# A Survey on Data Aggregation Mechanism in Wireless Sensor Networks

[1]Sayali Tembhurne, [2]Rajesh babu, [3]Roshani Talmale

[1]*M.Tech Student, Department of Computer Science & Engineering, Tulsiramji Gaikwad Patil College of Engineering and Technology, Nagpur, Maharashtra, India.*
[2,3]*Assistant Professor, Department of Computer Science & Engineering, Tulsiramji Gaikwad Patil College of Engineering and Technology, Nagpur, Maharashtra, India.*

**Abstract**— *Data aggregation is exceptionally critical methods in wireless sensor network. Since with the assistance of data aggregation we lessen the vitality utilization by wiping out excess. At the point when wireless sensor network sent in remote regions or threatening condition. In the wireless sensor network have the most difficult assignment is an existence time so with help of data aggregation we can improve the lifetime of the network .In this paper we talk about the data aggregation approaches based on the directing conventions, the calculation in the wireless sensor network. And furthermore talk about the favourable circumstances and detriments or different execution measures of the data aggregation in the network.*
**Keywords**— *Wireless sensor network, data aggregation, architecture, Network Lifetime, Routing, Tree, Cluster, Base Station*

## I.     Introduction

The wireless sensor network is specially appointed network. It comprises little light weighted wireless hubs called sensor hubs, sent in physical or natural condition. Furthermore, it gauged physical parameters, for example, sound, weight, temperature, and moistness. These sensor hubs conveyed in substantial or thousand numbers and work together to shape a specially appointed network fit for answering to data gathering sink (base station). Wireless sensor network have different applications like territory observing, building checking, wellbeing observing, military survival spear and target following. However wireless sensor network is an asset requirement in the event that we discuss vitality, calculation, memory and constrained correspondence abilities. All sensor hubs in the wireless sensor network are interface with each other or by moderate sensor hubs.

A sensor hubs that produces data, based on its detecting instruments perception and transmit detected data bundle to the base station (sink).
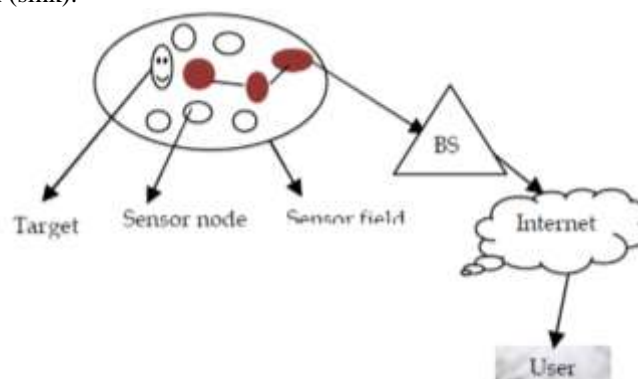


**Figure 1** Architecture of the Sensor network

This procedure fundamentally coordinate transmission since the base station may found extremely far from sensor hubs needs. More vitality to transmit data over long separations so a superior method is to have less hubs send data to the base station. These hubs called aggregator hubs and procedures called data aggregation in wireless sensor network.

## II.     Related Work

*A.*   Clustering In WSN

Sensor hub are thickly conveyed in wireless sensor network that implies physical condition would deliver fundamentally the same as data in near to sensor hub and transmitting such sort of data is pretty much

repetitive. So every one of these realities support utilizing some sort of collection of sensor hubs to such an extent that gathering of sensor hub can be consolidated or pack data together and transmit just smaller data. This can diminish limited activity in singular gathering and furthermore lessen worldwide data. This gathering procedure of sensor hubs in a thickly conveyed huge scale sensor hub is known as clustering. The method for brushing data and pack data having a place with a solitary cluster called data combination (aggregation).
Issues of clustering in wireless sensor network:-

1. What number of sensor hubs ought to be taken in a solitary cluster? Determination method of cluster head in an individual cluster
2. Heterogeneity in a network, it implies client can put some power full hubs, in term of vitality in the network which can carry on like cluster head and basic hub in a cluster fill in as a cluster part as it were. Numerous conventions and calculation have been proposed which manage every individual issue.

