A Review on Data Aggregation

[1]Krati Rastogi, [2]Kaushik Ghosh
[1,2]Dept. of the Computer Science of Mody University, Laxmangarh (Sikar), Rajasthan, India

Abstract
A Wireless Sensor Network (WSNs) is a network consisting of various nodes in the network. The application of this network is in various fields such as in health monitoring, environmental management, security and military, etc. Constraint for the network of wireless sensor nodes is the size of node as it is very small and has limited processing capability with low battery power. Collection of data from the nodes and sending to the other node is the main aim of the wireless sensor network. It basically collects information from various nodes and sends the collected data from source to the sink node. Wireless sensor network has been used to transmit data to base station from among the collected data of various sources it has sensed for decision making method. Transmitted information is collected to the base station which may act as a station for data aggregation that stores all the collected data and transmit further for more operations. With the help of data aggregation we can reduce the energy consumption by eliminating the redundant data, the implosion and overlap of data among the nodes in the network. For era of development in the field of the improvement in the energy of wireless sensor networks (WSN) data aggregation is an excellent technique. This paper reviews about the techniques of data aggregation efficiently for the energy conservation in the network.

Keywords
Wireless Sensor Network, Data Aggregation, Architecture, Energy Conservation.

I. Introduction

A Wireless Sensor Network (WSN) is a spatially distributed network; it is a wireless network which comprises of large number of nodes distributed in the network, self-organizing and, low powered devices called motes also known as sensor nodes . The most popular services in industrial and commercial applications of WSN because of its technical improvement in processor, communication, and usage of low power fixed computing devices. Sensor nodes are used to check environmental conditions like temperature, pressure, humidity, sound, position etc WSN naturally consist a large number of spatially dispersed, small, battery-operated, embedded devices that are networked to considerately collect, process, and convey data to the users, and it has controlled computing and processing capabilities. Motes are the small computers, which work cooperatively to form the networks. Motes used in WSN as energy efficient, to perform multi-functions. The requirements for motes in industrial applications are extensive. For performance improvement used to make links with each other in the network. Motes can communicate with each other using transceivers. In WSN the number of sensor nodes can be in the order of hundreds. In comparison with sensor networks, Ad Hoc networks will have less number of nodes without any infrastructure.

A. Data Aggregation

In wireless sensor networks, sensor nodes are generally resource-constrained and battery-limited. In order to save resources and energy, data must be aggregated to pass up overwhelming amounts of traffic in the system. Problems related to implosion and overlap can be solved using data aggregation. Data should be aggregated in WSN to reserve energy and to prevent burst traffic in communication network. This would reduce the data packet transmission to the sink. This is a must to eliminate redundancy of data transmission and enhance energy of node in wireless sensor network. The intend of data aggregation is to eliminate redundant data transmission and enhances the lifetime of energy in wireless sensor network. Data aggregation is the method of one or several sensors nodes then collects the uncover result from other sensor.

