# A Survey on Energy Efficient Secure Data Aggregation in WSNs

**Tanushree K N[1], Prasanna Kumar M[2]**

[1] M. Tech Scholar, Department of Computer, East West Institute of Technology, Bangalore, India

[2]Assistant Professor, Department of Computer Science, East West Institute of Technology, Bangalore, India

**Abstract:** ***Wireless Sensor Networks (WSNs) are special kind of ad-hoc networks, having abilities of sensing, processing and wireless connectivity contains hundreds or thousands of sensor nodes. Sensor nodes are usually constrained in energy, communication, storage, and computation capability. Due to the main characteristics of resource-constrained and battery-powered sensors in wireless sensor networks, energy consumption is always a major concern. WSNs usually generate large amounts of raw data in which there exists high redundancy. So, it is important to develop efficient data processing technique to reduce redundant data and the amount of transmission. Aggregation is an essential technique to reduce data redundancy and the communication overhead, also minimize the energy consumption and prolong network lifetime. Security is an important criterion to be considered because wireless sensor nodes are deployed in a remote or hostile environment area that is prone to attacks easily. So data aggregation and security are essential for WSN. In this paper we discuss the data aggregation approaches and advantages and disadvantages of data aggregation and secure data aggregation mechanisms in the network.***

**Keywords:** Data aggregation, Encryption, Sensor nodes, Wireless Sensor Networks. Energy Efficiency

## 1. Introduction

Wireless Sensor Network (WSN) is a collection of hundreds or thousands of, low cost, low-power wireless nodes called sensor nodes, deployed in physical or environmental condition for monitoring and collecting information from the surroundings. The data collected from sensor nodes should be transmitted to sink, wastes too much energy, since the base station may locate very far away from sensor nodes needs. More energy is required to transmit data over long distances so that a better technique is to have fewer nodes sends data to the base station. These nodes are called aggregator nodes and processes called data aggregation in wireless sensor network. However the sensor nodes may be deployed in remote and hostile environments where attackers may inject false information or forge aggregation values without being detected. Hence it is a challenging task to protect sensitive information transmitted by wireless sensor networks Thus security issue becomes an important research field in data aggregation for WSNs.[1][2]

## 2. Data Aggregation

In typical wireless sensor networks, sensor nodes are normally constrained of resource and battery-limited. In order to save resource and energy, data must be aggregated to avoid over whelming amounts of traffic in the network. The main aim of data aggregation is that eliminates redundant data transmission and enhances the lifetime of energy in wireless sensor network Data aggregation is a process of aggregating the sensor data using aggregation approaches. The general data aggregation algorithm are LEACH (low energy adaptive clustering hierarchy), TAG (Tiny Aggregation) and it is shown in the fig1 etc, where the data from the sensor aggregates the data, the aggregated data is transfer to the sink node by selecting the efficient path. Many effective type of data aggregation function are needed in wireless sensor network. These functions are very much related to sensor network application. Such as mean quintile, medium, count, average, max, and min. [3]

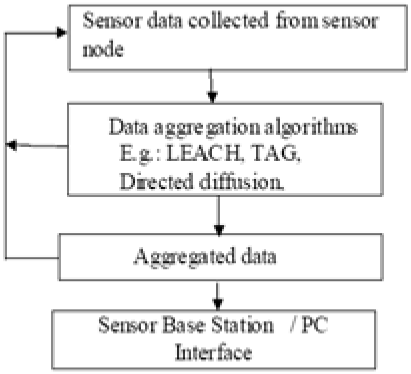

Sensor data collected from sensor node
Data aggregation algorithms E.g.: LEACH, TAG, Directed diffusion.
Aggregated data
Sensor Base Station / PC Interface

**Figure 1:** General architecture of the data aggregation algorithm

## 3. Why Data Aggregation

The main aim of secure data aggregation scheme is to maintain data privacy for each node in the WSNs. And also the scheme must consider the performance of efficiency, accuracy, and scalability. Hence a desired secure data aggregation should meet the following criteria.

**1) Efficiency:** data aggregation can achieve bandwidth efficiency through in-network processing. In integrity protecting private data aggregation schemes, additional communication overhead is unavoidable in order to achieve additional features. However, the additional overhead must be kept as much as small.
**2) Data privacy:** privacy concern is one of the important obstacles to civilian applications for wireless sensor networks. Curious individuals may attempt to get more detailed information by eavesdropping on the communications of their neighbors. It is increasingly

important to develop data aggregation schemes in order to ensure data privacy against eavesdropping.