*B.* Data Aggregation

In run of the mill wireless sensor networks, sensor hubs are generally asset compelled and battery-constrained. Keeping in mind the end goal to spare assets and vitality, data must be collected to abstain from overpowering measures of activity in the network. There has been broad work on data aggregation conspires in sensor networks, the point of data aggregation is that takes out excess data transmission and improves the lifetime of vitality in wireless sensor network. Data aggregation is the procedure of one or a few sensors at that point gathers the recognition result from other sensor. The gathered data must be prepared by sensor to diminish transmission load before they are transmitted to the base station or sink. The wireless sensor network has comprised three sorts of hubs. Straightforward standard sensor hubs, aggregator hub and queried. Normal sensor hubs sense data bundle from the earth and send to the aggregator hubs essentially these aggregator hubs gather data from various sensor hubs of the network, totals the data parcel utilizing a some aggregation work like total, normal, tally, max min and afterward sends totals result to upper aggregator hub or the queried hub who produce the question.
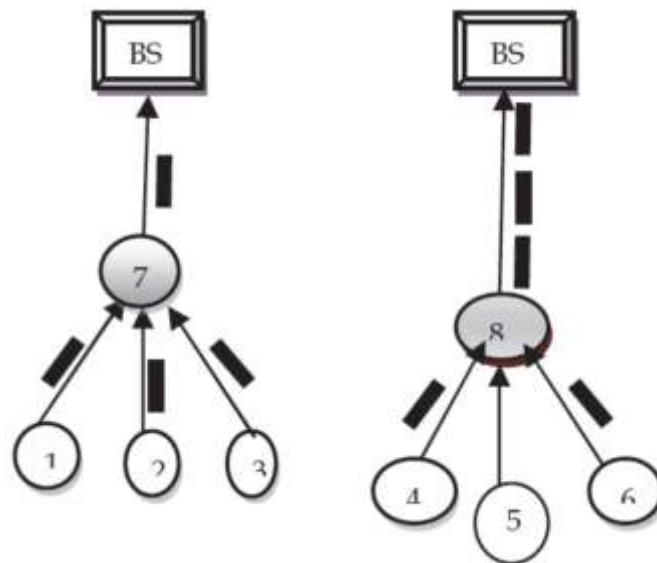


**Figure 2** Data aggregation model and Non data aggregation show

It can be the base station or in some cases an outer client who has authorization to associate with the network. Data transmission between sensor hubs, aggregators and the querier devours part of vitality in wireless sensor network. Figure 2 contain two models one is data aggregation model and second is non-data aggregation demonstrate in which sensor hubs 1, 2, 3, 4, 5, 6 are general hubs that gathering data bundle and announcing them back to the upper hubs where sensor hubs 7,8 are aggregators that perform detecting and totaling in the meantime. In this aggregation demonstrate 4 data parcel went inside the network and just a single data bundle is transmitted to the base station (sink). What's more, other non-data aggregation demonstrate likewise 4 data parcel went inside the network and all data bundles are sent to the base station(sink), implies we can state that with the assistance of data aggregation process we diminish the quantity of data bundle transmission. And furthermore spare vitality of the sensor hub in the wireless sensor network. With the assistance of data aggregation we upgrade the lifetime of wireless sensor network. Sink have a data parcel with vitality productive way with least data idleness. So data idleness is critical in numerous uses of wireless sensor network, for example, condition observing, wellbeing, checking, where the freshness of data is likewise an imperative factor.

It is basic to create vitality proficient data-aggregation calculations so network lifetime is upgraded. There are a few components which decide the vitality effectiveness of a sensor network, for example, network engineering, the data aggregation instrument, and the fundamental steering convention. Wireless sensor network has conveyed handling of sensor hub data. Data aggregation is the system. It depicts the preparing technique that is connected on the data got by a sensor hub and in addition data is to be steered in the whole network. In which lessen vitality utilization of the sensor hubs and furthermore decrease the quantity of transmissions or length of the data parcel.

