Data Aggregation Algorithms in Wireless Sensor Network:A Survey

Shikha
Department of Electronics and communication, Punjab Technical University, India.

Priyanka Dayal
Department of Electronics and communication, Punjab Technical University, India.

Abstract – Wireless sensor networks (WSN) have huge range of applications such as traffic analysis, localization, monitoring environment, industrialized monitoring, and tactical systems. Wireless sensor networks are large-scale networks; these are expected to play an important role in future inhabitant and military application. In wireless sensor network, Data aggregation is a fundamental problem that had attracted great attention in recent years. Energy consumption between sensors, with the help of data aggregation, is well balanced to achieve the satisfactory network lifetime. In this paper, a survey on different data aggregation algorithms is presented. To overcome the limitations like low computation ability, limited memory size, limited energy resources, and the make use of insecure wireless communication channels and also achieves reliable data in wireless sensor networks.

Index Terms – Wireless sensor network, cluster head, Data Aggregation Algorithm, energy consumption.

1. INTRODUCTION

Wireless sensor networks consists of large number of small sensing self-powered nodes that are deployed for the purpose to gather information (data) or sense some special actions and communicate in a wireless approach , with the end target of handling their processed data to a destination i.e. base station. Three key elements are sensing, processing and communication whose combination in one small device gives rise to a wide number of applications [2]. Low powered devices also known as smart sensors that are equipped with one or more sensors, a processor, memory, a power supply, a radio, and an actuator. Since the sensor nodes contain less memory and are typically deployed at those locations where these are impossible to deploy manually, a radio is implemented for wireless communication to transport the data (sensed information) to a base station. To determine the properties of the environment, a range of mechanical, biological, optical, thermal, chemical and magnetic sensors might be connected to the sensor node. WSNs are likely to play even more significant role in the next generation networks to sense the physical world.

WSN nodes are prone to failure due to less energy resources, hardware failure, errors in communication link, malicious attack, and so on. Thus, the reliability and precision of individual sensor node's evaluation is very crucial. Many different protocols have been developed or investigated to ensure secure data aggregation for WSN throughout the past couple of years. Many researchers have also addressed the possible cooperation among sensor nodes in data gathering and aggregating. However, sensor nodes are constrained in energy, computing ability, and bandwidth. Therefore, innovative routing techniques are main considerations that are used for the reduction of these constraints. Routing in WSN is very challenging and has distinguishing characteristics compared to other wireless networks.

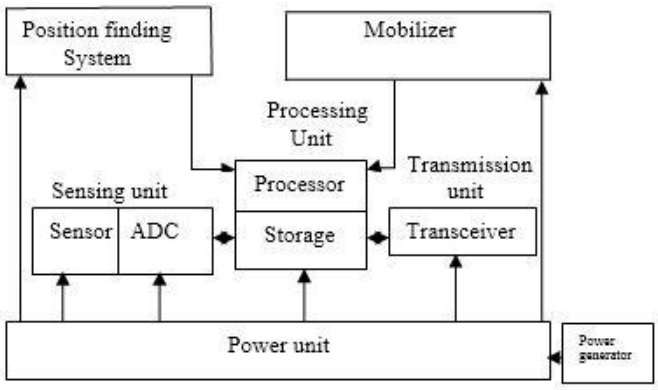

Figure. 1 The Components of sensor node

Figure. 1 shows a sensor field and the sensor node's components [1]. A sensor node is composed of four basic components: sensing unit, processing unit, transceiver unit, and power unit. Each sensor node has the potential to sense, collect and route data back to the base station (BS). Data can be routed back to the BS by using various routing protocols or data aggregation protocols. The BS may communicate with the user using available Internet connection [1].

As energy consumption is a main limitation in wireless sensor network. To enhance the lifetime of the network has to efficiently use the energy management by using data aggregation.

- The fundamental operation in such applications is data gathering, i.e., to gather sensing data from the sensor nodes and transmit to a base station for processing. The critical issue in data gathering is to conserve sensor energy and maximizing sensor lifetime and makes it accessible to the sink in an energy efficient manner with low data latency.
- Data aggregation, as a typical operation in data gathering applications, in which an intermediate node could first collect data from its preceding nodes, process the received data and aggregate it (e.g., the highest temp), and then forward the aggregated data to its parents (succeeding) nodes.

