

### Abstract

Data aggregation is very essential relevance practices in wireless sensor network. On accordingly by the help of data aggregation we minimize the energy utilization by terminating the redundancy. When the wireless sensor network distributed systematically in remote areas or hostile environment and the wireless sensor network have most challenging task is a life time so with help of data aggregation we can enhance the lifetime of the network .In this paper we discuss the data aggregation approaches based on the routing protocols and the algorithm in wireless sensor network. We also discuss the merits and demerits or various performance measures of the data aggregation in the network.

**Keywords:** Base Station, Cluster, Data aggregation, Network Lifetime, Routing, Structure, Tree, Wireless sensor network.

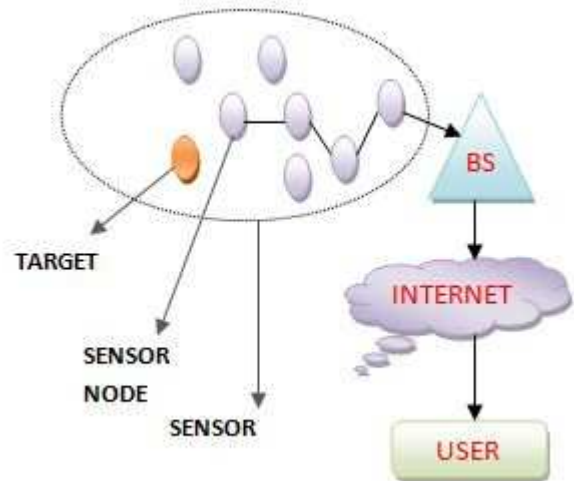
### Introduction

The wireless sensor network is also known as ad-hoc network and it consists of small light weighted wireless nodes called sensor nodes, distributed systematically in environmental or physical condition. They used to measure physical parameters such as pressure, sound, humidity and temperature. These sensor nodes are distributed systematically in large or thousand numbers and collaborate to form an ad hoc network capable of reporting to [base station] data collection sink. Wireless sensor network have various purposes like building monitoring, habitat monitoring, target tracking, health monitoring and military survival lance.

In accordingly the wireless sensor network is a reserve supply constraint if we analyse about memory, energy, computation and limited communication susceptibility. All sensor nodes in the wireless sensor network are Act towards with each other or by intermediate sensor nodes.

Sensor nodes that generate data based on its mechanism of sensing the observation and transmit sensed data packet to the [sink] base station. This action basically direct transmission since the base station may locate very far away from the sensor nodes requirements and they need more energy to transmit data over long distances so that a better technique is to have fewer nodes sends data to the base station [sink]. These are the

nodes called aggregator nodes and actions called data aggregation in wireless sensor network.



**Figure 1 Structure of the Sensor network**

### Clustering in WSN

Sensor node are densely distributed systematically in wireless sensor network that means physical environment would produce very similar data in close by sensor node and transmitting Such type of data is more or less redundancies. So all these facts Contribute to the progress using some kind of grouping

of sensor nodes such that group of sensor node can be combined or compress data together and transmit only data in a compact manner. This can minimize the localized traffic in distinct group and also minimize global data. This grouping action of sensor nodes in a densely distributed systematically large scale sensor node is known as clustering. The way of compress data and combing data belonging to a single cluster called data fusion (aggregation).

**Consequences of clustering in wireless sensor network:-**

1. How many sensor nodes should be taken in a unique cluster and selection procedure of cluster head in different cluster?
2. Heterogeneity in a network it means user can put some power full nodes in term of energy in the network which can behave like cluster head and cluster work contains simple node as a cluster member only. Many of the algorithm and protocols have been introduced which deal with each distinct issue.

