

Effect of Data Aggregation Protocols on Potency of Wireless Sensor Network

Vijay Nandal¹ and Sunita Dahiya²

¹Research Scholar ECE, DCRUST Murthal, India
Email: vijaynandal86@gmail.com

²Assistant Professor ECE, DCRUST Murthal, India
Email: sunitadahiya.ece@gmail.com

Abstract—The wireless sensor networks (WSNs) consists of nodes whose main task is to sense the particular phenomenon that is happening. Numerous uses for these networks exist, such as monitoring natural habitats, disaster preparedness, security, and military use. The size, processing power, and battery life of wireless sensor nodes are all incredibly small. Sensor networks may fail because the batteries they use have a very limited capacity. Data aggregation is a crucial method in these type of networks. With the aid of data aggregation, we lower energy consumption by removing redundancy. In this paper, we discuss the influence on various parameters of the wireless networks when data aggregation is used.

Index Terms—Data Aggregation, Wireless sensor network, Cluster.

I. INTRODUCTION

Sensor devices that are part of WSNs are primarily used for data sensing. These are employed to recognise and react to various environmental signals. Sensors consume less energy because they are small in size. As a result, the most crucial part in WSNs is to conserve energy. Sensors change the form of signals like humidity, pressure, temperature, voltage, light, and so forth. As battery-operated devices, sensors cannot have their batteries changed frequently. To increase battery life, energy consumption should be kept to a minimum for effective network communication. The three stages of a sensor node's primary operation are first is to sense the data by the help of transducers, second is to process the data, and third is to send the data to its desired location. Due to operations like data collision, idle channel listening, data overhearing, and other operations, the communication process uses a significant amount of the limited energy source. Before passing it along to other nodes, sensor nodes can convert unprocessed sensor data into useful information. In hierarchical WSNs, sensor nodes can be found in a variety of topological patterns, including tree-based topology, ring-based topology, and star-based topology. To ensure that tasks are successfully completed, sensor nodes in WSNs collaborate in clusters [11] [14]. One sensor node designated as a CH exists in each cluster (cluster head). The energy resources of the cluster head run out more quickly than those of the other sensor nodes. In order to send the data from source to the destination node many configurations are used like the inter-cluster and intra-cluster configurations use a significant amount of energy. As a result, periodic task reassignment for the cluster head uses high energy, which reduces efficiency that will shortens the network lifetime. By combining the sensed data from relaying nodes, data aggregation (DA) can significantly increase the energy efficiency of data collection [1].

Zone surveillance, environmental monitoring, and event tracking are just a few of the many uses for wireless

sensor networks (WSNs), especially when used in long-term operation modes. Low-power, battery-powered sensor devices with a limited energy source are what set WSNs apart [6]. It is a very hectic job to change the batteries of the sensor nodes due to the dense deployment of these devices. The limited amount of energy should be used wisely to maximise network lifetime overall.

A. Data Aggregation

In wireless sensor networks, data aggregation is a helpful tool for increasing energy efficiency. Data aggregation involves gathering numerous packets at a sink node before sending them to the base station. The same data can be sensed by many sensing components. Because sending the same amount of data to the base station will require more energy, this data is useless to the base station. A specific node will aggregate this data before sending the actual data to the base station. This will use less bandwidth on the network and save energy by reducing its energy consumption. As is well known, a sensor consists of a transducer that transforms an electrical signal from a physical state or quantity [7]. The sensor detects environmental physical conditions such as temperature, pH of soil, moisture, soil content, etc. The microprocessor or microcontroller located in the processing unit will then process this data. Using a variety of routing protocols, the base station will receive the data that the sensor node has collected. Because the communication component of the wireless sensor network uses more energy than the other processes, the data aggregation aids in lowering the battery's energy consumption. Because some nodes are far from and some are close to the base station, sending data from every node to the base station would require more energy [12]. Therefore, the farthest node will use more energy than the other nodes when sending data. Therefore, by aggregating the sensed data to a specific node and then having that node send the data to the base station, data aggregation will aid in reducing the communication energy. As a result of not having to transmit data to a distant node, the sensor nodes' battery life will be increased [9]. Additionally, the data aggregation reduces the number of transmissions, which lowers battery consumption.