In this paper, focus was on data aggregating routing techniques. These are categorized as 1) tree based 2) cluster based data aggregation routing techniques for transmitting data from all the deployed sensor nodes in sensor field to the BS. Tree based routing techniques are those in which tree is the combination of different nodes. Data transmission takes place between parent nodes and children nodes. And in cluster based techniques are those in which nodes make clusters and cluster heads and data is transferred from cluster heads to base station.

The remainder of the paper is organized as follows. In section II different data aggregation algorithms will be introduced, in section III comparison between the entire data aggregation algorithm will be introduced. The simulation results will be given in section IV. Finally, section V concluded the paper.

2. DATA AGGREGATION ALGORITHM IN WSN

Some data aggregation algorithms are explained below, which are used for reduction of data coming from different nodes. It is mainly used for reducing data and efficiently data transmission from normal nodes to the base station or sink with minimum energy consumption or minimum delay.

2.1 EDGE (Efficient data gathering protocol)

Edge [3] is the data gathering protocol which supports the limited resources as flooding and periodic updating of routing packets are avoided by it. EDGE is a tree-based topology rooted at a sink. Upon node failure or adding of new nodes, the tree created by EDGE will be reconstructed. EDGE is a multipoint-to-point approach designed for a source s and a destination d, where $s \epsilon \{S\}$ and $d \epsilon \{B\}$, namely, every sensor tries to transmit sensed data to the sink. In this protocol every new node is added to the tree by making requests and replies. Firstly through base station broadcasting of a child request (CRQ) packet is initiated. A non-member decides on its parent from received CRQ and while receiving different requests from different member of tree it will decide on their metric which member is best then it will send its reply (CRP) to that particular member. After reply to member will accept the child (CAC) then while joining child will broadcast request to parent and same process will undergo. In this way this protocol works. This protocol gives better results than the direct diffusion as it can deliver all of its packets. Delay is less as congestion in DD (Direct Diffusion) at sink due to flood messages in the network contention increases resulting packet loss. But in EDGE tree topology is used where new route found and tree is reconstructed whenever required.

2.2 Spanning tree algorithm for data aggregation

Another algorithm in [5] for data aggregation is spanning tree based on estimates cost. It's a well-known method for reducing energy consumption in the networks. Sensor should be a component of the tree, that is, an internal or leaf node, so that it can communicate with the sink. A spanning tree algorithm is based on cost estimated that consider the residual energy and approximate available degree of aggregation as key factor. This technique optimizes the creation of data aggregation tree. Tree is constructed by nodes and sink node is consider as the parent node for all nodes and different children make routes with their parent node to the sink. Nodes will form route with minimum cost so that energy consumption will be less and the lifetime of the network will extended.

2.3 LEACH (Low-Energy Adaptive Cluster Hierarchy)

LEACH is based on cluster topology data aggregation algorithm [9]. The most important goal of LEACH is to have a cluster heads to reduce energy cost of transmitting data from normal nodes to distance base station. Firstly nodes will organize in cluster with one cluster head. In this data aggregation process is done mainly by cluster heads where data is aggregated which is collected from other nodes which are present in that particular cluster and transmit data in one-hop or multi-hop to base station. Initially on the basis of probability p a node decides to be a CH and broadcasts its decision. The selection of CH is done on the basis of residual energy of a node. Each non-CH nodes determines its cluster by choosing appropriate CH using the least communication energy. The approach provides a balancing of energy consumption by random rotation of Cluster Heads. It forms clusters on the basis of received signal strength and CH nodes are considered as routers to the base-station.

LEACH provides the following energy saving key areas:

- No overhead is wasted throughout the process of decision making of which node will become cluster head.
- CDMA, as different code is assigned to each cluster, used during transmission information in clusters.
- Minimum transmission energy of each node is calculated to communicate with its cluster.

In LEACH, changing the CH is probabilistic; there is a good chance for a node with very less energy gets selected as a CH. When CH dies, that particular cluster becomes worthless.

LEACH also forms one-hop intra- and inter cluster topology where each node can transmit directly to the CH and aggregated data is transmitted to the base station.

2.4. PEGASIS (Power efficient gathering in sensor information system):

Power efficient gathering in sensor information system (PEGASIS) [7] prove to be more robust than the LEACH, since cluster formation data from a node is transmitted to any one of its 1-hop neighbor node where it is aggregated and passed on in the similar manner. PEGASIS proved to be more reliable and efficient, than LEACH in many real time applications. It is also known as chain based data aggregation algorithm. It is considered as near-optimal algorithm. PEGASIS algorithm selects cluster by transmitting testing signal with the energy decreases progressively. The optimal path is optimized by CH in relation to the position. Comparing with LEACH protocol, the PEGASIS's ability of data aggregation is stronger, and also the energy loss of PEGASIS is less.

