A Reliable Data Aggregation Forwarding Protocol for Wireless Sensor Networks

Basavaraj S. Mathapati 1†, Siddarama R. Patil 2†† and V. D. Mytri 3†††,

Appa Institute of Engineering & Technology, Gulbarga, Karnataka, India.[1]
PDA College of Engineering, Gulbarga, Karnataka, India. [2]
Shetty College of Engineering, Gulbarga, Karnataka, India.[3]

Summary
Extensive research has been done on reliable data forwarding in sensor networks. Existing protocols provide reliability at the cost of high energy consumption. In Wireless Sensor Networks (WSN), data aggregation is essentially used to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced. Power consumption is an important aspect to be considered in the data aggregation which is a scarce resource and they are irreplaceable. In this paper, we propose to design an energy efficient reliable data aggregation technique for wireless sensor networks. Initially we form clusters and a coordinator node (CN) is selected near the cluster in order to monitor the nodes in the cluster. The CN selects a cluster head (CH) in each cluster based upon the energy level and the distance to the CN. The packets sent by the sensor nodes are aggregated at the CH and transmitted to the CN. The CN measures the loss ratio and compares it with a threshold value of loss ratio. Depending upon this value, the forward node count is incremented or decremented and the cluster size is adaptively changed, ensuring reliability and balanced energy consumption. We have compared the new protocol with Distributed power control Energy Efficient Routing Protocol for wireless Sensor Networks (DPC-EERP) using NS-2, and evaluated the protocol by varying nodes. The result shows that our A Reliable Data Aggregation forwarding protocol for wireless sensor networks (RDAFP) outperforms the DPC-EERP.

Key words:
Reliable, Cluster head, Cluster Node, Cluster head, Data Aggregation

1. Introduction

In data aggregation, the aggregation processes are used to aggregate the sensor data effectively. Data aggregation techniques enhances the network lifetime by gathering and aggregating the data in an energy efficient manner. A striking method for data gathering in wireless sensor networks involves distributed system architectures and dynamic access via wireless connectivity. In the case of energy constraint wireless sensor networks, the data aggregation techniques intend to eradicate the redundant data transmissions thereby improving the lifetime of the network. [1]

Due to that the sensor nodes are tightly packed in the sensor networks, there are possibilities for the nearby sensor nodes to overlap sensing ranges. Because of this, redundant or correlated data are collected by the sensor networks. In order to
save the energy, the data correlation is subjugated which effectively reduces the amount of data transmitted in the network. In wireless sensor network routing, data aggregation proves to be an important aspect. The data originating from different sensor nodes aggregate together in the sink node during transmission. [2]

The main purpose of the data gathering in wireless sensor network (WSNs) is to obtain valuable information from the operating environment. It has been proven that the data redundancy can be eradicated and the communication load can be reduced using the data aggregation techniques. Multiple data sources and a data sink are included in the typical communication patterns of data aggregation. A data aggregation tree is constructed using the transmitted packet and this is similar to the reverse multicast structure. [3]

Advantage of Data Aggregation: Robustness and accuracy of information acquired from the network can be improved effectively. The data aggregation requires the data fusion processing in order to reduce the redundant information which is present in the data collected from the sensor nodes. Traffic load is minimized and the energy in the sensors can be conserved with the help of data aggregation. Disadvantage of Data Aggregation: The cluster heads are also known as the data aggregator nodes which combine the data in order to send it to the base station. There are chances of malicious attackers in the cluster head or the aggregator node. The accuracy of the aggregate data sent to the base station cannot be guaranteed when the cluster head is compromised. The uncompromised nodes send several copies of the aggregate result to the base station which increases the power consumed at these nodes [4].

In this paper, we proposed to develop A Reliable Data Aggregation forwarding protocol for wireless sensor networks (RDAFP) which is energy efficient and reliable. This technique is based on cluster formation and the loss ratios

Manuscript received May 5, 2012
Manuscript revised May 20, 2012

of the clusters are measured so that the energy consumption can be effectively reduced. Reliable transmission can be provided in the clusters using a coordinate node.

The reminder of this paper is organized as follows. In Section II, we introduce previous work related to our study. Section III gives description of proposed protocol. The simulation results in Section IV and in Section V paper is concluded.

2. Related Work

Ren P. Liu et al [5] have proposed an Efficient Reliable Data Collection (eRDC) algorithm. Maximum number of retransmissions is controlled in order to achieve energy savings. Dynamic programming concept is used to find the optimal solution. Implementation of eRDC is provided which uses next hop link quality and number of hops for determining number of retransmissions.