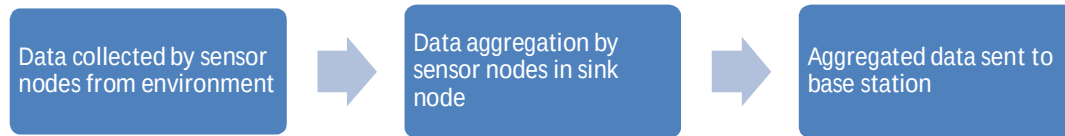


Figure 1. Data Aggregation

TABLE I. MERITS, DEMERITS & CHALLENGES OF DATA AGGREGATION

Data Aggregation	Merits	Demerits	Challenges
	It improves the efficiency of the network and it also enhance the accuracy of the data which is taken across the network.	It cannot be used in all measuring environments.	Data aggregation must be completed within specific time intervals.
	This technique will remove the unwanted data or information.	The cluster head could be subjected to failure.	There should be highly efficient protocols that can be used to save the energy of the nodes.
	There is a reduction in traffic.	Data collection may use more energy when duplicate data is received at the node.	Depending on the application, different compression rates must be used.

TABLE II: ABBREVIATIONS

Notations	Full Forms
WSN	Wireless sensor Network
DA	Data Aggregation
CH	Cluster Head
DP	Data packets
FN	Flat Network
CN	Centralised
IN	In-network
TB	Tree-based
SN	Source nodes
IoT	Internet of Things

B. Parameters

TABLE III. VARIOUS PARAMETERS IN DATA AGGREGATION

Parameters	Description
Confidentiality	Confidentiality is the property of transmitted messages that allows them to only be understood and accessed by the intended recipient. By doing this, it ensures that nefarious parties cannot eavesdrop on the sensor.
Integrity	Integrity is the capacity to prevent unauthorised parties from changing or erasing data packets while communication is occurring.
Availability	Users' ability to access the sensor services and benefits in the face of attacks is referred to as availability.
Lifetime	This is described as the longest period of time for which the deployed sensors can monitor the relevant phenomena.
Accuracy	Accuracy is the difference between the result and the actual data.
Robustness	A system's resilience to change without departing from its initially stable configuration is referred to as robustness, also known as fault-tolerance.

II. RELATED WORK

The following authors have done cluster-based data aggregation research.

Kemal et al [4] studied the impact on various parameters such as latency, security, fault tolerance mechanism on various data aggregation protocols and they also studied the aggregation of multimedia data. Lucy Dash et al [1] discussed that in order to increase the battery life the redundant data must be eliminated. The duplicate data must not be sent to the base station. The authors proposed a methodology to reduce the redundant data by using STCDRR protocol. Mousam et al [6] studied the way of communication in sensor network. They also studied that the data aggregation reduces the power consumption of the batteries of the nodes. Iman et al [2] surveyed different types of data aggregation protocols that are based on number of factors like algorithms. All techniques have the same motive to find out a way to increase the efficiency. Vijay Nandal et al [8] studied the various communication protocols that are used in the wireless sensor networks. They also studied that the data aggregation protocols can be a useful tool in the agriculture by the help of IoT. Bhaskar et al [3] modelled the performance of the various networks that use data aggregation. This model helped to identify the location of cluster whether they are located randomly or closely to each other. Ignacio et al [5] focused on the networks that use high number of sensing nodes and those nodes sensed a lot of data and that data is transmitted to sink nodes which are few in numbers. Ramesh et al [10] studied the merits and demerits of the various DA protocols. They also highlighted the need of energy efficiency in the DA protocols. They also focused on the security concerns in the DA networks. Suat Ozdemir et al [7] focused on the security in the data aggregation network. They provided various ways of security in the wireless sensor networks. They also summarized the different DA protocols. Sankardas et al [9] studied the various security concerns in the DA networks. They studied that a compromised node can degrade the whole system if it left unattended. Xiaowu et al [14] highlighted the various factors of security in the DA networks. They studied that security goals are different for different characteristics of WSNs. Hongjuan et al [13] focused on the communication overhead. It studied that the DA protocols can also reduce the communication overhead problem in the networks. Yao Lu et al [11] created a theoretical model of energy to analyse the merits of DA networks. The impact on DA by various parameters is also discussed in this paper. Vijay Nandal et al [12] proposed a DA protocol which is energy efficient and can be used for the smart farming. This protocol enhances the life cycle of the nodes. Sukhchandan et al [15] studied various methodologies of the aggregation of the data and also studied their various parameters and they compared the parameters of different approaches of the data aggregation.

III. PROTOCOLS FOR DATA AGGREGATION BASED ON NETWORK ARCHITECTURE

This section will describe how the two types of networks—flat networks and hierarchical networks—have an impact on data aggregation.

