Asian Journal of Science, Engineering and Technology (AJSET)
eISSN: 3006-6980 (online)
https://ideapublishers.org/index.php/ajset
Vol. 4, No. 1 (January-December), 2025, pp. 173-182
https://doi.org/10.47264/idea.ajset/4.1.11
Original Research Article

Reference pattern-based data aggregation in a wireless sensor network

Tahir Saleem[1]*, Khadija Sarwar[2], Umar Ayaz Khan[3], Muhammad Shaheer[1], Hannan Adeel[4], Inamur Rehman Rao[5]

1. Department of Computing, Hamdard University, Islamabad Campus, Islamabad, Pakistan.
2. Faculty of Computer Science & Engineering, Ghulam Ishaq Khan Institute of EST, Swabi, Pakistan.
3. Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan.
4. Department of Intelligent Computing and Analytics, Universiti Teknikal Malaysia, Melaka, Malaysia.
5. Department of Information Technology, Hamdard University, Islamabad Campus, Pakistan.

*Corresponding Author Email: tahir.saleem@hamdard.edu.pk

Article History

Received:
25-Feb-2025

Revised:
07-Aug-2025

Re-revised:
28-Aug-2025

Accepted:
29-Aug-2025

Published:
12-Dec-2025

Abstract:
Data aggregation is a proven technique for mitigating issues in wireless sensor networks. In this paper, reference pattern-based data aggregation in a wireless sensor network is utilised for data aggregation to conserve energy and eliminate redundancy in a wireless sensor network. In traditional data aggregation, each node forwards its data to the cluster head, which aggregates it and then forwards it to the base station. This process takes higher energy. However, in energy-efficient secure pattern-based data aggregation (ESPDA), each sensor node initially generates a pattern code from a lookup table based on the sensed data and then transmits the data to the cluster head. Pattern codes are indicative of sensed data and compact in size. Production of pattern codes from a lookup table is problematic. First, creating a lookup table consumes more computational resources; second, searching for critical values within the corresponding interval also consumes more computational resources and energy. In this research, a new mechanism has been proposed to reduce the consumption of computational and energy resources. We used a reference-based mechanism to create the pattern codes. The simulation results indicate that the proposed mechanism is more energy and computation-efficient than ESPDA.

Keywords: Pattern code, Reference value, Cluster, Computing hardware, Simulation results, Energy consumption, Energy conservation, Data transmission, Sensed data.

How to Cite: Saleem, T., Sarwar, K., Khan, U. A., Shaheer, U., Adeel, A., & Rao, I. R. (2025). Reference pattern-based data aggregation in a wireless sensor network. *Asian Journal of Science, Engineering and Technology (AJSET), 4*(1), 173-182. https://doi.org/10.47264/idea.ajset/4.1.11

1. Introduction

Advances in computing hardware have enabled the development of wireless sensors that can rapidly sense and report various real-world phenomena. Such systems, however, are plagued by bandwidth, power, and throughput limitations, which restrict the volume of information that can be delivered end-to-end. Wireless sensor networks (WSNs) have emerged recently and gained widespread recognition. Such a network comprises sensors, and sensor nodes are organised in different topologies. The design of sensor nodes has been altered by the emergence of micro-electromechanical systems (MEMS) and wireless communications. Thus, the capabilities of each node have been significantly enhanced by computation, communication, and memory. Sensor nodes are still devices that must contend with energy limitations, unlike other wireless devices such as laptops (Sharmin *et al.*, 2023; Yuan & Gao, 2025). The process of inventing the battery has not been as quick as the process of computer inventions. The lifetime of WSNs depends on the battery life. WSN nodes will not have the chance of battery replacement and recharging (Sharmin *et al.*, 2023). Extending the lifetime and conserving energy in WSNs are significant research challenges due to the small, non-replaceable batteries in the sensors (Khedhiri *et al.*, 2025). A sensor node measures environmental conditions, such as temperature and pressure. WSN is a highly resource-strained network. These sensors are produced in large quantities and cooperate to form an ad hoc network that reports data to a sink or base station, where the data is collected (Murthy & Manoj, 2004).

WSNs mainly use a hierarchical structure. Sensors are partitioned into clusters, and a single node is chosen as the cluster head. Sensors transmit data to the base station through the cluster head. Cluster heads are dynamically changed based on the residual energy of nodes in a cluster (Reddy *et al.*, 2023). The cluster head is also responsible for relaying the sensor status to the base station (Reddy *et al.*, 2023). WSNs present numerous challenges due to their wireless nature. Hostile deployment is one of them, and nodes are not safe against physical access. The other one is network communication instability. The critical issue is how to extend the lifetime of sensor nodes, as it would not be practical to change the batteries of many sensors. However, all these difficulties are somehow overcome, and some of them are under investigation. Their resources are scarce and in short supply, with limited battery power, computing power, memory, and channels. This will lead to focusing on minimising storage overhead, computational cost, and energy consumption by reducing the transmission of unnecessary information over the network (Saadallah & Alabady, 2024). The fundamental theme is to maximise network lifetime by minimising resource depletion in sensor nodes. Therefore, there is an immediate need to reduce the size of transmitted data by removing redundant information, thereby enabling data aggregation.