Elena Fosolo et al in [5] portray "In network aggregation is the selective procedure of gathering and directing data through a multi hop network. Handling of data parcel with the assistance of halfway sensor hubs. The goal of this approach is expanding the life time of the network and furthermore diminishes asset utilization. There are two sort of approach for in network aggregation. With estimate diminishment and without measure lessening .In network aggregation with estimate decrease. It is the procedure in which join and compacting the data got by a sensor hub from its neighbors keeping in mind the end goal to lessen the length of data bundle to be sent towards the base station. Case, in some situation a hub gets two data bundles which have a corresponded data. In this condition it is pointless to send the two data parcels. At that point we apply a capacity like MAX, AVG, and MIN and again send single data parcel to base station.

With help of this approach we lessen the quantity of bit transmitted in the network and furthermore spare a considerable measure of vitality. In network aggregation without measure lessening is characterized during the time spent data bundles got by various neighbors in to a solitary data parcel however without preparing the estimation of data. This procedure additionally decreases vitality utilization or increment life time of the network.

### a. Advantage and Disadvantage of Data Aggregation in WSN

Preferred standpoint: With the assistance of data aggregation process we can improve the power and precision of data which is acquired by whole network, certain excess exists in the data gathered from sensor hubs subsequently data combination handling is expected to lessen the repetitive data. Another preferred standpoint is those diminishes the activity load and monitor vitality of the sensors.

Disservice: The cluster head implies data aggregator hubs send meld these data to the base station .this cluster head or aggregator hub might be assaulted by noxious assailant. In the event that a cluster head is traded off, at that point the base station (sink) can't be guarantee the accuracy of the total data that has been send to it. Another downside is existing systems are a few duplicates of the total outcome might be sent to the base station (sink) by uncompromised hubs. It increment the power expended at these hubs.

### b. Performance measure of Data Aggregation

There are critical execution measures of data combination calculation. These exhibitions are exceptionally subject to the coveted application.

- Vitality Efficiency: By the data-aggregation plot, we can build the usefulness of the wireless sensor network. In which each sensor hubs ought to have spent a similar measure of vitality in each datum assembling round. A data aggregation conspire is vitality proficient in the event that it augments the usefulness of the network. Network lifetime, data exactness, and inertness are a portion of the critical execution measures of data-aggregation calculations. The meanings of these measures are exceptionally subject to the coveted application.
- Network lifetime: The network lifetime is characterizing the quantity of data combination rounds. Till the predefined level of the aggregate hubs passes on and the rate rely upon the application .If we discuss some application, simultaneously working of the all the sensor hubs is vital thus the lifetime of the network is number of round until the main hubs which enhances the vitality proficiency of hubs and improve the lifetime of entire network.
- Idleness: Latency is assess data of time defer encounters by system, implies data send by sensor hubs and got by base station (sink).basically postpone associated with data transmission, steering and data aggregation.
- Correspondence overhead: It assesses the correspondence unpredictability of the network combination calculation.
- Data precision: It verifies the proportion of aggregate number of perusing got at the base station (sink) to the aggregate number of created. There are diverse writes data aggregation conventions like network engineering based data aggregation conventions, network-stream based data aggregation conventions and nature of administration (QOS) - mindful data-aggregation conventions intended to ensure QOS measurements. Here network design based conventions are depicted in detail.

**c.  Impact of Data Aggregation in WSN**

In this section we talk about the two principle factors that influence the execution of data aggregation strategies in wireless sensor network, Such as vitality sparing and delay. Data aggregation is the procedure, in which collecting the data bundle originating from the distinctive sources; the quantity of transmission is decreased. With the assistance of this procedure we can spare the vitality in the network. Deferral is the inactivity associated with aggregation data from nearer sources may need to kept down at transitional hubs keeping in mind the end goal to join them with data from source that are more remote away. Fundamentally aggregation technique based on the position of the sources in the network, number of sources and the network topology. On the off chance that it analyse the variables, we think about the two models of the source arrangement. The occasion sweep (ER) model and arbitrary source demonstrate [14]. The demonstrating says us that where the source are clustered close to each other or found haphazardly, noteworthy vitality picks up are conceivable with data aggregation. These increases are most noteworthy when the quantity of sources is vast, and when the sources are found generally near each other and a long way from base station. The displaying through, additionally appears to the recommend that aggregation inertness could be non-immaterial.