A. Flat Network

Since all sensor nodes in a WSN have batteries that are similarly powered and perform similar functions within the network, flat networks play a crucial role in WSNs. This network requires data that will be obtained by centrally routing in the process of DA while the sink sends data packets to the sensor nodes, much like it does in floods. Flooding sensors send information that corresponds to the DP, and they return the packets of response data to the sink.

B. Hierarchical Networks

The fact that any type of computing and communication puts a strain on the sink is a common cause of the extensive energy use rates in FN. Data packets sent to the sink in hierarchical networks are fewer because a special node aggregates the data. As a result, hierarchical networks, such as CN, IN, TB, and cluster-based networks, have a structure that increases their overall energy efficiency.

IV. DATA AGGREGATION TECHNIQUES

Data aggregation can be performed using a variety of methodologies, including first is centralised, second is tree-based, third is cluster-based, and fourth is in-network aggregation.

TABLE IV. DATA AGGREGATION TECHNIQUES

Technique	Description
Centralised Approach	The central node, which is typically the most powerful node, receives the sensed data from each node using this method. It can receive many parameters like how much energy is left or consumed or the bandwidth used. The cost metric is taken into account at each intermediate node when the multi hop algorithm is combined with address-centric routing. Data from other nodes must be combined and reported to the base station by the central node.
In-Network Aggregation	The processing of data takes place at nodes which is between the nodes to save important resources like energy and processing time. By attempting to lower energy consumption at each node, this strategy also lengthens network lifetime.
Tree Based Approach	This type of method uses an aggregation tree—typically a minimum spanning tree—is first built. In this approach, the nodes that act as a leaf of the tree act as the SN, the nodes that are between the nodes act as the parent nodes, and the node that acts as root as the base station. The leaf node sends its sensed node to its parent node using a path that was found between it and the base station.
Cluster based approach	With this method, the interest area is divided into many clusters. A cluster head will be elected; whose primary responsibility is to aggregate the data. Instead of sending the information directly to the base station, the node that has given the task of sensing detects the desired parameter and send it to the cluster head. A network can thus conserve a significant amount of energy.

V. IMPACT OF DATA AGGREGATION ON WIRELESS SENSOR NETWORK EFFECTIVENESS

A. Data confidentiality

Data security is significantly influenced by data confidentiality. The data must not be seen or felt by parties who are not authorised. The sensor node that is gathering data cannot authorise other nodes to share the data it has sensed. Between the transmitter and receiver, a suitable communication path must be established. This path must be secure to prevent interference from unauthorised nodes. The sensed data can be encrypted to maintain confidentiality. It implies that the sensed information must be transformed into a code that can only be decoded by the receiver. By doing this, the true significance of the sensed data will be concealed. The encrypted data will not be aggregated during the data gathering process. The protocols will decrypt any encrypted data that is being sent to the node for aggregation. The data will then be combined and encrypted after that. Because the data should be encrypted before transmission and should be decrypted at the cluster head, this process jeopardises the security of the data. It will jeopardise complete confidentiality. Additionally, there is a delay and more energy is used.

B. Data integrity and freshness

In terms of the security of the data, data integrity is also very important. The data that is being sent cannot be altered in any way. The process of encryption and decryption won't shield the data from being altered. At that node, the data that has been gathered at the cluster head can be adjusted. Data will be altered or completely changed in meaning if the cluster head is corrupt. Data can be altered either when there are outside interferences or not. This will occur as a result of the communication channels being unreliable. Data integrity can be compromised if the encryption process is not carried out correctly. The base station will only collect the data that is sent to it by the various cluster heads if the cluster head is corrupt and is sending malicious data to the sink node or base station. In this case, the base station has no way to verify the accuracy of the data. There is no procedure in place to verify the validity of the data. Data freshness ensures that the transmitted data is up-to-date, protecting data aggregation schemes from replay attacks.

C. Verification of Source

In order to send the correct data to its intended recipient, the source's verification is also crucial. The corrupted data may occasionally be sent to the cluster head or the base station by malicious nodes that pose as cluster nodes. The malicious node starts to send the malicious data to the network while concealing its identity. As a result of the network being unable to recognise the genuine node, the entire system will be compromised. On occasion, very sensitive data is sent over the network, and if a corrupted node enters the system, it will have tampered with the data and corrupted the entire process. By utilising source authentication, this issue can be resolved. The sender's identity must be verified by the receiving node. Whether the sender is authorised or not must be verified. A method of encryption must be used by the sender and the receiver. For the data to be authorised, they need a secret key or code. By doing this, false nodes won't be able to join the system.