Volker Turau et al [6] have presented the design and preliminary evaluation of a reliable data gathering service of periodic data in the face of poor link quality and frequent disconnects. The data is buffered by persistent storage provided by the nodes using services based on a packet-level, and hop-by-hop routing protocol. This design also provides an upper limit for sampling rate that is handled reliably.

Hemant Sethi et al [7] have proposed an Energy Efficient Interest Based Reliable Data Aggregation (EIRDA) Protocol for WSNs. Here each cluster considers the uniform distribution of sensor nodes using EIRDA which is a static clustering scheme. Beta-distribution function is used to provide reliability with the help of Functional Reputation concept. The overall impact of all measures taken at each phase of protocol implementation is clearly visible on the energy spent in the setup phase of the protocol.

David Gugelmann et al [8] have presented a novel data dissemination protocol with a focus on reliability and energy-efficiency. Scheduling image dissemination only during reserved time slots eliminates interference with the regular data gathering protocol and increases the observability during the network reprogramming phase.

Mingming Lu et al [9] have proposed the data-gathering problem in wireless sensor networks from the maximization of the expected network utility point of view. The resource scarcity and the unstable nature of wireless channels are considered here. Data gathering problem is designed as an optimization problem and the NP-hard problem is proved here. For both the broadcast tree and reverse multicast tree problems several heuristics were proposed.

Numerous sensor nodes are deployed densely in WSN. The sensor nodes forward the packets generated during the event detection to the base station. The neighboring nodes help in the transmission of packets to the sink. These packets are processed after they reach the base station. Thus reliable event detection requires reliable data transport in the WSNs. The reliable data deliveries in wired networking are not apt for TCP/IP protocols due to its inaccurate nature. [10]

In wireless sensor networks, the data aggregation needs to be processed based upon the energy and the reliability due to the following reasons:

- Operate the energy resources over a long period.
- Efficient energy is required for the sensor nodes in order to schedule their transmission strictly. Improper transmissions may lead to idle listening and overhearing causing energy wastage.
- The aggregation increases the amount of data concentrated in a single message which requires alteration in reliability.
- In event detection, the packets are transmitted from the sensor nodes to the base station and then to the neighboring which is possible through reliable data transport.

To overcome our previous certain issues on data aggression, we provide an efficient energy based and reliable data aggregation technique in the wireless sensor network. This technique is based on cluster formation and the loss ratios of the clusters are measured so that the energy consumption can be effectively reduced. Reliable transmission can be provided in the clusters using a coordinate node.

3. Design of RDAFP

3.1 Overview

In this paper, we propose to develop a data aggregation technique which is energy efficient and reliable. Initially a cluster is formed and the cluster head is selected based upon the cost value which is explained in section 3.2. The nodes in the cluster maintain a Neighbor information table (NIT) containing Node id, Distance and Cost. This NIT information is sent to the cluster head. Each cluster selects a coordinator node (CN) randomly in the network which is closer to the cluster and monitors the operations of the sensor nodes and commands them for specific operations. The cluster head aggregates the data and sends it to the CN.

The CN calculates the loss ratio which is the ratio of number of packets dropped and total packets broadcast from the source. Based upon the loss ratio, the cluster size can be modified and the forward node count of each node can be incremented or decremented which is explained in

section 3.4. Once the cluster size is changed, the CN gathers the information again from the cluster head compresses it and sends it to the sink.
Since the loss ratio is measured at the CN itself, the energy consumption can be effectively reduced. Also the reliability can be increased due to altering the cluster size before the data is transmitted to the sink.

3.2 Cluster Head Selection

- Initially the sensor nodes are arranged into clusters and the CN selects the cluster head for each cluster.
- Numbers of neighboring nodes M are determined by the CN based upon the node density.
- The sensor nodes transmit the M number of nearest neighbors to the CN.
- Received signal strength indicator (RSSI) estimates the distance to the nodes.
- K-theorem is used in each cluster by the CN to select the candidate set of cluster heads (S_{CH}).
- The request for the candidate set of cluster heads is sent by the CN and the sensor nodes reply their cost value (CV).
- Each candidate cluster head node calculate its own CV based on residual energy, and distance to coordinator node, and send it to CN. Calculation of cost value is explained in section 3.2.2.
- The coordinator node selects a node as cluster head among candidate set of cluster heads for each cluster based on CV. The higher the CV a node has; greater the chances of being cluster head. The CN confirms each cluster about their CH.

In this figure 1, we assume a cluster with three sensor nodes. The CN sends a request for the nodes in the cluster. The cost value of the sensor nodes 1, 2, and 3 are sent back to the CN. The node having higher CN becomes the Cluster head. Here the node 1 has CV value as 7 and it is elected as the CH for the cluster and the CN sends this information to other Clusters.