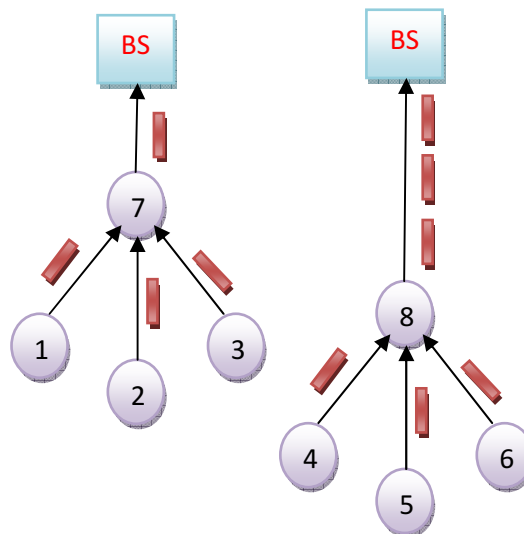
**Data Aggregation**

In typical wireless sensor networks the sensor nodes are usually reserve supply-constrained and battery-limited. In order to save reserve supplies and energy, data must be aggregated to avoid overwhelming amounts of traffic occurring in the network. There has been extended work on data aggregation methods in sensor networks. The scope of data aggregation is that to eliminate redundancy data transmission and enhances the lifetime of energy in wireless sensor network. Data aggregation is the action of one or several sensors then collects the detection result from other another sensor. Then the collected data must be processed by sensor to diminish transmission burden before they are transmitted to the base station [sink]. The wireless sensor network has consisted three different types of nodes.

1. Simple regular sensor nodes
2. Aggregator node
3. Queried.

The Simple regular sensor nodes sense data packet from the environment, send to the aggregator nodes initially these aggregator nodes gather the data from multiple sensor nodes of the network and aggregates the data packet using a some aggregation function like sum, count, average, max min, then sends the aggregate result to upper queried node or the aggregator node who generate the query. It can be the base station or else sometimes an external user who has permission to interact with in the network. Data transmission between aggregators, sensor nodes and the queried consume a lot of energy in wireless sensor network. Fig. 2 contain two types of models in which one of them is data aggregation

model and another one is non data aggregation model in which sensor nodes like 1, 2, 3,4,5,6 are regular nodes that collecting data packet and reporting them back to the upper nodes in which sensor nodes like 7, 8 are aggregators that perform aggregating and sensing at the same time.



**Fig 2: Data aggregation and non data aggregation models.**

In these aggregation model 4 data packet travelled within the network and only one data packet is transmitted to the base station [sink]. Then other non data aggregation model also 4 data packet travelled within the network and all data packets are sent to the base station [sink], means we can able to explain that with the assist of data aggregation action we can decrease the number of data packet transmission. They can also save energy in the sensor node which occurs in the wireless sensor network. On accordingly with the help of data aggregation we can enhance the lifetime of wireless sensor network. Sink have a data packet with energy in accuracy manner with minimized data latency. So the data latency is very important in many applications of wireless sensor network such as health monitoring, environment monitoring where the freshness of data is also an important factor. It is very crucial to develop energy efficient data aggregation algorithms so that network lifetime is enhanced. There are so many factors which analyse the energy efficiency of a wireless sensor network, such as the data aggregation mechanism, network Structure and the underlying protocol for routing techniques. Wireless sensor network has distributed processing of sensor node contains data. Data aggregation is the technique in which it describes the processing method that is applied on the data received by a sensor node as well as data is to be routed in the entire network. In which minimize energy utilization of the

sensor nodes and also minimize the number of transmissions or length of the data packet. In network aggregation is the exclusive action of collecting and routing information through a multi hop wireless network, by processing the data packet with help of intermediate sensor node. The objective of this technique is increasing the life time of the network and also minimizes reserve supply for utilization.

There are two types of techniques for in network aggregation.

1. With size reduction.
2. Without size reduction.

In network aggregation with size reduction, the action in which combine and compressing the data received by a sensor node from its neighbours in order to minimize the length of data packet to be circumstance a node receives two data packets which have a correlated data. In these conditions it is useless to send both data packets. Then we apply a function like MAX, MIN and AVG then it again send single data packet to base station [sink]. With help of these approaches we minimize the number of bit transmitted in the network and also save a lot of energy. In the network aggregation without size reduction is defined in the action of data packets received by different neighbours in to a single data packet but without processing the value of data. This action also minimizes energy utilization or increase life time of the network.