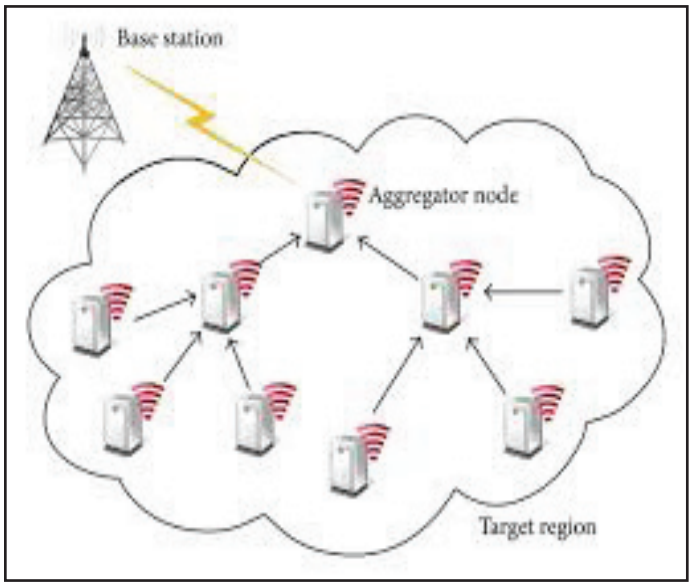

Fig. 1: Data Aggregation

II. Impact of Data Aggregation in WSN.

In this paper we tend to discuss the two main factors that have an effect on the performance of data aggregation in wireless sensor network, strategies like energy saving and delay. Data aggregation is that the method, during which aggregating the data packet returning from the various sources, the amount of transmission is reduced. With the existence of this method we tend to save the energy within the network. Delay is that the latency connected with aggregation knowledge from nearer sources might have to power back at intermediate nodes so as to combine them with knowledge from supply that square measure farther away. Basically aggregation technique supported the position of the sources within the network, variety of sources and also the network topology. If the examine the factors, we tend to think about the two models of the supply placement. The event radius model and random supply model. The modeling says us that wherever the supply square measure clustered close to one another or located at random, vital energy gains square measure attainable with knowledge aggregation. These gains square measure greatest once the number of sources is massive, and once the sources square measure located comparatively near one another and much from base station. The modeling through, conjointly looks to the recommend that aggregation latency may not be negligible.

III. Data Aggregation Based Network

A. Flat Network

In flat networks, each node typically plays an equivalent role and sensing element nodes collaborate along to perform the sensing task. Due to the large number of such nodes, it is not possible to assign a worldwide symbol to every node. This consideration led to central routing, wherever the basestation sends queries to bound regions and waits for knowledge from the sensors settled within the chosen regions. Since knowledge is being requested through queries, attribute-based naming is important to specify the properties of information. Early works on knowledge central routing, e.g., SPIN and directed diffusion was shown to save lots of energy through knowledge negotiation and elimination of redundant data. These two protocols actuated the planning of the many alternative protocols that follow an analogous idea.

1. Data Diffusion

Data diffusion may be a well-liked data aggregation paradigm for wireless device networks. It is a data-centric and application aware paradigm, within the sense that each one info generated by sensing element nodes is termed by attribute-value pairs. Such a theme combines the data returning from all totally different sources en-route to the sink by eliminating redundancy and minimizing the quantity of transmissions. throughout this suggests, it saves the energy consumption and can increase the network period of time of WSNs. throughout this theme typically base station broadcast the message to the interested provide node. Afterward each node receives interest. These interests define the attribute value like name of object. Each node get the interest can cache it for later use. As a result of the interest is broadcasted by the network hop by hop, gradient area unit setups to draw info satisfying the question toward the requesting node. A gradient is also a reply link to the nearer from that the interest was received

2. SPIN

SPIN is an adaptive routing protocol, which transmits the knowledge initial by negotiating. As described earlier, transmission of knowledge consumes additional energy. To cope up with this downside SPIN create use of information of the particular information to be sent. Assume a node must send a perceived image file it initial generates the information for image, and this information is broadcast. Information can contain the outline of the message that the node needs to send. The particular information are transmitted providing the node desires to receive it. For this purpose SPIN makes use of three messages specifically, 1. ADV 2. REQUEST 3. Data. ADV: Before causing a message, a node initial generates the descriptor of the message to be sent. This information is changed by creating use of ADV message. ADV message informs the scale, contents and needs of the message. This helps the receiving node on deciding transmission of the message. When receiving the ADV message receiver node verifies the descriptor whether or not the message may be a duplicate and whether or not receiver node's battery capabilities area unit enough to transmit the info. If the node is fascinated by information, it replies with asking message to the sender node. DATA: If the sender node receives asking message, it starts the particular transmission of knowledge by creating use of knowledge message. This is often the particular information transfer part

B. Hierarchical Network

Hierarchical routing protocols are referred to as cluster-based routing, planned in wireless networks. They are well known techniques having special blessings associated with quantifiability and economical communication. The concept of hierarchical routing is additionally utilized to perform energy efficient energy economical routing in WSNs. In a hierarchical design, higher energy nodes are often accustomed method and send the knowledge whereas low energy nodes are often accustomed perform the sensing within the proximity of the target. This implies that creation of clusters and assigning special tasks to cluster heads will greatly contribute to overall system quantifiability, lifetime, and energy efficiency. Ranked routing is associate economical thanks to lower energy consumption among a cluster and by performing knowledge aggregation and fusion so as to decrease the quantity of transmitted messages to the basestation. Hierarchical routing is principally two-layer routing wherever one layer is employed to pick cluster heads and therefore the alternative layer is employed for routing. However, most techniques during this class don't seem to be regarding routing, rather on "who and when to send or process/aggregate" the knowledge, channel allocation etc., this may be orthogonal to the multihop routing perform.