2.5 PEDAP (Power–Efficient Data Gathering and Aggregation Protocol):

In this paper two algorithms are proposed under the name of Power –Efficient Data Gathering and Aggregation Protocol (PEDAP) [11]. These are data aggregation schemes based on near optimal minimum spanning tree. These two algorithms are same but one is the power-aware version of the other. First, while the tree starts to build, a node is selected as base station and at iteration the minimum weighted edge from a vertex in the tree is selected and new vertex gets added to the tree. Here, the newly added node will transmit its data through the indicated edge. This process is repeated until all nodes get attached to the tree. In this work, the sensors are in fixed location. These are in direct communication range and can transmit to and receive from the base station. The sensors periodically sense data from environment and send to base station in round basis. Before sending to base station, the nodes aggregates their data with other data received from other nodes. This proposal shows that it can save much energy and much improved than LEACH and PEGASIS protocols.

2.6 DEDA (Delay-minimized Energy-efficient Data Aggregation Algorithm):

In this paper [4] author proposed a data aggregation technique DEDA. DEDA is a distributed, energy efficient technique for gathering data from sensor nodes with minimum delay. As energy-efficiency and delay are crucial issues in wireless sensor networks. All the sensor nodes are deployed in the field as they sense data and transmit it to the base stations. Its structure is in form of tree, its root is cluster head (CH) and other nodes are considered as cluster members (CM). Data link is formed between CH and base station (BS) directly. In this approach firstly one or many data aggregation trees are built, network structure is form. Main is to join any two same sized clusters together to form one bigger cluster. The process is repeated until anymore clusters cannot join with any other same sized cluster, then these clusters form connections with the BS directly. Based on this approach, the final network will consists of clusters of different sizes. Then straight forward scheduling algorithm is applied. Each node is assigned time slot to transmit sensed data by its rank. Every parent node's children have different ranks corresponding to different transmitting time slots. Multiple receptions thus could not happen at parent node, as all clusters have different sizes, their corresponding different ranks. Hence, these CH's also broadcast data to the BS in different timeslots. By constructing a delay-efficient network structure minimum delay is achieved. Energy consumption is saved thus an acceptable network lifetime could be obtained.

2.7 L4DAS (Long- lifetime and Low-latency Data Aggregation scheduling algorithm):

In this L4DAS [6], author proposed the algorithm called long lifetime and low latency data aggregation scheduling. In this data aggregation tree is constructed. Firstly, the field is break up into clusters with the diameter equal to transmission range r and then in each cluster spanning tree is constructed. After that a global tree over the clusters constructed and connects global tree to the spanning tree in each cluster. In every cluster there is a representative means a cluster head which is used further for constructing global tree. If some representatives are not able to connect with each other, some connecting nodes are chosen to connect them. Then nodes are being scheduled in aggregation tree according to scheduling scheme i.e. maximum interference priority scheduling scheme. At each time slot, nodes are scheduled as senders if an earlier time slot they are leaf-nodes that are not being scheduled or an intermediate node who's all children nodes have been scheduled and also choose node which interferes with number of receivers while transmitting as maximum interference priority is considered. Finally, through the simulation and comparisons, this proves that protocol outperforms then other schemes.

2.8 Bilayer-based Data Aggregation Algorithm:

Energy efficiency is the most important issue in wireless sensor networks. The main objective is to enhance the

Data Aggregating algorithm	**Energy efficient**	**Hierarchy**	**Cluster head**	**Data transmission**	**Enhancement in Lifetime**	**Topology**	**Delay**	**Better results than other algorithm**
EDGE [3]	YES	NO	NO	MULTI-HOP	—	TREE	LESS	DIRECT DIFFUSION (DD)
Spanning tree algorithm for data aggregation [5]	YES	—	—	MULTI-HOP	YES	TREE	—	—
LEACH [9]	—	YES	YES	ONE HOP	YES	STAR	—	—
PEGASIS [7]	YES	YES	YES	MULTI-HOP	YES	CHAIN	—	LEACH
PEDAP [11]	YES	—	—	MULTI-HOP	—	TREE	—	LEACH, PEGASIS
DEDA [4]	YES	YES	YES	MULTI-HOP	YES	MESH, STAR, TREE	LESS	LEACH, PEDAP (SOMEHOW)
L4DAS [6]	YES	YES	YES	ONE-HOP	YES	STAR, TREE	LESS	—
Bilayer-based data aggregating algorithm [8]	YES	YES	YES	MULTI-HOP	YES	STAR	LESS	—