## III.  Literature Survey

Data aggregation process is performed by particular steering convention. Our point is totalling data to limit the vitality utilization. So sensor hubs should course bundles based on the data parcel content and pick the following hop with a specific end goal to advance in network aggregation. Essentially steering convention is isolated by the network structure that is the reason directing conventions is based on the thought about methodologies.

**a.  Tree-Based Approach**

The tree based approach is characterizing aggregation from building an aggregation tree. The type of tree is least traversing tree, sink hub consider as a root and source hub consider as a clears out. Data streaming of data begin from surrenders hub over to root implies sink (base station).Disadvantage of this approach, as we probably am aware like wireless sensor network are not free from disappointment .if there should arise an occurrence of data bundle misfortune at any level of tree, the data will be lost for single level as well as for entire related sub tree too. This approach is reasonable for outlining ideal aggregation procedures'. Enrage et al. in [6] data driven convention know as Tiny aggregation (TAG) approach. The working of TAG is relying upon two stages: circulated stage and accumulation stage.
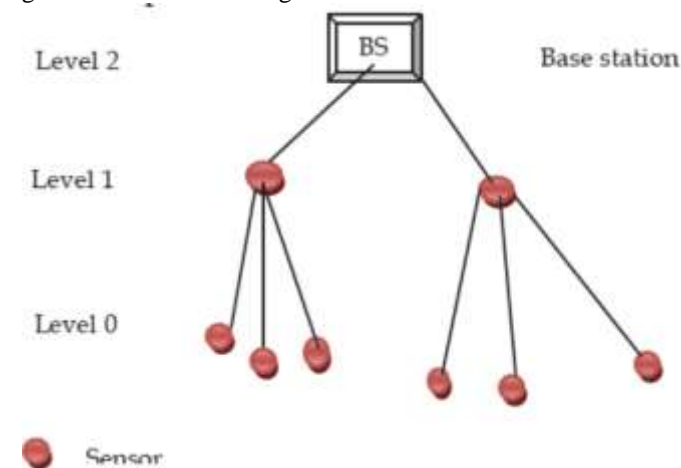


**Figure 3** Tree based data aggregation in wireless sensor network

**b.  Cluster-Based Approach**

In vitality obliged sensor networks of expansive size, it is wasteful for sensors to transmit the data straightforwardly to the sink in such situations, Cluster based approach is various leveled approach. In cluster-based approach, entire network is isolated in to a few clusters. Each cluster has a cluster-head which is chosen among cluster individuals. Cluster-heads do the part of aggregator which total data got from cluster individuals locally and after that transmit the outcome to base station (sink). As of late, a few cluster-based network association and data-aggregation conventions have been proposed for the wireless sensor network. Figure 4 demonstrates a cluster-based sensor network association. The cluster heads can speak with the sink specifically by means of long range transmissions or multi hopping through other cluster heads.
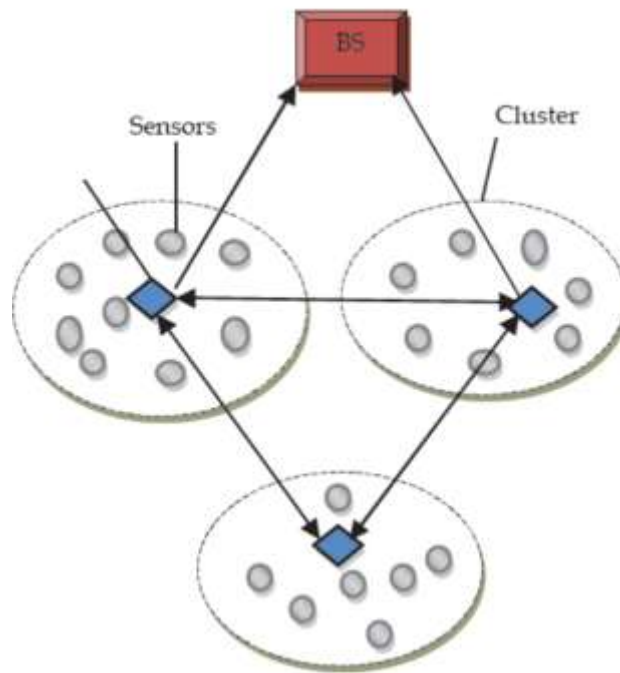
**Figure 4** Cluster based sensor network.