1. Cluster based

These Wireless device network is resource constraint that's why device cannot directly transmit information to the bottom station. Within which all regular sensors will send information packet to a cluster head (local aggregator) that aggregates information packet from all the regular sensors in its cluster and sends the succinct digest to the bottom station. With the assistance of the theme we have a tendency to save the energy of the sensors. LEACH: Low energy reconciling clustering has been projected to organize a device network into a collection of clusters so the energy consumption is often event distributed among all the device nodes.

2. Chain based

In which every sensor sends information to the nearer neighbor. Power- economical Data-Gathering Protocol for detector data Systems (PEGASIS) is kind of chain based mostly data aggregation. In PEGASIS, all sensors area unit structured into a linear chain for information aggregation. The nodes will type a sequence by using a greedy algorithmic program or the sink can decide the chain during a centralized manner. Within the Greedy chain formation assumes that all sensors have inclusive data of the network. The farthest node from the sink initiates chain formation and, at every step, the nearest neighbor of a node is selected as its successor within the chain. In every data-gathering spherical, a node receives data packet from one amongst its neighbors, aggregates the info with its own, and sends the aggregates information packet to its different neighbor on the chain. Eventually, the leader node within the area unit like cluster head sends the aggregate information to the base station.

3. Tree based

In which all node organized in kind of tree suggests that they belong to Hierarchical routing protocols, with then facilitate of intermediate node we are able to perform information aggregation method and information transmit leaf node root node. Tree based mostly information aggregation is appropriate for applications that involve in network data aggregation. One in every of the most aspects of tree-based networks is that the construction of tree based network is energy efficient data-aggregation tree.

IV. Proposed Methods of Data Aggregation

Kalyan Sasidhar et.al.[4]"A WSN lifetime improvement algorithm reaping profit of information aggregation and position transitions", Wireless sensor system consist of various subsystems such as sensing, transmission, reception, power and processing systems. The life of sensor nodes is one of the important factors to consider in a wireless sensor system. The algorithm basically combines data aggregation and state transition to increase the network lifetime. Bhaskar Krishnamachari et al. proposed data-centric routing and its performance with traditional end-to end routing schemes. The authors also discuss about the communication network density on the energy costs and impact of source destination placement and delay related with data aggregation. They also find out the complexity of optimal data aggregation. The authors planned and modeled performance of data aggregation in resource-constrained distributed event-based system. A method was proposed by Y. E. Massad et al[10], a linear distributed algorithm to aggregate data and consumes energy in a regular way and pretends it for different scenarios. The proposed method is distributed, there is no sensor synchronization and there is total absence of comprehensive knowledge of the wireless sensor network. The absence of comprehensive knowledge differentiates this scheme from others as PEGASIS being a significant contribution. K. Sruthi et al. [6] describes Enhanced Link aware Clustering Mechanism, a data aggregation algorithm based on passive clustering method that focuses both on the efficient CH and GW selection process and diverse transmission power during the data aggregation process. Passive clustering concentrates on the link condition and state of the nodes. Swapna B. Sasi et al. [8] shows the aim provide the locked communication in the wireless sensor networks. For that, several cryptography using optimization algorithms is investigated. Numerous optimization algorithms are discussed for cryptography to make the keys solution for the encryption. Al-Karaki et al. [9] suggest method to find the least number of aggregation points in order to maximize the system lifetime. Shah and Rabaey proposed to set sub-optimal path rarely in order to increase the lifetime of the system. These paths are select by means of a probability method, it depends on the energy utilization of each path. System survivability is the main metric that the approach is concerned with. This approach argues that using the least amount energy path all the time will reduce the energy of nodes on that path. Instead, one of the multiple paths is used with an influenced probability so that the whole network lifetimes increase. Schurgers et al. suggested a slightly changed version of Directed Diffusion, called Gradient-based routing (GBR). The idea is to keep the quantity of hops when the interest is diffused through the network. Hence, each node can find out the least number of hops to the sink, as they are called height of the node. Younis et al. [10] works for diverse hierarchical routing algorithm based on three-tier architecture. Sensors are arranged into clusters earlier to network operation. The algorithm employ cluster heads, namely gateways, which are less energy reserved than sensors and assumed to know the location of sensor nodes. Gateways preserve the states of the sensors and sets up multi-hop routes for collecting sensors data. Subramanian and Katz [11] proposed self-organizing protocol but develop taxonomy of sensor applications as well. Based on such classification, they have suggested architectural and infrastructural components essential for building sensor applications. The architecture supports heterogeneous sensors that can be mobile or stationary. Some sensors, which can be moreover stationary or mobile, look into the surroundings and forward the data to selected set of nodes that beehive as routers. Router nodes are stationary and form the backbone for interaction. The data which is collected are forwarded through the routers to stronger sink nodes. Each sensing node should be available to a router node in order to be part of the network. Ganesan et al. [17] proposed employing multiple paths in advance so that in case of a crash of a path, one of other paths is chosen without any cost for finding for another one. There is of course extra overhead of keeping these another paths alive by using low data rate, which will definitely use further energy but more energy can be saved when a path fails and a new path should be selected.