Table 1: Comparison between data aggregation techniques

lifetime of the networks by utilizing bilayer-based data aggregation GA optimized [8]. In this scheme, we divide the wireless network to two layers (bi-layer) and each layer has numerous of cluster heads optimized by GA, nodes near that cluster head will make cluster with that head. In the detection region of each layer transmit data to the related head which is optimized by GA.

And these cluster heads are responsible for data aggregation and further transmission to the base station. As the simulations clarified that the algorithms can decrease the amounts of transmitted data and system delay. And energy is also saved in this algorithm as compared to other.

3. COMPARISON BETWEEN DATA AGGREGATION ALGORITHM

In this paper some data aggregation algorithms are observed that can be used in wireless sensor networks are shown in Table 1. By comparing these algorithms the metrics involve are saving energy, robustness, hierarchy, lifetime, delay. Different data aggregating techniques in WSN are summarized in Table 1.

Some of the algorithms have hierarchy based routing where as others have both hierarchy as well as tree-based routing algorithms. Cluster-based routing protocols combine sensor nodes to efficiently relay the sensed data to the sink. Since sensor nodes are energy constrained device, almost all of the discussed algorithms focused on energy consumption and targeted to achieve balance energy consumptions.

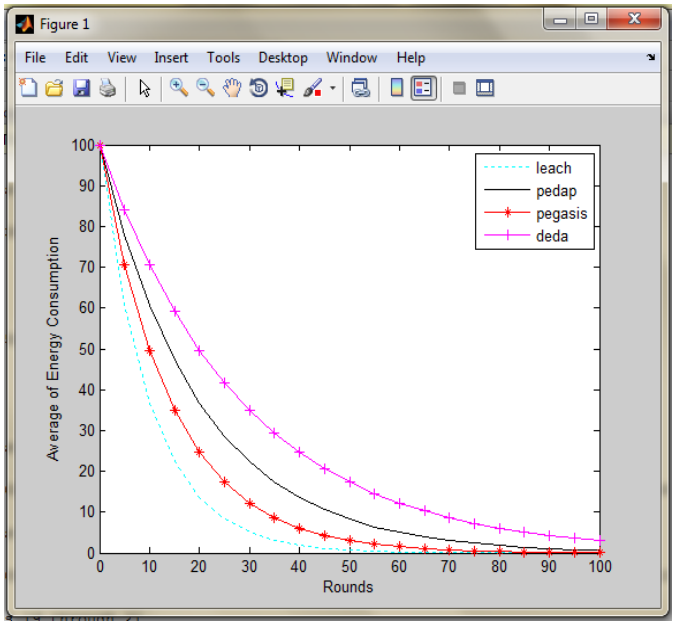

Figure. 2 Average of WSN energy consumption in different data aggregation algorithms.

4. SIMULATION ANALYSIS

In order to compare different data aggregation algorithms, the performance of algorithms are evaluated in form of their lifetime and energy consumption. LEACH algorithm as the basis of cluster-based algorithm in order to reduce the energy consumption by creating clusters and cluster head. Compared with PEGASIS which is the chain based algorithm, PEDAP is optimal spanning tree based algorithm and DEDA algorithm in which one or multiple data aggregation techniques are built and same sized clusters will merge to form bigger cluster. From Figure 2 PEGASIS is better than LEACH and PEDAP gives better result than both PEGASIS and LEACH. And algorithm is somehow better than all of three (LEACH, PEGASIS, PEDAP).

5. CONCLUSION

Data aggregation in sensor network is a challenging area of research. This paper concludes comprehensive survey of different data aggregation algorithm in wireless sensor network. The most common trend in these algorithms is to reduce the energy consumption and increase the lifetime of the network. Each and every discussed algorithm has their advantages and limitations. The use of above discussed data aggregation algorithms are according to application. Different algorithms discussed in this paper are suitable for different applications according to requirement in that particular area. For multi-hop transmission DEDA data aggregation algorithm has better lifetime response than other algorithms.

