

# Data Aggregation Techniques in Wireless Sensor Networks: Literature Survey

Ameya S. Bhatlavande  
PG Student

N.B. Navale Sinhgad College of Engineering,  
Solapur

Amol A. Phatak  
Assistant Professor

N.B. Navale Sinhgad College of Engineering,  
Solapur

## ABSTRACT

The wireless detector network is wide employed in totally different areas of application like military applications etc. Wireless detector network is nothing however the cluster of detector nodes those area unit human action to every alternative mistreatment radio frequencies through the bottom station. Every detector node in network has the battery that is important and crucial to outline the lifetime of detector node. Hence it's needed to possess economical energy utilization with aim of extending the lifetime of detector networks. One approach that greatly reduces this consumption of power is reduction of within detector networks information traffic that successively reduces the information sent to the bottom station. This method of collection information and causation it to the bottom station is termed as information aggregation. There area unit has several information aggregation techniques conferred by varied researchers with aim of rising the performance of wireless detector network in terms of energy consumption, network out-turn, packet delivery quantitative relation etc. primarily the information aggregation algorithms collection and aggregating data in economical utilization of power with aim of extend the period. Throughout this review paper, we tend to area unit planning to present the various existing strategies for in-network information aggregation in WSNs.

## Keywords

Wireless sensor networks, data aggregation, data collection, routing protocols, energy utilization, SPT, LEACH, Latency.

## 1. INTRODUCTION

A detector network is outlined as being composed of an out sized range of nodes with sensing, process and communication facilities that square measure densely deployed either within the development or terribly near it. Every of those nodes collect knowledge and its purpose is to route this data back to a sink. The network should possess self-organizing capabilities since the positions of individual nodes don't seem to be planned. Cooperation among nodes is that the dominant feature of this sort of network, wherever teams of nodes join forces to spread the data gathered in their neighborhood to the user. Recent advances in micro-electric mechanical systems (MEMS) technology, wireless communications, and digital natural philosophy have created potential to develop inexpensive, low-power, multi functional detector nodes that square measure little in size and communicate freely in brief distances [1]. These little detector nodes, that incorporates sensing, processing, and communication parts, leverage the thought of detector networks supported cooperative effort of an out sized range of nodes. Detector networks represent a major improvement over ancient sensors, that square measure deployed within the following 2 ways in which [1]:

- Sensors may be positioned removed from the particular development, i.e., one thing legendary by sense perception. During this approach, massive sensors that use some advanced techniques to tell apart the targets from environmental noise ar needed.

- Many sensors that perform solely sensing may be deployed. The positions of the sensors and communications topology ar rigorously built.

They transmit statistic of the perceived development to the central nodes wherever computations ar performed and knowledge ar amalgamated. The position of device nodes needn't be built or pre-determined. this enables random readying in inaccessible terrains or disaster relief operations. On the opposite hand, this conjointly implies that device network protocols and algorithms should possess self-organizing capabilities.

The main goal information|of knowledge|of information| aggregation algorithms is to assemble Associate in Nursing d combination data in an energy economical manner so network time period is increased. WSN supply Associate in Nursing more and more enticing methodology of information gathering in distributed system architectures and dynamic access via wireless property [2]. information aggregation {attempts|makes Associate in Nursing attempt|tries} to gather the foremost crucial information from the sensors and build it out there to the sink in an energy economical manner with minimum information latency. information latency is vital in several applications like setting observation wherever the freshness of information is additionally a vital issue. it's crucial to develop energy economical information aggregation algorithms so network time period is increased. There square measure many factors that verify the energy potency of a sensing element network like spec, the info aggregation mechanism and also the underlying routing protocol. we have a tendency to currently gift a proper definition of energy potency.

**Energy Efficiency:** The practicality of the sensing element network ought to be extended as long as attainable. In a perfect information aggregation theme, every sensing element ought to have exhausted identical quantity of energy in every information gathering spherical. a knowledge aggregation theme is energy economical if it maximizes the practicality of the network. If we have a tendency to assume that each one sensors square measure equally vital, we should always minimize the energy consumption of every sensing element. this concept is captured by the network time period that quantifies the energy potency of the network.

Network time period, information accuracy, and latency square measure a number of the vital performance measures of information aggregation algorithms. The definitions of those measures square measure extremely hooked in to the

required application. we have a tendency to currently gift a proper definition of those measures.

**Network time period:** Network lifetime is outlined because the numbers of information aggregation rounds until awake to sensors die wherever is given by the system designer. for example, in applications wherever the time that each one nodes operate along is significant, time period is outlined because the range of rounds till the primary sensing element is drained of its energy. the most plan is to perform information aggregation such there's uniform energy drain within the network. additionally, energy potency and network time period square measure substitutable in this up energy potency enhances the time period of the network [3].