V. Advantages and Disadvantages of data Aggregation

A. Advantages

1. Data aggregation process we can improve the robustness and correctness of informative data which is obtained by entire network, certain redundancy exists in the data collected from sensor nodes consequently data fusion processing is needed to reduce the redundant information.
2. It reduces the traffic load and conserves energy of the sensors.

B. Disadvantage

The cluster head send fuse these data to the base station .this cluster head or aggregator node may be attacked by cruel attacker. If a cluster head is compromised, then the base station (sink) cannot be sure of the correctness of the aggregate data that has been send to it. These performances were very much dependent on the beloved application.

Energy Efficiency by the data-aggregation technique, we can increase the functionality of the wireless sensor network. During which every sensor nodes should have spent the equal amount of energy in each and every data gathering round. A data aggregation scheme is energy efficient if it maximize the functionality of the system. Network lifetime, data accuracy, and latency are some of the considerable measures of data aggregation algorithm.

VI. Data Aggregation Function in Wireless Sensor Network

Many efficient type of data aggregation function is needed in wireless sensor network. These functions are closely connected to sensor network application. Such as mean quintile, medium, count, average, max, and min.

A. Duplicate Sensitive and Duplicate Insensitive

Aggregation function may be average and minimum, maximum. If we use average function, it takes as a duplicate sensitive and minimum function is taking as duplicate insensitive function. Data aggregated in the network on that point of time, same data consider several times. We can used duplicate function then the final result depends on the number of the times and that value is considered otherwise aggregation function is said to be duplicate and considered to be insensitive.

B. Lossy and Lossless

Data packet knows how to be aggregated with the help of lossy aggregation or by lossless aggregation. Lossy aggregation approach does not follow an ideal reconstruction but lossless aggregation ensures a complete recovery of all individual sensor data at base station (sink) [2].

VII. Conclusion

In this paper we have studied the data communication in sensor networks i.e. data aggregation and we realize how transfer among sensor nodes is different from other wireless sensor networks. Wireless sensor networks are energy controlled system. Since most of the energy consumed for sending and receiving data, the process of data aggregation becomes an important matter and optimization is needed. Well-organized data aggregations not only provide energy preservation but also remove redundancy data and hence provide useful data only. When the data from source node is send to sink through surrounding nodes in a multi hop network by dropping transmission and receiving power, the energy used is low as compared to that of sending data directly to sink that is aggregation reduces the data transmission without aggregation. Our scheme integrates energy efficient and data storage mechanism. It also shows that these techniques not only reduces power consumption but also prolongs the lifetime constraint. That is why lifetime of the network is limited so the diverse approach or protocol has been proposed for increasing the lifetime of the wireless sensor network.