#### **Merits and Demerits of Data aggregation in wireless sensor network**

**Merits:** With the help of data aggregation action we can enhance the robustness and accuracy of information which is obtained by entire network certain redundancies exists in the data collected from sensor nodes thus data fusion processing is needed to minimize the redundancies information. Another merit is those minimizes the traffic load and conserve energy of the sensors.

**Demerits:** The cluster head means data aggregator nodes send fuse these data to the base station .this cluster head or aggregator node may be attacked by deliberately harmful attacker. If a cluster head is compromised, then the base station cannot be ensure the correctness of the aggregated data has been send to it. Another consequences is existing systems are several copies of the aggregate result may be sent to the base station [sink] by uncompromised nodes .It increase the power consumed at these nodes.

#### **Performance measure of data aggregation**

There are very important performance measures on data fusion algorithm techniques. These performances are highly relying on the desired application.

**Energy Efficiency:** By the data-aggregation scheme maximize the functionality of the wireless sensor network. Every sensor nodes must have spent the same

amount of energy in every data together round. A data aggregation scheme is energy efficient if it maximizes the functionality of the network. Data accuracy, Network lifetime and latency are some of the significant performance measures for algorithm in data-aggregation. The concise explanation of these measures is highly relying on the desired application.

**Network lifetime:** The network lifetime is defining the number of rounds in data fusion. Till the particular percentage of the total nodes dies and the percentage depend on the application .If we talk about some application simultaneously working of the all the sensor nodes is crucial hence the lifetime of the network is number of round until the first nodes which improves the energy efficiency of nodes and enhance the lifetime of whole network.

**Communication overhead:** They analyse the complexity in communication of the network fusion algorithm.

**Latency:** Latency is evaluate data of time delay experienced by system that means data send by sensor nodes and received by base station [sink]. Initially delay involved in routing, data transmission and data aggregation.

**Data accuracy:** It is to evaluate the ratio of total number of reading received at the base station (sink) to the total number of generated. There are different type's data aggregation protocols like network-flow-based data aggregation protocols, network Structure based data aggregation protocols and quality of service (QOS)-aware Data aggregation protocols designed to guarantee QOS metrics. Here network Structure based protocols are described in detail.

#### **Impact of data aggregation in wireless sensor network**

In this paper we discuss the two main factors that affect the performance of data aggregation methods in wireless sensor networks, Such as delay and energy saving. The action of data aggregation in which aggregating the data packet coming from the different sources; the number of transmission is minimized. With the help of this action we can save the network's energy. The Delay is called latency linked together with aggregation data from closer sources may have to held back at intermediate nodes in order to combine them with data from source that are greater extent away. Initially methods of aggregation based on the position of the sources in the number of sources, network and topology of the network. If we analyze the factors can consider the two models of the source placement. The event radius (ER) model and random source model [14]. The modelling says us that where the source is clustered located near randomly or each other important energy gains are possible with aggregation of data. These gains are highest when the number of source is very large and when the sources are located relatively close to each

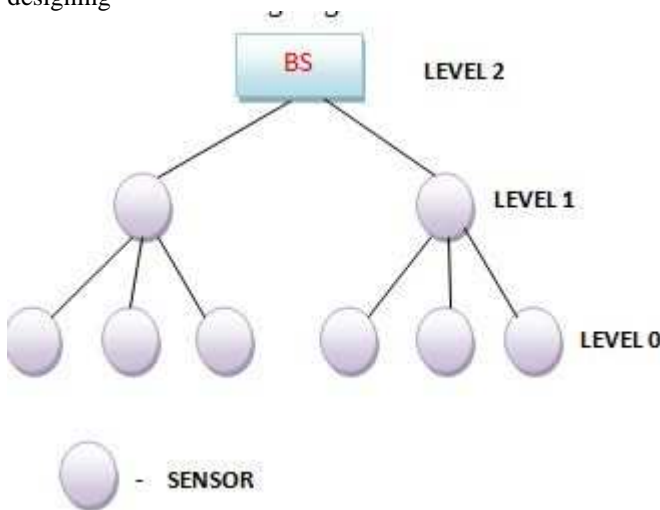
other and far from base station [sink]. The modelling also seems to suggest that aggregation latency could be non negligible.