This paper has two main objectives. The first to eliminate redundant data from WSN using a reference pattern-based data aggregation technique. The second objective is to extend the lifetime of the wireless sensor network by preventing duplicate transmissions on the sensor side through data aggregation. Additional care is taken to maintain the energy-efficient aggregation technique, aiming to achieve the maximum network lifetime compared with other aggregation techniques. In our method, we assume that a cluster has already formed.

2. Literature review

Minimising storage overhead and computational complexity, and reducing energy consumption in data transmission, are the primary concerns in WSNs. Among the plans, one

option is to use data aggregation. A survey of protocols from the prior literature on secure data aggregation in WSN is presented below.

The first one is the LEACH protocol proposed by Heinzelman (2000). In this protocol, nodes are distributed and arranged inside a cluster to fuse the information. LEACH has two phases: the setup phase and the stabilisation phase. The first step is to group the network into clusters and elect a cluster head for each cluster. The second stage consists of aggregating data and transmitting it to the base station. LEACH improves data accuracy and extends the WSN lifetime. However, this protocol has some restrictions. One is that LEACH assumes that every node must be capable of aggregating data and acting as the cluster head. This assumption might not hold for sensors with limited energy.

SDA is proposed by Hu and Evans (2003). This protocol achieves this by delaying aggregation and authentication to higher-level nodes, thereby increasing the compromised node's flexibility. Sensor nodes transmitted data without any aggregation or change at the second node. Data cannot be aggregated at the closest node in immediate proximity. Sensors waiting for data until the base station cannot reveal the shared key. Data freshness and integrity are ensured by this protocol. Once the root (parent) and child nodes in the hierarchy are violated, data certainty increases, and integrity can be compromised. If a sensor node is identified as having violated the protocol, no action is taken to control the violation. Therefore, this disrupts data availability. Furthermore, the memory overhead at the sensor level is increased because sensors cache data until the base station discloses the key.

The SDAP protocol is proposed by Ozdemir and Xiao (2009). SDAP relies on the following two principles: Divided and conquer, committed and attentive. SDAP divided the network tree into a set of logically sub-trees so that the number of aggregators can be increased and the number of nodes in each sub-tree can be reduced. This incorporation of the committed-and-attested principle provokes the hop-by-hop design and assists at the base station in the correctness of aggregation. After a node in a subtree commits its result, it cannot be rejected. To provide sufficient security, this scheme must convey additional data, which increases the communication cost.

Chan *et al.* (2006) build on the SIA and instantiate it with the aggregate-commit-prove technique. In this protocol, the sensor utilises the Merkle hash tree to apply the aggregation function when retrieving data from its child nodes. It provides assurance about the input data utilised to calculate the aggregation. Then it forwards the commitment and aggregated data to parent nodes until the base station receives them. When the base station gets the last commitment values, it rebroadcasts them to the entire network. Each node needs to verify whether its commitment is attached. When it becomes aware that its commitment has been added, the node sends a message authentication code to the base station. Ignorance of one code results in the rejection of the aggregation result. This is a key disadvantage, and another is the high computational and transmission costs.

Çam *et al.* (2006) present ESPDA. This plan combines security and aggregation in cluster-based WSNs. This protocol avoids redundancy between sensor nodes and cluster heads. With security challenges in place, pattern code comparisons are performed, and based on these comparisons, the cluster head determines which sensor should send data to the base station. This contributes to increased energy efficiency and the effective use of bandwidth. Between nodes and the cluster head, redundant transmissions are ignored. In this protocol, a sleep-active

mode is shown. In this mode, the number of active nodes is reduced through overlapping-area monitoring. This protocol is secure to the extent that it reduces the amount of data sent from the sensors to the cluster head using pattern codes. This protocol employs a lookup table that contains a large number. The figures are split into ranges. The threshold values fix the width and boundaries of the interval. The PG algorithm utilises the seeds of the lookup table pattern to map the critical values of the correspondence interval. The lookup table interval defines the range of all the intervals. The critical values in the lookup table assign an essential value to every interval.