**3) Data integrity:** Usually data aggregate results may be used to make critical decisions; a base station should give guarantee of the integrity of the aggregated result before accepting it. Hence, it is difficult that data aggregation schemes can protect the aggregated results from\ being polluted by attackers.

**4) Accuracy:** An accurate aggregate result of sensed data is generally desired. Hence , we should take accuracy as a criterion to evaluate the performance of integrity-protecting private data aggregation schemes. When accurate aggregate results are required, schemes based on randomization techniques could not be applicable.

**5) Scalability**: the cheap sensor nodes are prone to attack, which makes WSNs dynamic network changes. When some nodes fail or new nodes are deployed, it is very much needed for the secure data aggregation scheme to continue to be implemented correctly. A good secure data aggregation scheme needs to have easy and good scalability.

## 4. Advantage and Disadvantage of data aggregation in WSN

Data aggregation gives several advantages.

- It mainly reduces the number of transmission thereby improving the bandwidth and energy utilization in the network.
- Redundancy exists in the data collected from sensor nodes thus data fusion processing is needed to reduce the redundant information
- Another advantage is that it reduces the traffic load and conserves energy of the sensors
- Data aggregation process enhance the robustness and accuracy of information which is obtained by entire network.
- Data aggregation has several disadvantages also.
- The cluster head means data fusion nodes send fuse these data to the base station .this cluster head or fusion node may be attacked by malicious attacker.
- If a cluster head is compromised, then the base station (sink) cannot be ensure the correctness of the fusion data that has been send to it.
- Another drawback in existing systems are several copies of the fusion result may be sent to the base station (sink) by uncompromised nodes .It increase the power consumed at these nodes.

## 5. Approaches for Data Aggregation

There are many types of aggregation techniques are present some of them are listed below.

- Centralized
- Decentralized
- Tree based
- Cluster based
- Chain based
- Grid based

### 5.1 Centralized Approach

Centralized approach is an address centric approach using a multi-hop wireless protocol each node sends data to a central node via the shortest possible route. The sensor nodes simply send the data packets to a leader node. The leader node is the strong node. The leader aggregates the data which can be queried. Each intermediate node has to send the data packets addressed to leader from the child nodes. Hence a large number of messages have to be transmitted for a query in the best case equal to the sum of external path lengths for each node.

### 5.2 Decentralized Approach

In the decentralized approach, centralized node is not be maintained. Data fusion occurs locally at each node on the basis of local observations and the information obtained from neighboring nodes..The advantage of this architecture is scalable and tolerant to the addition or loss of sensing nodes or dynamic changes in the network.

### 5.3 Tree-Based Approach

In the tree based approach, aggregating the data from constructing an aggregation tree. The form of tree is minimum spanning tree, sink node is taken as a root and source node is taken as a leaves. The flow of data start from leaves node up to root means sink (base station).Disadvantage 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, the data will be lost not only for single level but for whole related sub tree as well.

### 5.4 Cluster-Based Approach

In cluster-based approach, whole network should be divided in to several clusters. Each cluster has its own cluster-head which is selected among cluster members. Cluster heads perform the role of aggregator which aggregate data received from cluster members locally and then transmit the result to sink.