D. Availability

Another crucial element is the nodes' availability. In the contemporary world, denial-of-service attacks are quite common. The network is made up of many layers, and hackers can attack any layer. Any sensor node has a very short battery life. When a particular cluster of sensor nodes is working harder than necessary, the battery will discharge quickly and the system will suffer. Because the system as a whole will be compromised by the network segment whose battery life is low in comparison to the other network segments. The likelihood of losing the war will increase if a network of sensor nodes is deployed at the battlefield and is transmitting vital information about the advancing force when, due to overwork or subpar protocol measures, the sensor node's battery quickly discharges. Therefore, availability is crucial to the network system. In order to improve energy efficiency, proper data aggregation protocols are required.

VI. COMPARISON OF VARIOUS METRICS OF DATA AGGREGATION PROTOCOLS

There are numerous data aggregation methods that influence the system's latency, fault tolerance, energy use, and accuracy. The use of different data aggregation protocols may have an impact on these parameters. Some protocols may impact latency or accuracy, while others may impact energy consumption. The following table lists the various protocols that have an impact on various parameters. This table compares various protocols that have an impact on a specific metric.

TABLE V. COMPARISON OF DATA AGGREGATION TECHNIQUES IN TERMS OF DIFFERENT METRICS

S.No.	Protocol	Metrics			
		Latency	Energy consumption	Accuracy	Fault Tolerance
1	Optimal data aggregation [16]		Low		
2	QAM-based aggregation [17]	Low			
3	Energy-accuracy tradeoff [21]		Low	High	
4	Synopsis diffusion [24]				High
5	RBC [19]				High
6	Multi-level fusion [25]	Low		High	
7	Data aggregation with path sharing [26]		Low		
8	EScan [20]		Low		
9	Data centric routing [27]	Low	Low		
10	WFQ-based aggregation [22]	Low	Low		
11	Lites [18]				High
12	TAG [28]	Low		High	
13	Fjords [29]	Low		High	
14	TiNA [23]	Low		High	

VI. CONCLUSION

The energy conservation of the sensor nodes is greatly aided by data aggregation techniques. These methods make use of a number of protocols to extend battery life. The Data Aggregation Techniques do, however, have a drawback. This paper provides an overview of how data aggregation techniques affect wireless sensor network performance. These techniques have an impact on a number of parameters. While creating any data aggregation technique, there are a few restrictions that should be taken into consideration based on the survey. We hope to take these limitations into account when creating any protocol for data aggregation.