The bolts show wireless correspondence joins K. Dasgupta et al. in [7] proposed a most extreme lifetime data aggregation (MLDA) calculation which discovers data gathering plan gave area of sensors hub and base-station, data bundle size, and vitality of every sensor hub. A data gathering plan determines how data bundle are gathered from sensors hub and transmitted to base station for each round. A calendar can be thought of as a gathering of aggregation trees. In [4], they proposed heuristic-avaricious clustering-based MLDA based on MLDA calculation. In this they parceled the network in to cluster and alluded each cluster as super-sensor. They at that point register greatest lifetime plan for the super-sensors and after that utilization this timetable to build aggregation trees for the sensors. W. Choi et al. in [1] introduce a two-stage clustering (TPC) conspire. Stage I of this plan makes clusters with a cluster-head and every hub inside that cluster frame a direct associates with cluster-head. Stage I the cluster-head revolution is restricted and is done based on the rest of the vitality level of the sensor hubs which limit time change of sensors and this prompt vitality sparing from pointless cluster-head turn. In stage II, every hub inside the cluster looks for a neighbour nearer than cluster-head which is called data transfer point and setup up a data hand-off connection. Presently the sensor hubs inside a cluster either utilize guide connection or data hand-off connection to send their data to cluster head which is a vitality proficient plan. The data transfer point totals data at sending time to another data hand-off point or cluster-head. In the event of high network thickness, TPC stage II will setup superfluous data transfer connect between neighbours as nearly conveyed sensor will detect same data and this prompt a misuse of vitality.

### c. Multi-Way Approach

The disadvantage of tree based approach is the constrained power of the system. To conquer this disadvantage, another approach was proposed by numerous analysts .in which sending somewhat collected data to single parent hub in aggregation tree, a hub could send data over different ways. In which every single hub can send data bundles to its conceivably various neighbours. Hance data parcel spill out of source hub to the sink hub along numerous way, part of transitional hub between source hub to sink hub so aggregation done in each middle hub. Utilizing this approach we will make the system hearty yet some additional overhead. The case of this approach like ring topology, where network is separated in to concentric hover with characterizing level levels as indicated by hop remove from sink.[3]propose another procedure have the two issues : vitality proficiency and power. In which single way to interface every hub to the base station it is vitality sparing however high danger of connection disappointment. Yet, on the other head multipath approach would require more hubs to take an interest with ensuing misuse of vitality. Creators show a cunning utilization of multi-way just when there is loss of parcel which is actualized by shrewd reserving of data at sensor hubs. Creators additionally contend that in numerous down to earth circumstance data might be assembled just from a specific area, so they utilize an alternate approach that depends on a spreading over tree and gives elective ways just when a failing is identified. Calculation embraces a tree-based approach for sending bundles through the network. In the perfect circumstance when no disappointments happen, this is positively the best decision, as the

base number of hubs is occupied with the transmission stage. Within the sight of connection or hub disappointments, the calculation will find elective ways, so as guarantee the conveyance of however many bundles as would be prudent inside the time limitations. The issue with this approach is that it might cause the emerging of problem areas and hubs along favoured ways will devour their vitality assets rapidly, conceivably causing disengagement in the network.