**Data Aggregation Approaches In Wireless Sensor Network**

Data aggregation action is performed by specific protocol for routing. Our Scope is aggregating data to minimize the energy utilization. So sensor nodes should route packets based on the data packet content and choose the next hop in order to promote aggregation in the network. Initially routing protocol is separated by structure in the network so that the routing protocols are based on the considered approaches.

**Tree-Based Approach**

The tree based approach is defining aggregation from constructing a tree in aggregation form. The aggregation form of tree is minimum spanning tree and considering the sink node as a root and considering source node as a leaves. Flowing information of data start from leaves node up to root means sink (base station). Demerits of this approach, as we know like wireless sensor network are not free from failure .in case of data packet loss at any level of tree and the lost in data will be not only for single level but for whole related to the sub tree as well. These approaches are suitable for designing



**Fig: 3 Data aggregation in Tree based for wireless sensor network**

optimal techniques for aggregation. Data centric protocol knows as [TAG] Tiny aggregation approach. The working of TAG is depending on two ways of phases,

1. Distributed phase
2. Collection phase.

**Cluster-Based Approach**

In energy-constrained sensor networks avail in large size and is inefficient for sensors to transmit the data directly to the sink in such scenarios. Cluster based approach is also known as hierarchical approach. This approach deals entire network is divided in to clusters in several. These clusters have a cluster head which is picked among cluster members. Cluster heads perform the role of aggregation in which they aggregate data received from cluster members locally and then transmit the result to base station (sink). Recently, several cluster-based network organization and data-aggregation protocols in wireless sensor network have been proposed. The Fig: 4 shows a sensor network in cluster-based the cluster heads can communicate with the sink directly via long range transmissions or multi hopping through other cluster heads.

The arrows indicate wireless communication links proposed a maximum lifetime data aggregation (MLDA) algorithm which finds data gathering procedure provided location of sensors node and base-station, energy and data packet size of each sensor node. Then data gathering procedure denotes how data packet are collected from sensors node and transmitted to base station for each and every round. Scheduling can be assumed as a combination of aggregation trees. In [4], they proposed heuristic-greedy clustering-based MLDA based on algorithm of MLDA. They partitioned the network in to cluster and referred super-sensor for each cluster. Then they analyse maximized lifetime procedure for the super-sensors and then use this procedure to construct aggregation trees for the sensors. Phase I of Two Phase Clustering [TPC] scheme creates clusters with a cluster-head and each node within that cluster form a direct connects with cluster head. Cluster Head rotation in phase-I is localized and is done based on the remaining energy level of the sensor nodes which minimize time variance of sensors and this lead to energy saving from unnecessary rotation in cluster-head. Each node in phase II are within the cluster searches for a neighbour closer than cluster-head which is called data relay point and setup up a data relay link.

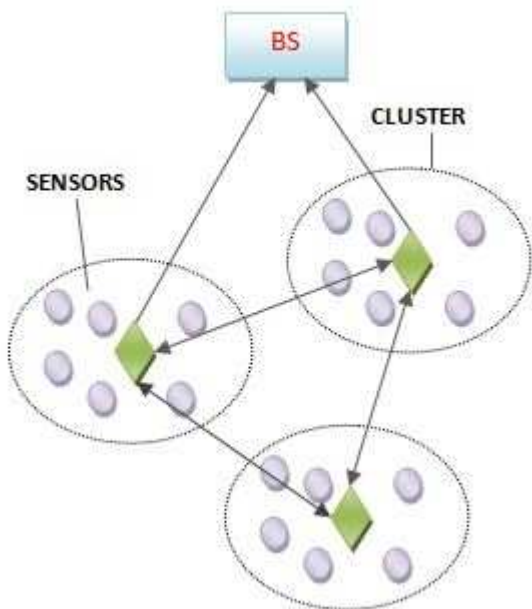


Fig: 4 Sensor network in Cluster based.