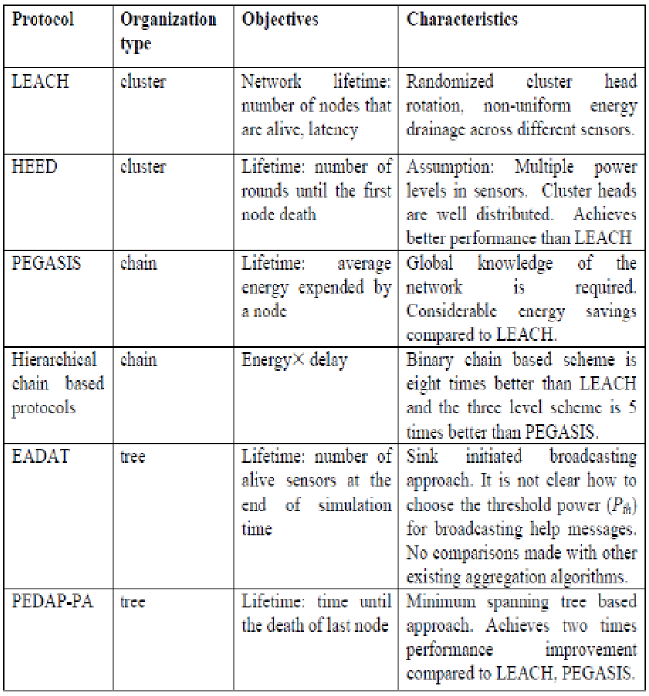
### 5.5 Chain based approach

In chain based approach, each sensor sends data to the closer neighbor. All sensors are structured into a linear chain for data aggregation. The nodes can form a chain by employing a greedy algorithm or the sink can decide the chain in a centralized manner. In the Greedy chain formation assumes that all sensors have inclusive knowledge of the network. The farthest node from the sink initiates chain formation and, at each step, the closest neighbor of a node is selected as its successor in the chain. In each data-gathering round, a node receives data packet from one of its neighbors, aggregates the data with its own, and sends the aggregates data packet to its other neighbor along the chain. Eventually, the leader nodes are similar to cluster head sends the aggregated data to the base station.

### 5.6 Grid based approach

In which a set of sensors is assigned as data aggregators in fixed regions of the sensor network. The sensors in a grid send the data packet directly to the aggregator of that grid. Hence, the sensors within a grid do not communicate with each other. In-network aggregation is similar to grid-based data aggregation with two major differences; each sensor within a grid communicates with its neighboring node. Any node within a grid can assume the role of aggregator node in terms of rounds until the last node dies. This is similar to cluster-based data aggregation in which the cluster heads are fixed [4][5][6]

**Table1:** Protocol based on different data aggregation approach



| Protocol | Organization type | Objectives | Characteristics |
|---|---|---|---|
| LEACH | cluster | Network lifetime: number of nodes that are alive, latency | Randomized cluster head rotation, non-uniform energy drainage across different sensors. |
| HEED | cluster | Lifetime: number of rounds until the first node death | Assumption: Multiple power levels in sensors. Cluster heads are well distributed. Achieves better performance than LEACH |
| PEGASIS | chain | Lifetime: average energy expended by a node | Global knowledge of the network is required. Considerable energy savings compared to LEACH. |
| Hierarchical chain based protocols | chain | Energy× delay | Binary chain based scheme is eight times better than LEACH and the three level scheme is 5 times better than PEGASIS. |
| EADAT | tree | Lifetime: number of alive sensors at the end of simulation time | Sink initiated broadcasting approach. It is not clear how to choose the threshold power ($P_{th}$) for broadcasting help messages. No comparisons made with other existing aggregation algorithms. |
| PEDAP-PA | tree | Lifetime: time until the death of last node | Minimum spanning tree based approach. Achieves two times performance improvement compared to LEACH, PEGASIS. |

## 6. Secure Data aggregation

In sensor network, aggregation is used to aggregate the data from several sensors, which substantially reduces the communication overhead. The security issues such as data integrity, confidentiality and freshness in data aggregation become crucial when the WSN are deployed in a remote or hostile environment where sensors are prone to node failures and compromises. In sensor network the method of data aggregation is not only important providing security is also very important. Secure data aggregation is the method of providing security to the aggregated data, several secure aggregation algorithms have been proposed assuming that the base station is the only aggregator node in the network, Where the each node directly send an authentication message to the base station, which is a very expensive solution, another general approach for providing security is the encryption technology. Secure data aggregation is classified into two types they are: Hop-by Hop and End-to-End. In Hop-by Hop secure data aggregation, aggregator node must decrypt all data what they receive and encrypt the aggregation result before send to next hop by using aggregation function. In End-to-End secure data aggregation scheme performs data aggregation through homomorphic encryption technology. One intermediate (cluster head) node receives the cipher texts from leaf (cluster) nodes and then aggregates them with its own encrypted sensor data; the result will finally be sent to a next node SEEDA is the secure end-to-end data aggregation protocol, this protocol provides security for aggregated data. Here the data encryption performed at the source node and it is decrypted by the sink node and the cipher text is added at the aggregator node. This protocol is based on the additive homomorphic encryption; this encryption allows addition of cipher text which when decrypted results in addition of the respective plan. In the end-to-end secure protocol data is data is encrypted by adding respective secret key for all responding nodes. upon receiving the cipher text sink node is able to decrypt the data using the information of non responding node that is send along with the cipher text This method of forwarding the node information along with the cipher text will increases the number of bits transmitted. This has to be overcome by SEEDA protocol. In SEEDA protocol instead of sending the non responding node data only compute a cipher text for non responding nodes considering this data is 0. Even though some nodes are not responded the cipher text received by sink node is added with keys of all the nodes. Aggregated data is obtained by the sink node By Subtracting respective keys of all the nodes. The scheme ensures end-to-end data privacy with less number of bits transmitted compared to end-to-end aggregation scheme. RDAS-Reputation-based-Resilient-Data Aggregation in Sensor Network. This protocol is based on the hop-by hop data aggregation. It is a robust data aggregation protocol that uses a reputation based approach in order to identify the and locate malicious nodes in a sensor network. RDAS is based on the clustered sensor arrangements, where the cluster head is going to analyzes data from cluster nodes in order to determine the location of an event and also it easily fin\d the redundancy of multiple nodes sensing an event to determine what data have been reported by each node. The main advantages of RDAS. RDAS develops and effectively uses a reputation system to add resiliency to the common function of event localization. It provides concrete answers to the questions of how to effectively generate, propagate and use reputation ratings such that RDAS can handle both colluding and non-colluding faulty nodes as well as lying nodes trying to compromise the reputation system. RDAS can provide security in a network where there does not exist any trusted node and under compute, memory and bandwidth constraints.[7][8][9][10].