REFERENCES

[1] Akyildiz, I., Su, W., Sankara subramaniam, Y. and Cayirci, E. "A Survey on Sensor Networks. IEEE Communication Magazine", 40, 102-114,2002.

[2] Daniele Puccinelli and Martin Haenggi, "Wireless Sensor Networks: Applications and Challenges of Ubiquitous Sensin", IEEE CIRCUITS AND SYSTEMS MAGAZINE, 2005.

[3] Niwat Thepvilojanapong, Yoshito Tobe, Kaoru Sezaki, " On the Construction of Efficient Data Gathering Tree in Wireless Sensor Networks", circuits and systems, ISCAS IEEE International symposiumon, 2005.

[4] Huu Nghia Le, Vyachelav Zalyubovskiy, Hyunseung Choo, " Delay-minimized energy-efficient data aggregation in wireless sensor networks.", international conference on cyber-enabled distributed computing and knowledge discover, 2012.

[5] Jie Shao, Ning Ye, "A Spanning Tree Algorithm for Data Aggregation in Wireless Sensor Networks", Proceedings of the 7th World Congress on Intelligent Control and Automation, Chongqing, China, June 25 - 27, 2008.

[6] Zhengyu Chen, Geng Yang, Lei Chen and Jin Wang, "An algorithm for data aggregation scheduling with long-lifetime and low-latency in wireless sensor networks" international journal of future generation communication and networking vol. 5, no. 4, December, 2012.

[7] Mohammad Abdus Salam, Ttanjima Ferdous, "Tree- based data aggregation algorithms in wireless sensor networks: a survey" proceedings of the 2012 international conference on industrial engineering and operations management Istanbul, Turkey, July 3 – 6, 2012.

[8] Tian Bao Wang, Chengdong Wu, Peng Ji, Jian Zhang, "Bilayer-based data aggregation algorithm for low latency in wireless sensor networks" IEEE national nature science foundation , 2010.
[9] Olutayo Boyinbode, Hanh Le, Audrey Mbogho, Makoto Takizawa, Ravi Pilah, "A survey on clustering algorithms for wireless sensor networks" 13th international conference on network-based information systems, 2010.
[10] Mousam Dagar and Shilpa Mahajan, "Data aggregation in wireless sensor network: a survey" international journal of information and computation technology ISSN 0974-2239 volume 3, number 3 pp. 167-174, 2013.
[11] Korpeoglu, I., and Tan, H., "Power efficient data gathering and aggregation in wireless sensor networks." *ACM SIGMOD Record* 32.4, pp. 66-71, 2003.
[12] Chen Tian, Hongbo Jiang, Chonggang Wang, Zuodong Wu, Jinhua Chen, and Wenyu Liu, ''Neither Shortest Path Nor Dominating Set: Aggregation Scheduling by Greedy Growing Tree in Multihop Wireless Sensor Networks'', IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 60, NO. 7, SEPTEMBER 2011.
[13] D. Chinh Hoang R. Kumar S. Kumar Panda, ''Optimal data aggregation tree in wireless sensor networks based on intelligent water drops algorithm'', IET Wireless Sensor Systems, 17th May 2012.
[14] Junchao Ma, Wei Lou, Xiang-Yang Li, ''Contiguous Link Scheduling for Data Aggregation in Wireless Sensor Networks'', IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 25, NO. 7, JULY 2014.
[15] Joohwan Kim, Xiaojun Lin, Ness B. Shroff, Prasun Sinha, ''Minimizing Delay and Maximizing Lifetime for Wireless Sensor Networks With Any cast'', IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 18, NO. 2, APRIL 2010.
[16] Sunil R. Gupta, Dr. N.G. Bawane, Dr. Sudhir Akojwar, ''A Clustering Solution for Wireless Sensor Networks based on Energy Distribution & Genetic Algorithm'', 2013 Sixth International Conference on Emerging Trends in Engineering and Technology.

Authors

Shikha is a PG student in the Deptt. of Electronics and Communication at Punjab Institute of Technology (PTU Main Campus), Kapurthala, Punjab, India. She obtained B.tech (Electronics and Communication) from Punjab Technical University. Her research area interest includes wireless communication networks.

Priyanka Dayal is Assistant Professor in the Department of Electronics and Communication Engineering at Punjab Institute of Technology, Kapurthala, and Punjab. She obtained her B.Tech from PTU and M.Tech from LPU. Her research area interest includes VLSI Design and wireless communication.