**Data accuracy:** The definition of information accuracy depends on the particular application that the detector network is intended. as an example, in a very target localization drawback, the estimate of target location at the sink determines the information accuracy.

**Latency:** Latency is outlined because the delay concerned in knowledge transmission, routing and knowledge aggregation. It may be measured because the time delay between {the knowledge|the info|the information} packets received at the sink and therefore the data generated at the supply nodes [4].

In this paper we tend to square measure discussing the various aspects of in-network information aggregation, totally different strategies bestowed already, and its options. the most aim of this paper is to spot however precisely the method of information aggregation works for WSN, and what square measure the vital performance metrics those square measure settled with use of economical information aggregation methodology. In next section, section II bestowed the review of information aggregation and its totally different techniques. In section III we are going to discuss regarding the In-Network Aggregation Technique with its benefits.

## 2. REVIEW OF DATA AGGREGATION TECHNIQUES

Data aggregation could be a method of aggregating the sensing element knowledge victimization aggregation approaches. the overall knowledge aggregation rule works as shown within the below figure. The rule uses the sensing element knowledge from the sensing element node and so aggregates the info by victimization some aggregation algorithms like centralized approach, LEACH(low energy adjustable bunch hierarchy),TAG(Tiny Aggregation) etc. This aggregative knowledge is transfer to the sink node by choosing the economical path. Figure one is showing the fundamental design of information aggregation technique. There area unit many sorts of aggregation techniques area unit gift a number of them area unit listed below.

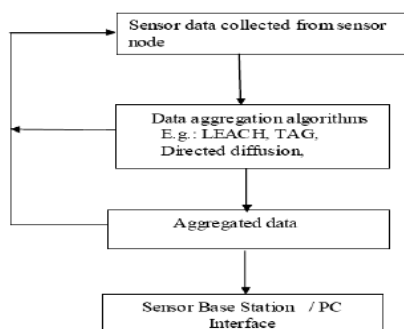


Fig 1: Architecture of Data Aggregation Technique

### 2.1 Centralized Approach

This is AS address central approach wherever every node sends information to a central node via the shortest potential route employing a multihop wireless protocol. The device nodes merely send the information packets to a pacesetter, that is that the powerful node. The leader aggregates the information which may be queried. every intermediate node must send the information packets self-addressed to leader from the kid nodes. thus an out sized range of messages ought to be transmitted for a question within the best case up to the total of external path lengths for every node.

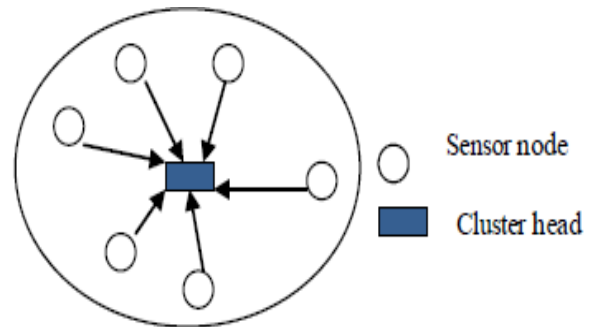


Fig 2: Centralized Data Aggregation Technique Architecture

### 2.2 In-Network Aggregation [4]

In-network aggregation is that the international method of gathering and routing info through a multihop network, process information at intermediate nodes with the target of reducing resource consumption (in explicit energy), thereby increasing network period. There square measure 2 approaches for in-network aggregation: with size reduction and while not size reduction. In-network aggregation with size reduction refers to the method of mixing & pressure the information packets received by a node from its neighbors so as to scale back the packet length to be transmitted or forwarded towards sink. In-network aggregation while not size reduction refers to the method merging information packets received from completely different neighbors in to one information packet however while not process the worth of information. Researchers at the University of Pittsburgh propose Associate in Nursing improved synchronization theme known as TiNA (Temporal Coherency-Aware in-Network Aggregation) in [1]. The goals of TiNA square measure to scale back the quantity of data that has got to propagate through the network, and to extend the accuracy of according results even with some sensors have lost communication with the network.

### 2.3 Tree- based Approach [5]

Within the tree-based approach perform aggregation by constructing Associate in Nursing aggregation tree, that might be a minimum spanning tree, stock-still at sink and supply nodes square measure thought-about as leaves. every node incorporates a parent node to forward its information. Flow of information starts from leaves nodes up to the sink and in this the aggregation done by parent nodes. In most cases, tree-based protocols build a conventional shortest path routing tree. for example, the Shortest Path Tree (SPT) rule uses a really easy strategy to create a routing tree in an exceedingly distributed fashion [10].