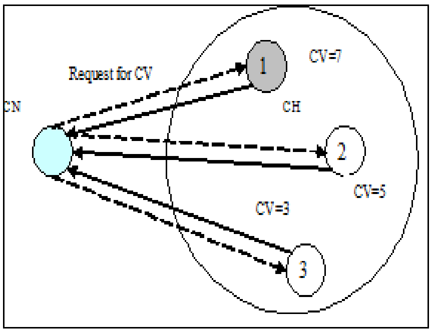

Figure 1: Selection of Cluster Head.

Selection of M for clustering:

In each cluster, value of M is set by the CN and this M value is relative to the node density and the ratio of cluster heads in a WSN.
The ratio should be below 0.50 ranging from 0.01 to 0.99. Many local optima can be obtained when the value of M is lesser. The number of best sensor nodes which are suitable for CH can be determined using M value. Optimal sensor node for cluster head can be selected by providing alternate suboptimal options by M.
M nearest neighbors is selected for each sensor node deployed in the cluster which is based on the distance. The received signal strength indicator calculates the distance between the sensor nodes. For larger distance, nearest neighboring nodes can be determined using multihop communication route. Energy consumption is less in choosing a neighbor in multi-hop connection when compared to the direct communication.
Every sensor node calculates its frequency of occurrence and minimum frequency required for a cluster to become a CH is also calculated. The weighted mean of frequencies is calculated and it is enhanced by adding 1 to it.
The product of each frequency of occurrence and number of sensor nodes having that frequency is calculated as weighted mean. The frequency value is rounded to its nearest integer. The value of Sensor nodes having frequency F or greater are identified and they become the candidates for cluster head (CH). The candidate cluster head nodes would always be equal to value of F [10].

Cost Value Calculation:
The Cost value (CV) is calculated based on following criterion:
Residual energy (E): The residual energy of a node preferably is greater than the approximate energy dissipated in previous round by the cluster head.
Distance to coordinator node (D) :The nodes having less distance from coordinator node should have higher probability to become cluster head. As energy consumption is directly proportional to the square of distance.

Cost value is based on the residual energy and the distance to the coordinate node. The cost is high, when the residual energy is high and the distance to the coordinator node is less.

$$CV = (a \times E) + (b \times \frac{1}{D}) \quad \ldots (1)$$

Where a and b are normalization constants.

3.2 Loss Ratio Calculations

Each node maintains a forward node count (C_{FN}), which denotes the broadcast or rebroadcast probability.
Initially C_{FN} [Nk] = C_{FN}min, for all nodes Nk, k=1, 2…….
C_{FN}min is the minimum number of forwarding nodes. Without loss of generality, we can assume that
C_{FN}min =1. The steps involved in the adaptive energy efficient forwarding phase are given below:

- Suppose N wants to send the collected data to the sink, it attaches its cost to the data packet and broadcast the packet to the nearest neighbors.
- When a neighbor N1 receives the packet from N, it first checks its cost is less than that of N. If it is less, it further forwards the packet. Otherwise it drops the packet, since N1 is not towards the direction of the sink.
- When the packet reaches the destination D, it measures the loss ratio (LR), which is the ratio of number of packets dropped and total packets broadcast from the source.
- Then D sends this LR value as a feed back to the source N.

When N receives this value, it checks the value of LR. It then modifies the value of C_{FN} as

$$C_{FN} = C_{FN} + \gamma, \text{ if } LR > LRmax. \quad \ldots (2)$$

Where γ is the minimum increment of decrement count and LRmax is the maximum threshold value of loss rate.

It then rebroadcast the data packets with the incremented C_{FN}, so that increasing the reachability of the sink. The total power required to reach the sink is thus calculated based on the cost field of all the nodes in C_{FN}. For example, if C_{FN} =4, then the minimum required power will be 4 * cost of each neighbor node in the NIT.

When the rebroadcast packets reach the destination D, it again calculates the loss ratio LR and sends back to N. It then reassigns the value of C_{FN}, depending on the value of LR. Once LR < LRmax , then

$$C_{FN} = C_{FN} - \gamma, \text{ until } C_{FN} >= C_{FN}min \quad \ldots.. (3) \ [11]$$

This data aggregation technique proves to be efficient in terms of energy and reliability since,

- Delay can be reduced due to that the loss ratio is measured in the CN itself. Measuring loss ratio at the sink causes high delay.
- Energy is reduced effectively when the size of the cluster is altered based upon the loss ratio.

The reliability can be maintained due to the change in the size of the clusters.