REFERENCES

- [1] Lucy Dash , Binod Kumar Pattanayak , Sambit Kumar Mishra , Kshira Sagar Sahoo, "A Data Aggregation Approach Exploiting Spatial and Temporal Correlation among Sensor Data in Wireless Sensor Networks," *Electronics*, 2022.
- [2] Iman Dakhil Idan Saeedi, Ali Kadhum M. Al-Qurat, "A Systematic Review of Data Aggregation Techniques in Wireless Sensor Networks," *Iraqi Academics Syndicate International Conference for Pure and Applied Sciences*, 2021.
- [3] Bhaskar Krishnamachari, Deborah Estrin, Stephen Wicker, " The Impact of Data Aggregation in Wireless Sensor Networks," *International Conference on Distributed Computing Systems Workshops*, IEEE, 2002.
- [4] Kemal Akkaya, Murat Demirbas and R. Savas Aygun, "The impact of data aggregation on the performance of wireless sensor networks," *WIRELESS COMMUNICATIONS AND MOBILE COMPUTING*, 2006.
- [5] Ignacio Solis and Katia Obraczka, " The Impact of Timing in Data Aggregation for Sensor Networks," *IEEE*, 2004.
- [6] Mousam Dagar, Shilpa Mahajan, " Data Aggregation in Wireless Sensor Network: A Survey," *International Journal of Information and Computation Technology*, 2013.
- [7] Suat Ozdemir, Yang Xiao, " Secure data aggregation in wireless sensor networks: A comprehensive overview," *Computer Networks*, 2009.
- [8] Vijay Nandal, Sunita Dahiya, " IoT Based Energy-Efficient Data Aggregation Wireless Sensor Network in Agriculture: A Review," *PSYCHOLOGY AND EDUCATION*, 2021, 58(1): 2985-3007.
- [9] Sankardas Roy, Mauro Conti, Sanjeev Setia, and Sushil Jajodia, " Secure Data Aggregation in Wireless Sensor Networks," *IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY*, 2012.
- [10] Ramesh Rajagopalan, Pramod K. Varshney, "Data aggregation techniques in sensor networks: A survey," *Surface*, 2006.
- [11] Yao Lu , Pierre Kuonen, Beat Hirsbrunner, Mingwei Lin, " Benefits of data aggregation on energy consumption in wireless sensor networks," *IET Communications*, 2017.
- [12] Vijay Nandal, Sunita Dahiya, "Energy Efficient Data Aggregation Protocol Based Iot-WSN Framework For Smart Agriculture," *Webology*, Volume 18, Number 6, 2021.
- [13] Hongjuan Li, Kai Lin, Keqiu Li, "Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks," *Computer Communications*, 2011.
- [14] Xiaowu Liu, Jiguo Yu, Weifeng Lv, Yinglong Wang, "Data Aggregation in Wireless Sensor Networks: From the Perspective of Security," *IEEE INTERNET OF THINGS JOURNAL*, 2019.
- [15] Sukhchandan Randhawa, Sushma Jain, "Data Aggregation in Wireless Sensor Networks: Previous Research, Current Status and Future Directions," *Wireless Pers Commun*, 2017.
- [16] Enachescu M, Goel A, Govindan R, Motwani R., "Scale free aggregation in sensor networks," *Proceedings of the First International Workshop on Algorithmic Aspects of Wireless Sensor Networks (Algosensors)*, 2004.
- [17] Yu Y, Krishnamachari B, Prasanna VK, "Energy-latency tradeoffs for data gathering in wireless sensor networks," *Proceedings of IEEE Infocom*, 2004.
- [18] Arora A, Dutta P, Bapat S, Kulathumani V, Zhang H, Naik V, Mittal V, Cao H, Demirbas M, Gouda M, Choi Y-R, Herman T, Kulkarni SS, Arumugam U, Nesterenko M, Vora A, Miyashita M, "A line in the sand: a wireless sensor network for target detection, classification, and tracking," *In computer networks (Elsevier)*, 2004.
- [19] Przydatek B, Song D, Perrig A, "SIA: secure information aggregation in sensor networks," *Proceedings of the 2nd ACM Conference on Embedded Network Sensor Systems (SenSys)*, 2004.
- [20] Zhao YJ, Govindan R, Estrin D, "Residual energy scan for monitoring sensor networks," *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'02)*, 2002.
- [21] Boulis A, Ganeriwal S, Srivastava MB, "Aggregation in sensor networks: an energy-accuracy tradeoff," *Elsevier Ad-hoc Networks Journal (special issue on sensor network protocols and applications)*, 2003.
- [22] Akkaya K, Younis M, Youssef M, "Efficient aggregation of Delay-Constrained Data in wireless sensor networks," *In Proceedings of Internet Compatible QoS in Ad Hoc Wireless Networks*, 2005.
- [23] Sharaf MA, Beaver J, Labrinidis A, Chrysanthos PK, "TiNA: a scheme for temporal coherency-aware in-network aggregation," *Proceedings ACM MobiDE'03*, 2003.
- [24] Nath S, Gibbons PB, Seshan S, Anderson Z, "Synopsis diffusion for robust aggregation in sensor networks," *Proceedings ACM SenSys*, 2004.
- [25] Yuan W, Krishnamurthy SV, Tripathi SK, "Synchronization of multiple levels of data fusion in wireless sensor networks," *Proceedings IEEE GLOBECOM*, 2003.
- [26] Intanagonwiwat C, Estrin D, Govindan R, Heidemann J, "Impact of network density on data aggregation in wireless sensor networks," *Proceedings of the 22nd International Conference on Distributed Computing Systems*, IEEE, 2002.
- [27] Krishnamachari B, Estrin D, Wicker S, "Modeling data-centric routing in wireless sensor networks," *Proceedings IEEE Infocom 2002*, New York, NY, 2002.
- [28] Madden S, Franklin M, Hellerstein J, Hong W, "TAG: a tiny aggregation service for ad hoc sensor networks," *Proceedings of OSDI Conf.*, 2002.
- [29] Madden S, Franklin M, "Fjording the stream: an architecture for queries over streaming sensor data," *Proceedings 18th International Conference on Data Engineering (ICDE'02)*, 2002.