Referneces

[1] Chandrakasan, R. Amirtharajah, S.-H. Cho, J. Goodman, G.Konduri, J. Kulik, W. Rabiner, A. Wang,"Design Considerations forDistributed Microsensor Networks," Proc. IEEE Custom Integrated Circuits Conference (CICC), pp. 279-286, 1999.

[2] M. Dong, K. Yung, W. Kaiser,"Low Power Signal Processing Architectures for Network Micro sensors," Proc. Int'l Sump. Low Power Electronics and Design, pp. 173-177, 1997.

[3] V. Vaidehi,"Secure Data Aggregation in Wireless Sensor Networks", Computing for Sustainable Global Development (INDIACom), pp. 2179 – 2184, 2015.

[4] Fei Yuan, Yiju Zhan, Yonghua Wang,"Data Density Correlation Degree Clustering Method for Data Aggregation in WSN", IEEE SENSORS, pp. 29-35, 2014.

[5] Rabia Noor Enam,"An adaptive data aggregation technique for dynamic cluster based Wireless Sensor Networks", 2014 23rd International Conference on Computer Communication and Networks, pp. 1 – 7, 2007.

[6] Manish V. Bhosle,"An Energy Efficient and Reliable Location wise Data Aggregation In WSN", International Conference on Computing Communication Control and Automation, pp. 45-50, 2015.

[7] Nandini. S.,"Research in Data Aggregation in Wireless Sensor Network", 2010 IEEE International Conference on Computational Intelligence and Computing Research, pp. 21-30, 2010.

[8] Kiran Maraiya,"Wireless Sensor Network: A Review on Data Aggregation", International Journal of Scientific & Engineering Research, Vol. 2, Issue 4, pp. 67-72, 2011.

[9] Yuling Lei, Yan Zhang; Yanjuan Zhao,"The Research of Coverage Problems in Wireless Sensor Network", pp. 31-34, IEEE, 2009.

[10] Yonghui Shim, Younghan Kim,"DataAggregation with Multiple Sinks in Information-Centric Wireless SensorNetwork", IEEE Conf. on Data Aggregation, pp. 13-17, 2014.

[11] J. Rabaey, J. Ammer, J.L. da Silva Jr., D. Patel,"Pico Radio: Ad-hoc Wireless Networking of Ubiquitous Low-Energy Sensor/Monitor Nodes", Proc. IEEE Computer Society Annual Workshop on VLSI, 2000.

[12] Estrin, R. Govindan, J. Heidemann, S. Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," Proc. ACM/IEEE Mobicom, Aug. 1999.

[13] C. Intanagonwiwat et al.,"Directed Diffusion for Wireless Sensor Networking," IEEE/ACM Trans. Net., Vol. 11, No. 1, Feb. 2002, pp. 2–16.

[14] S. Lindsey, C. Raghavendra, K. M. Sivalingam, "Data Gathering Algorithms in Sensor Networks using Energy Metrics," IEEE Trans. Parallel Distrib. Sys. Vol. 13, No. 9, Sept. 2002, pp. 924–35.

[15] H. Luo et al.,"Energy Efficient Routing with Adaptive Data Fusion in Sensor Networks," 3rd ACM/SIGMOBILE Wksp. Foundations of Mobile Comp., Cologne, Germany, Aug. 2005.

[16] H. Albert, R. Kravets, I. Gupta,"Building Trees Based On Aggregation Efficiency in Sensor Networks," Med-Hoc-Net 2006, Lipari, Italy, June 2006.

[17] D. Ganesan et al.,"Highly resilient, energy efficient multipath routing in wireless sensor networks", Mobile Computing and Communications Review 5 (4), pp. 11–25, 2002.

[18] Sasidhar, Sreeresmi R, Rekha P,"A WSN lifetime improvement algorithm reaping benefits of data aggregation and state transitions", Global Humanitarian Technology Conference - South Asia Satellite, pp. 201-205, 2014.