4. Simulation Results

4.1 Simulation Setup

We evaluate our RDAFP scheme through NS2 simulation [12]. We considered a random network deployed in an area of 500 X 500 m. The number of nodes is varied as 100,125,150,175 and 200. Initially the nodes are placed randomly in the specified area. The sink is assumed to be situated 100 meters away from the above specified area. The initial energy of all the nodes assumed as 8.1 joules. In our simulation, the channel capacity of mobile hosts is set to the same value: 2 Mbps. We use the distributed coordination function (DCF) of IEEE 802.11 for wireless LANs as the MAC layer protocol. The simulated traffic is CBR with UDP source and sink. All experimental results presented in this section are averages of five runs on different randomly chosen scenarios. The Table- I summarize the simulation parameters used.

4.2 Performance Metrics

We compare RDAFP with the DPC-EERP [13] scheme. We evaluate mainly the performance according to the following metrics.

- **Average end-to-end delay:** The end-to-end-delay is averaged over all surviving data packets from the sources to the destinations.

- **Average Packet Delivery Ratio:** It is the ratio of the number .of packets received successfully and the total number of packets transmitted

TABLE I SIMULATION PARAMETERS

No. of Nodes	25, 50, 75, 100, 125, 150, 175 and 200
Area Size	500 X 500
Mac	802.11
Simulation Time	50 sec
Traffic Source	CBR
Packet Size	512
Transmit Power	0.360 w
Receiving Power	0.395 w
Idle Power	0.335 w
Initial Energy	8.1 J
Transmission Range	75m

- **Drop:** It is the total number of packets dropped during the transmission.
- **Energy Consumption**: It is the average energy consumption of all nodes in sending, receiving and forward operations.

The simulation results are presented in the next section.

4.3 Simulation Results

In our first experiment we vary the number of nodes as 25, 50, 75, 100, 125, 150, 175 and 200.

The figure 2, we can see that the Energy consumption of proposed RDAFP is less than the existing DPC-EERP [13] protocol

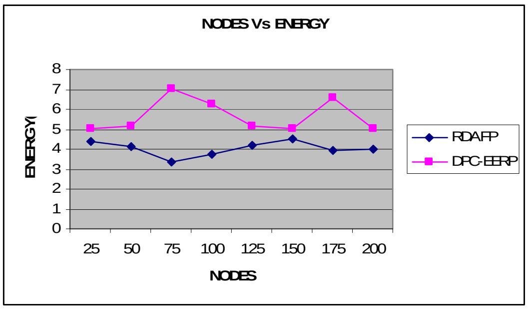

Fig 2: Nodes Vs Energy

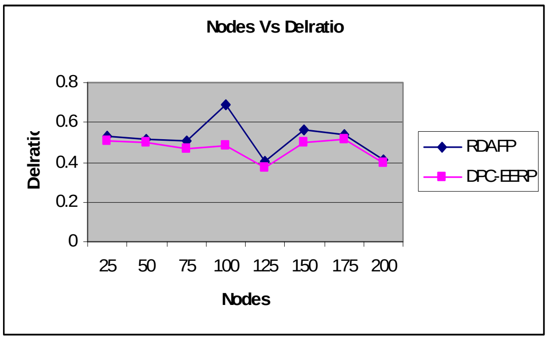

Fig 3: Nodes Vs Delivery Ratio

From the figure 3, we can see that the Delivery Ratio of our proposed RDAFP is higher than the existing DPC-EERP protocol.

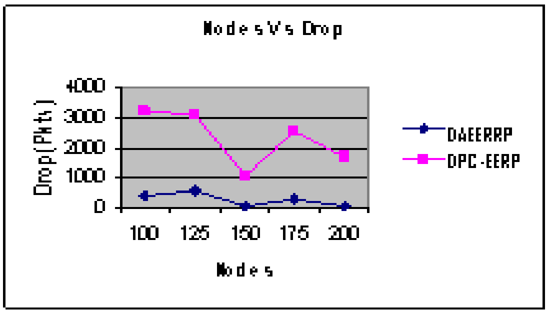

Fig 4: Nodes Vs Drop

From figure 4, we can see that the packet drop of our proposed RDAFP is lower than the existing DPC-EERP protocol by considering 100,125,150,175 and 200 nodes.

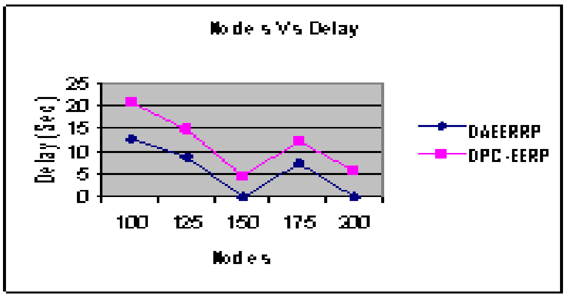

Fig 5: Nodes Vs Delay

From figure 5, we can see that the average end-to-end delay of our proposed RDAFP protocol is less than the existing DPC-EERP protocol obtained by considering 100,125,150,175 and 200 nodes.