Now the sensor nodes within a cluster either use direct link or data relay link to send their data to cluster head which is an efficient scheme in energy. The relay points in data aggregates data at forwarding time to another cluster-head or data relay point. In case of high density in network and TPC phase II will setup unnecessary data relay link between neighbours as closely distributed systematically sensor will sense same data and this lead to a waste of energy.

**Multi-path Approach:**

The drawback of tree based approach is the limited robustness avail in the system. For overcoming this drawback, the new approach was proposed by many inventors in that they sending some aggregation data to aggregation tree which has single parent node. Node can send data over many types of paths. On these all node can send data packets to its possibly multiple neighbours. Hence data packet flow from source node to the sink node along many way path, then lot of negotiate node between source node to sink node so aggregation done in every negotiate node. By applying this approach we make this system robust but some overhead. Example of this approach like ring topology in which network is separated in to circle in concentric with explaining level levels according to hop distance from sink and introduce a new technique have both issues : robustness and energy efficiency. The single path is to establish connection each node to the base station it is energy saving but high risk of failure in link. On the other head approach of multipath would require more nodes to participate with

consequent wastage in the energy. The clever usage of multi-path only when there is loss of packet which is implemented by smart caching at sensor nodes data. On many implemented situation data may be gathered only from a specific region. These different approaches rely on a spanning tree and provide alternative paths only when it was not functioned well is detected. By adopting algorithm a tree based approaches for forwarding packets on the network. During the ideal situation when failures not occur and this will be certainly the best choice because the minimal number of nodes are already in use for this phase in transmission. On the presence of node failures or link, the algorithm will discover paths in alternative manner. So as to aware the dispatch of many packets by possible within the certain constraint in time. The issues in this techniques is that it may cause the arising of hot spots and nodes along preferred paths will consume their energy reserve supplies quickly, possibly causing disconnection in the network.

**Hybrid Approach**

Hybrid approach followed between tree, multipath scheme and cluster based. The data aggregation structure can adjusted according to specific network situation and to some performance statistics.

| Protocols/ algorithms  | Tree | Cluster | Multipath | Hybrid |
|------------------------|------|---------|-----------|--------|
| TAG                    | ✓    | -       | -         | -      |
| Directed Diffusion     | ✓    | -       | -         | -      |
| PEGASIS                | ✓    | -       | -         | -      |
| DB-MAC                 | ✓    | -       | -         | -      |
| EADAT                  | ✓    | -       | -         | -      |
| LEACH                  | -    | ✓       | -         | -      |
| Cougar                 | -    | ✓       | -         | -      |
| Synopsis               | -    | -       | ✓         | -      |
| Diffusion              | -    | -       | ✓         | -      |
| Tributaries and Deltas | -    | -       | -         | ✓      |

Table 1: Various approaches for routing protocol

**Data Aggregation Function in Wireless Sensor Network**

Many effective type of data aggregation function in wireless sensor network is needed. These functions are closely related to network applications sensor. Such as mean medium, count, quantile, average, min and max.

**Duplicate sensitive and duplicate insensitive**

Aggregation function may be average and minimum. If we use average function , it take as a

duplicate sensitive and minimum function is take as duplicate insensitive function in WSN. Aggregation of data in the network on that time same data consider many number of times. By using duplicate function then the final result depends on the number of the times and same value has been considered otherwise aggregation function is said to be duplicate insensitive.