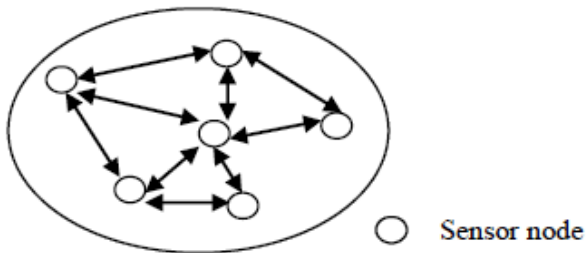


Fig 3: Decentralized Technique Aggregation

## 2.4 Cluster-based Approach [3]:

In cluster-based approach, whole network is split in to many clusters. every cluster features a cluster-head that is chosen among cluster members. Cluster heads do the role of person that mixture information received from cluster members domestically so transmit the result to sink. The Low-Energy adaptive cluster prelate (LEACH) algorithmic rule, clustered structures area unit exploited to perform information [11].

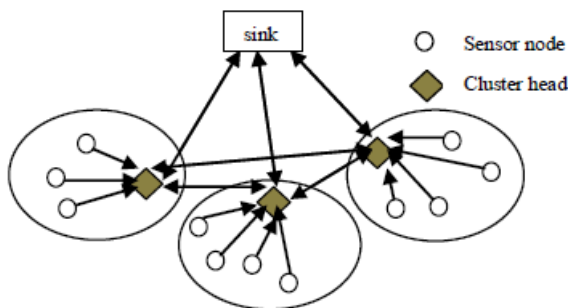


Fig 4: Cluster Based Technique Architecture

## 3. OVERVIEW OF IN-NETWORK AGGREGATION

A query requesting combination information is injected into the device network at a number node, conjointly referred to as a sink. The question is forwarded by the host to the opposite nodes within the network. the only and least optimum question arrange would need every node to report its own readings back to the host node for process. when receiving all information packets from the supply nodes, the host node would combination all of the info into a final worth and report the worth back to the user. This approach, referred to as direct delivery [6], encompasses a range of disadvantages. One downside is that an out sized range of packets should be sent to the host node. Since every node sends its own information to the host, there should be a minimum of one packet of information sent per node. in addition, since some nodes might not be able to communicate directly with the host, their information packets should be forwarded by different nodes till they reach the host. A connected downside is that the dimensions of every packet is comparatively little since it solely contains readings from one device. The raised range of little packet transmissions necessary to propagate required information to the host severely limits the lifetime of the network.

In order to conserve each energy and information measure, it's helpful to maneuver the mixing and filtering of device information into the network itself. In-network aggregation could be a mechanism for reducing the general quantity of power and information measure needed to method the user's

question by permitting device readings to be mass by intermediate nodes.

### 3.1 Advantages

**1. A discount within the range of packets that has to be sent through the network.** As packets propagate from supply nodes they'll be combined along into fewer packets containing mixture values. as a result of the user isn't inquisitive about individual values there's no loss within the quality of the result came back.

**2. A discount within the chance of packet collisions.** Since fewer packets area unit sent through the network, it's less doubtless that a device node can have to be compelled to resend a message lost in a very network collision [7].

**3. A discount within the quantity of redundancy received at the host node.** within the case wherever individual readings area unit rumored to the host node, it's doubtless that multiple intermediate nodes can hear a message from a supply node and every forward the message toward the host. during this case the host would receive multiple copies of identical message and would want to separate out the redundant knowledge [8].

**4. a rise in accuracy of results.** device nodes that become briefly disconnected from the network aren't ready to answer queries. during this case, the parent node can notice that the kid node didn't report a worth and should be ready to estimate the worth supported previous readings. this is often particularly useful in non-volatile environments wherever readings area unit unlikely to alter considerably over short periods of your time (e.g. the temperature in one area of AN workplace building).

### 3.2 Properties of Aggregates

In order to talk of mixture functions normally terms, it's useful to spot properties common among teams of aggregates. Researchers at UC Berkeley denote many dimensions accustomed classify aggregates [9].

**1. Duplicate sensitivity.** This property specifies whether or not AN mixture operate can come back identical result once the data set contains duplicate values. samples of duplicate sensitive aggregates area unit MEDIAN, AVERAGE, and COUNT. samples of duplicate insensitive aggregates embrace MIN, MAX, and COUNT DISTINCT.