## 7. Conclusion

In this paper we present wireless sensor network is consist a hundreds or thousands of sensor node. And these nodes are constraint of resource such as Energy, Storage, and Communication. Especially one powered by battery cannot be replacing so easily hence the various approaches or protocol has been proposed for increasing the lifetime of the wireless sensor network. In this paper we present the data aggregation method is one of the important techniques for enhancing the life time of the network. And we also present the various approaches, advantages and disadvantages for data aggregation. And also we discussed various approach for secure data aggregation in wireless sensor networks

## References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci,"A Survey on Sensor Networks", IIEEE Commun. Mag, 40(8), 102-114, 2002.

[2] J. Yick, B. Mukherjee, D. Ghosal, Wireless sNetworks Survey, Comput. Networks, 52(12), 2292-2330, 2008.

[3] K. Akkaya, M. Demirbas, R.S. Aygun, The Impact of Data Aggregation on the Performance of Wireless Sensor Networks, Wiley Wireless Commun. Mobile Comput. (WCMC) J. 8 (2008) 171–193.

[4] S. Chatterjea and P.Havinga, "A Dynamic data aggregation scheme for wireless sensor networks," *Proc. Program for Research on Integrated Systems and Circuits*, Veldhoven, The Netherlands, Nov. 2003

[5] Lindsey, S.; Raghavendra, C.; Sivalingam, K.M. Data gathering algorithms in sensor networks using energy metrics. *IEEE Trans. Parallel Distrib. Syst.* 2002, *13*, 924–935.

[6] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A Scalable and robust communication paradigm for sensor networks," *Proceedings of the Sixth Annual International Conference on Mobile Computing and Networking (MobiCOM '00),* August 2000.

[7] Suat Ozdemir, Yang Xiao, Secure data aggregation in Wireless Sensor Networks: A Comprehensive Overview The International Journal of Computer and Telecommunications Networking vol.53 (12), 2022-2037, 2009

[8] Hani Alzaid, Ernest Foo, Juan Gonzalez Nieto Secure data aggregation in wireless sensor network: a survey, sixth Australasian conference onInformation security, Wollongong, NSW, Australia, 2008.

[9] Carlos R. Perez-Toro, Rajesh K. Panta, Reputation-based Resilient Data Aggregation in Sensor Network, IEEE Communications Society,2010.

[10] A.S.Poornima, B.B.Amberker, Secure End-to-End Data Aggregation in Wireless Sensor Networks, IEEE Communications Society, 2010.

## Author Profile

**Tanushree K N** received the B.E degree in Computer Science and Engineering from Visvesvaraya Technological University in 2011. She is currently pursuing her M.Tech degree in Computer Science and Engineering under the same University.

**Prassanna Kumar M** received the B.E and M.Tech degrees in Computer Science and Engineering from Visvesvaraya Technological University. With 12 years of experience in lecturing field, presently he is working as Assistant Professor in Computer Science Dept East West Institute of Technology Bangalore.