4. Conclusion

In this paper, we have provided Reliable and efficient data forwarding and aggregation technique . Initially, the network is partitioned to various clusters .Each cluster selects a coordinator node (CN) randomly in the network which is closer to the cluster and monitors the operations of the sensor nodes and commands them for specific operations. In each cluster, the cluster head is selected based upon the cost value. The nodes in the cluster maintain a Neighbor information table (NIT) containing Node id, Distance and Cost. This NIT information is sent to the cluster head.. The cluster head aggregates the data and sends it to the CN. The CN calculates the loss ratio which is the ratio of number of packets dropped and total packets broadcast from the source. Based upon the loss ratio, the cluster size can be modified and the forward node count of each node can be incremented or decremented. Once the cluster size is changed, the CN gathers the information again from the cluster head compresses it and sends it to the sink. This technique proves to be efficient since the delay is reduced due to that the loss ratio is measured in the CN itself. Energy is reduced effectively and reliability is maintained when the size of the cluster is altered based upon the loss ratio. From our simulation results we prove that this technique is efficient in energy consumption and reliability.

References

[1] Nandini. S. Patil, and Prof. P. R. Patil "Data Aggregation in Wireless Sensor Network", IEEE International Conference on Computational Intelligence and Computing Research, 2010.

[2] Kevin Yuen, Baochun Li, and Ben Liang "Distributed Minimum Energy Data Gathering and Aggregation in Sensor Networks," *IEEE International Conference on Communications ICC' 06*, Vol. 8, pp . 3536-3541, June 2006.

[3] Yang Yu, Bhaskar Krishnamachari, and Viktor K. Prasanna "Energy-Latency Tradeoffs for Data Gathering in Wireless Sensor Networks," *proceedings of International journal of Sensor Networks*, Volume 4 Issue 1/ 2, pp. 48-54, July 2008.

[4] Kiran Maraiya, Kamal Kant, and Nitin Gupta "Wireless Sensor Network: A Review on Data Aggregation" International Journal of Scientific & Engineering Research Volume 2, Issue 4, April -2011.

[5] Ren P. Liu, John Zic, Iain B. Collings, Alex Y. Dong, and Sanjay Jha "Efficient Reliable Data Collection in Wireless Sensor Networks," in proceedings of *IEEE 68th Vehicular Technology Conference, VTC2008* , 2008. .

[6] Volker Turau and Christoph Weyer "Long-term Reliable Data Gathering Using Wireless Sensor Networks," *proceedings of 4th International conference on networked Sensing Systems INSS '07,* pp. 252-259, June 6-8 2007.

[7] Hemant Sethi, Devendra Prasad, and R. B. Patel "EIRDA: An Energy Efficient Interest based Reliable Data Aggregation Protocol for Wireless Sensor Networks," *in proceedings of International Journal of Computer Applications,* Volume 22– No.7, May 2011.

[8] David Gugelmann, Philipp Sommer, and Roger Wattenhofer "Poster Abstract: Reliable and Energy-Efficient Bulk-Data Dissemination in Wireless Sensor Networks," *in proceedings of SenSys'10*, November 3–5, 2010.

[9] Mingming Lu and Jie Wu "Utility-Based Data-Gathering in Wireless Sensor Networks with Unstable Links," *proceedings of the 9th international conference on Distributed Computing and Networking ICDCN' 08,* pp. 13-24, 2008.

[10] Muhammad Imran, Asfandyar khan, and Azween B . Abdullah "Energy Balancing Through Cluster Head Selection Using K-Theorem In Homogeneous Wireless Sensor Networks*", International Conference on Science & Technology: Applications in Industry & Education (2008).*

[11] Basavaraj.S. Mathapati, Dr. V. D. Mytri, "An Adaptive Energy Efficient Forwarding Routing Protocol for Wireless Sensor Network", *IEEE 2nd International Conference On Information Management and Engineering (ICIME) , China,* pp. 299-305, 2010.

[12] Network Simulator: http://www.isi.edu.nsnam/ns.

[13] Basavaraj.S.Mathapati,,Siddaram.R.pati,V.D.Mytri,"Distributed power control with Energy Efficient Routing Protocol for wireless Sensor Network", International Journal of Information Technology and knowledge Management ,2012.