense the event periodically or continuous and transmit the data packets to the respective cluster head. When there is buffer full condition at sensor node, incoming data is lost. So this scenario will reduce data loss by caching the incoming data and hence improving the reliability. Another one point is congestion at collector node due to duplicate data, due to this the collector nodes will consume more energy and die early. So this scenario will aggregate the data at collector node to remove data duplication and send this data to the sink, ultimately this will increase the reliability and network lifetime. Third point is that retransmission of lost data in multihop network will take more time, so the scenario includes data broadcasting technique which will reduce end-to-end delay in the network.

In future we will try to check security of the network and data. Also there is need to check an intermediate node or source node is really a part of our network (A security challenge).

REFERENCES

[1] ChangheeJoo, Ness B. Shroff, "On the Delay Performance of In-Network Aggregation in Lossy Wireless Sensor Networks", IEEE/ACM Transactions onnetworking, vol. 22, no. 2, April 2014.

[2] Arun Kumar, Dr. V.K.Katiyar, "A Survey on Challenges in Wireless Sensor Networks", IJETCAS,5(5), June-August 2013, pp. 471- 475.

[3] Meng, Jin-Tao, Jian-Rui Yuan, Sheng-ZhongFeng, and Yan-JieWei. "An Energy efficient Clustering Scheme for Data Aggregation in Wireless Sensor Networks", Journal of Computer Science and Technology 28, no. 3 (2013): 564-573.

[4] Sumit Chaudhary, Neha Singh, Avinav Pathak and A.K Vatsa," Energy Efficient Techniques for Data aggregation and collection in WSN", International Journal of Computer Science, Engineering and Applications (IJCSEA) vol.2, no.4, August 2012.

[5] SaminaEhsan and BechirHamdaoui, "A Survey on Energy-Efficient Routing Techniques with QoS Assurances for Wireless Multimedia Sensor Networks", IEEE Communications Surveys Tutorials, Vol. 14, No. 2, Second Quarter 2012.

[6] K. Pradeepa,W. R. Anne, and S. Duraisamy, "Design and Implementation Issues of Clustering in Wireless Sensor Networks", Int. J. Comput. Applications, vol. 47, no. 11, pp.2328, 2012.

[7] K. Khedo, R. Doomun and S. Aucharuz, "READA: Re- dundancy Elimination for Accurate Data Aggregation in Wireless Sensor Networks", Wireless Sensor Network, Vol. 2, No. 4, 2010, pp. 302- 308.

[8] Dae-Young Kim and Jinsung Cho, "Active Caching: A Transmission Method to Guarantee Desired Communication Reliability in Wireless Sensor Networks", IEEE Communications Letters, vol. 13, no. 6, June 2009

[9] Chang Wu Yu, Shan Shiung Wei, Bing JiunShie, "Activating the Hot Spot Problem by Switching Multiple Sinks in Wireless Sensor Networks", IEEE Mobile Ad-hoc and Sensor Networks, pp. 204, 14-16 Dec. 2009.

[10] Y. Wang, G. Attebury, and B. Ramamurthy, "Survey of Security Issues in Wireless Sensor Networks", IEEE Commun. Surveys Tuts., vol. 8, no. 2, pp. 223, 2006.

[11] H. S. Kim and W. H. Kwon, "Spatial and Temporal Multi- Aggregation for State Based Sensor Data in Wireless Sensor Networks", Telecommunication Systems, Springer Netherlands Volume 26, No. 2-4 / June, 2004 161-179.