#### **Lossy and lossless**

Data packet can be aggregated with the help of lossless aggregation or by lossy aggregation. The approach of Lossy aggregation does not follow a perfect reconstruction but lossless aggregation ensures a complete recovery of all distinct sensor data at base station [sink].

#### **Data Representation in Wireless Sensor Network**

Data representation is the effective way to the data representation. The Wireless sensor network consists of a large number of small sensor nodes. These are reserve supply constraint, due to limited reserve supply constraint it needs to decide whether to store, compress, discard and transmit data. This entire requirement wants a suitable way to represent the information any type of structure are common to all sensor node in the network.[14]

#### **Security Issues in Data Aggregation for Wireless Sensor Network**

There are two type of securities in data aggregation are require for wireless sensor network. Listed as,

1. Confidentiality.
2. Integrity.

The issue in security is data confidentiality to protecting the sensitive passive attacks and data transmission like eavesdropping. On discuss the hostile environment so data confidentiality is mainly used because wireless channel is vulnerable to eavesdropping by the method of cryptography. The Difficult is to analyze encryption and decryption operations like modular multiplication and modular division. The issues in security is data integrity with the help of integrity we minimize the compromised sensor source nodes or aggregator nodes from significantly altering the final value in aggregation. In sensor network the sensor node is compromised easily. The nodes which are compromised have a capable to discard or modify messages.

#### **Method of secure data aggregation:**

They having two type of method for securing data hop by hop encryption and end to end encryption, both methods follow some step.

1. Encryption action has to be done by sensing nodes in wireless sensor network.

2. Decryption action has to be done by aggregator nodes.

3. After that aggregator nodes aggregates the result and then encrypt the result again.

4. The sink node gets final aggregated result and decrypt it again.

#### **Conclusion**

In this paper we present wireless sensor network is consist a large number of sensor node. And these nodes are reserve supply constraint. That's why lifetime of the network is limited so the various approaches or protocol has been proposed for increasing the wireless sensor network lifetime. We discuss the data aggregation are one of the important techniques in this paper for enhancing the network life time. Also discuss the various techniques for data aggregation or also discuss the merits and demerits and various performance measures of the data aggregation.

#### **References**

- [1] T. Arampatzis, and S. Manesis "Applications Survey in Wireless Sensors and Wireless Sensor Networks" In Mediterranean Conference on Control and Automation MED '05, Nicosia, Cyprus, 2005.
- [2] P.Havinga and S.Chatterjea "Dynamic data aggregation technique for WSNs " Research Program on Circuits and Integrated Systems, Veldhoven, Netherland, Nov. 2003.
- [3] S.B.Wicker and A.J.Gold smith "Design Challenges for Energy-constrained Wireless Sensor Networks in Ad Hoc" Communications in IEEE Magazine.
- [4] L. Gattani, G. Lo Re, and M. Ortollani, "Efficient and Robust Data Gathering for Wireless Sensor Networks" in Proceeding of the 39th Hawaii International Conference on System Science in 2006
- [5] B. Hong, V.K. Prasana, "Optimizing System lifetime for Data Gathering in network sensor systems" Workshop on Algorithms for Wireless and Ad-hoc Networks (A-SWAN), August 2004, Boston.
- [6] K. Kalpakakis, K. Dasgupta and P. Namjoshi, "Efficiency Algorithm for maximum lifetime data gathering and aggregation in wireless sensor network" Computer Network, vol. 42, August 2003, pp.697-716.
- [7] Ramesh Rajaghobalan and Pranhod K. Varshney, "A Survey: Data Aggregation Techniques in sensor network" Computer

Science and Electrical Engineering. Paper – 22,  
2006

- [8] N. Shrivastaava, C. Buragohain, D. Aghrawal, and S. Soori, "Medians and beyond: new aggregation techniques for sensor network" Proceeding of the second international conference on Embedded networked sensor system. New York, ACM Press, 2004.
- [9] W. Zhanng and G. Caaao, "Group rekeying for filtering false data in sensor networks: A pre distribution and local collaboration based approach" IEEE March 2005.