**2. Exemplary/Summary.** Exemplary aggregates perpetually come back a representative price gift within the data set whereas outline aggregates perform some calculation over the whole data set and come back the calculated price. outline values (such as AVERAGE and COUNT) area unit additional simply calculable even in a very loss network wherever all knowledge packets aren't received. Exemplary aggregates, on the opposite hand, could also be extremely inaccurate if even many messages area unit lost. Such aggregates embrace MIN, MAX, and MEDIAN.

**3. Monotonic aggregates.** Aggregates that permit early testing of predicates within the network area unit monotonic. for instance, assume the user requests the liquid ecstasy temperature reading within the network. As supply nodes report their values toward the host node, alternative nodes could listen and solely report their own values if they're larger than the present liquid ecstasy. This provides savings within the overall range of messages sent through the network while not poignant the result.

**4. Partial state necessities.** the quantity of partial state info needed differs among mixture functions. Aggregates like total and COUNT need partial state records that area unit identical size because the final mixture. the common operate needs a partial state record containing 2 values (both the total and COUNT). alternative aggregates like MEDIAN and bar graph need that the whole data set be came back to the host node unless some style of compression or estimation is employed.

#### **4. CONCLUSION AND FUTURE WORK**

In this review paper we have given the elaborate insights for method of knowledge aggregation and its totally different techniques of doing. The survey is given on {different|totally totally different|completely different} aspects of wireless sensing element networks with different knowledge aggregation architectures. This all knowledge aggregation architectures ar specializing in improves the performance of power utilization in WSN. Energy is extremely essential resource of wireless sensing element network that defines the period of time of WSN. in conjunction with the energy, there are another performance metrics that additionally thought of whereas planning totally different knowledge aggregation techniques like output, latency, knowledge accuracy, delay etc. Recently several researches are going over the info aggregation techniques with aim of up the routing performance of WSN. For the longer term work, we are able to recommend to figure over recently given technique and appraise it beneath totally different network conditions.

#### **5. REFERENCES**

- [1] H. Cam, S. Ozdemir, P. Nair, and D.Muthuavina shiappan, “ESPDA: Energy-Efficient and Secure Pattern-based Data Aggregation for Wireless Sensor Networks”, in *Proceedings of IEEE Sensor- The Second IEEE Conference on Sensors, Toronto, Canada, Oct. 22-24, 2003*, pp. 732-736.
- [2] Chalermek Intanagonwivat, Ramesh Govindan, and Deborah Estrin, “Directed diffusion: a scalable and robust communication paradigm for sensor networks”, (MobiCom 2000) pp 56-67.
- [3] K. Dasgupta, K. Kalpakis, and P. Namjoshi, “An Efficient Clustering-based Heuristic for Data Gathering and Aggregation in Sensor Networks”, IEEE 2003.
- [4] J. E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, “In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey”, IEEE Wireless communication 2007.
- [5] M. Lee and V.W.S. Wong, “An Energy-aware Spanning Tree Algorithm for Data Aggregation in Wireless Sensor Networks,” *IEEE PacRim 2005*, Victoria, BC, Canada, Aug. 2005.
- [6] Yong Yao and Johannes Gehrke, “Query processing in sensor networks,” in *Proceedings of the First Biennial Conference on Innovative Data Systems Research (CIDR)*, 2003.
- [7] Dragan Petrovic, Rahul C. Shah, Kannan Ramchandran and Jan Rabaey, “Data Funneling: Routing with Aggregation and Compression for Wireless Sensor Networks,” in *Proceedings of the IEEE Sensor Network Protocols and Applications (SNPA)*, May 2003.
- [8] B. Krishnamachari, D. Estrin, and S. Wicker, “The Impact of Data Aggregation in Wireless Sensor Networks,” *International Workshop on Distributed Event-Based Systems, (DEBS '02)*, Vienna, Austria, July 2002.
- [9] S. R. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, “TAG: A tiny aggregation service for ad-hoc sensor networks,” in *Proceedings of the Symposium on Operating Systems Design and Implementation, OSDI*, December 2002.
- [10] B. Krishnamachari, D. Estrin, and S.B. Wicker, “The Impact of Data Aggregation in Wireless Sensor Networks,” Proc. 22nd Int’l Conf. Distributed Computing Systems (ICDCSW ’02), pp. 575-578, 2002.
- [11] A.P. Chandrakasan, A.C. Smith, and W.B. Heinzelman, “An Application-Specific Protocol Architecture for Wireless Microsensor Networks,” IEEE Trans. Wireless Comm., vol. 1, no. 4, pp. 660-670, Oct. 2002.