TABLE I ROUTING PROTOCOL FOR TREE, CLUSTER, MULTIPATH AND HYBRID APPROACH

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

## IV.    Conclusions

In this paper we show wireless sensor network is comprise a substantial number of sensor hub. Furthermore, these hubs are asset limitation. That is the reason lifetime of the network is constrained so the different methodologies or convention has been proposed for expanding the lifetime of the wireless sensor network. In this paper we talk about the data aggregation are one of the vital systems for upgrading the life time of the network. Furthermore, security issues is data respectability with the assistance of honesty we decrease the bargained sensor source hubs or aggregator hubs from altogether adjusting the last aggregation esteem. Sensor hub in a sensor network is effectively to bargained. Bargained hubs have a capacity to adjust or dispose of messages. Technique for secure data aggregation: There are two kind of strategy for securing data hop by hop encryption and end to end encryption, the two strategies takes after some progression.
1.  Encryption process must be finished by detecting hubs in wireless sensor network.
2.  Unscrambling process must be finished by aggregator hubs.
3.  After that aggregator hubs totals the outcome and afterward encode the outcome once more.
4.  The sink hub gets last amassed outcome and unscramble it once more.
We additionally examine the different methodologies for data aggregation or likewise talk about the favorable position and disservices and different execution measures of the data aggregation.

## References
[1].  I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey On Sensor Networks", IEEE Communications Magazine, Volume 40, Number 8, pp.102-114, 2002.
[2].  T. Arampatzis , J. Lygeros, and S. Manesis, "A Survey Of Applications Of Wireless Sensors And Wireless Sensor Networks", In Mediterranean Conference On Control And Automation MED05, Nicosia, Cyprus, 2005.
[3].  L. Gatani, G. Lo Re, and M. Ortolani, "Robust and Efficient Data Gathering for Wireless Sensor Networks", in Proceeding of the 39th Hawaii International Conference on System Sciences – 2006
[4].  K. Dasgupta, K. Kalpakis, and P. Namjoshi, "An Efficient Clustering-based Heuristic for Data Gathering and Aggregation in Sensor Networks", IEEE 2003

[5]. E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey", IEEE Wireless communication 2007.

[6]. S. Madden et al., "TAG: a Tiny Aggregation Service for Adhoc Sensor Networks," OSDI 2002, Boston, MA, Dec. 2002.

[7]. K. Dasgupta et al., "Maximum Lifetime Data Gathering and Aggregation in Wireless Sensor Networks", In Proc. of IEEE Networks'02 Conference, 2002.

[8]. M. Ding, X. Cheng and G. Xue, "Aggregation Tree Construction in Sensor Networks," 2003 IEEE 58th Vehic. Tech. Conf., vol. 4, no. 4, Oct. 2003, pp. 2168–72.

[9]. K. Vaidhyanathan et al., "Data Aggregation Techniques in Sensor Networks," Technical Report, OSU-CISRC-11/04-TR60, Ohio State University, 2004

[10]. V. Raghunathan, C. Schugers, S Park, and M.B. Srivastava,"Energy-Aware Wireless Microsensor Networks", IEEE Signal Processing Magazine, Volume 19, Number 2, pp. 40-50, 2002.

[11]. Lin F.Y.S., Yen H.H., Lin S.P., Wen Y.F.: 'MAC aware energyefficient data-centric routing in wireless sensor networks'. Proc. IEEE Int. Conf. Commun. (ICC), 2006, vol. 8, pp. 3491–3496

[12]. Krishnamachari B., Estrin D., Wicker S.: 'Modeling datacentric routing in wireless sensor networks', USC Computer Engineering Technical Report, CENG 02-14, 2002.

[13]. H.H. Yen, C.L. Lin: 'Integrated channel assignment and data aggregation routing problem in wireless sensor networks', IET Communications, 2009, Vol. 3, Iss. 5, pp. 784–793.

[14]. Ramesh Rajagopalan and Pramod K. Varshney, Syracure University "Data aggregation techniques in sensor network : A survey" "IEEE communication survey & tutorial.4th quarter 2006