When the detected information is sent to the environment, the parameter interval is compared against the interval defined in the lookup table of the pattern-generating algorithms, and each parameter is assigned a critical value. These values are being cross-referenced in light of the provided parameter, which is the pattern code of the sensed parameter. The sensor also inserts its ID, the timestamp, and the pattern codes, and transfers them to the cluster head. As a result of the lookup table, the following problem occurs. The interval size is unequal, which creates an inaccurate situation, and the lookup table requires significant storage overhead. Every sensor node calculates the lookup table and re-computes it after a preset time period, thereby increasing the computational cost. Later on, OzgurSanli et al (2004) proposed Secure Reference-Based Data Aggregation (SRDA). In this protocol, the base station sends the reference value to the sensor network. Sensor nodes compute the difference between the sensed and reference data and relay the differentially data to the base station instead of relaying the actual data. Differential data occupy less space than the original data, thereby reducing energy consumption.

3. Reference pattern-based data aggregation (ESRPDA)

ESPDA can secure information using pattern codes, as the adversary cannot determine whether the data is real. Having more intervals and making the interval size smaller will increase efficiency. This, however, complicates the search mechanism and also increases the computational cost of the lookup table. So, we applied the mechanism of reference value. In the reference value mechanism, global reference values are transmitted by the base station in encrypted form to all cluster heads during a specific period, and the cluster heads then forward these values to the sensor nodes. The reference value must be chosen so that it reflects the actual parameter in the environment, and it should be higher than the sensor's normal reading, except during critical events (e.g., a fire). The base station should update the reference value several times at set time intervals. When sensor nodes receive the reference data, they decrypt it and subtract their sensor data from it. Once subtraction occurs, the nodes select the leftmost digit of the result and integrate it with the results of other parameters to form pattern codes. The required pattern codes are the digits concatenated. In the same way, every node reacts and produces the pattern codes of its sense data. The algorithm is described below.

Pattern generation algorithm: At the node side:

Step 1: Initialise variable PC // The pattern code is initialised

Step 2: If (time ‘T’ received reference value)

Then,

Step 3: For i=1 to n

A: Each parameter of sense data is subtracted from reference values

B: PC = Left-most digit of the result

C: PC = PC + [PC] // Concatenate to create pattern code

End of

Step 4: PC = PC+ [Timestamp]+ [SensorID]

End of If

The execution of this algorithm occurs in all nodes of the cluster, creating the pattern codes. When pattern codes are generated, the nodes forward them to the cluster head, and the cluster head executes the pattern comparison algorithm to identify nodes that generate redundant information. In each cluster head, a pattern comparison algorithm is performed. The algorithm can be seen below.

Pattern comparison algorithm: Operate at the cluster head

Step 1: Broadcasted reference values.

Step 2: While (T is timeout)

Go to step 1.

A: Get Pattern code, Timestamp, SensorID

B: Compare pattern codes and select nodes of different codes

C: If (sensor is selected)

Then, request to send actual data.

End of if

End of while

D: Remaining nodes are put to sleep mode.

The chosen nodes transmit actual data, while the others are abandoned by transmitting an acknowledgement and then placed into sleep mode to conserve energy. Sleep mode implies switching off the node's sensing unit for an extended period of time, rather than turning off the radio unit. If the periods of the sleep nodes have lapsed, these nodes turn on. If the same node produces the same pattern codes, the sleep time will be doubled during the succeeding period.

Suppose five sensor nodes sense Temperature p1, humidity p2, and pressure p3. All cluster heads broadcast the reference values for all parameters to the nodes. Then, at the node side, it subtracts the sensed value from the reference value. The left-most digit of the result of each parameter is picked after subtraction, and the digits of the results are concatenated. For example, the cluster head sends a reference value (100, 100, 100). The pattern generation is presented in Table-1.

Table-1: Pattern codes

Sensor	Reference value	Sense values	Subtraction	Pattern code
Sensor1 (p1, p2, p3)	(100,100,100)	(56,92,70)	(44,04,30)	403
Sensor2 (p1, p2, p3)	(100,100,100)	(56,92,70)	(44,04,30)	403
Sensor3 (p1, p2, p3)	(100,100,100)	(56,92,70)	(44,04,30)	403
Sensor4 (p1, p2, p3)	(100,100,100)	(56,92,70)	(44,04,30)	403
Sensor1 (p1, p2, p3)	(100,100,100)	(56,92,70)	(44,04,30)	403

The leftmost digits in Table-1, 4, 0, and 3, correspond to sensor 1. Hence, these are taken and concatenated to form the pattern code of this node. The same process is applied to the remaining nodes. Thus, in Table-1, nodes 1 and 3 produce identical patterns, and nodes 2, 4, and 5 also form similar patterns. That implies that nodes 1 and 3 detected redundant data, and nodes 2, 4, and 5 also detected redundancy.

4. Simulation and result discussion

The simulation is performed using NS2 version 2.33 using the SensorSim extension. The modelled system is supposed to be a homogeneous WSN. In the simulation, the network has one hundred sensor nodes and three clusters. The physical environment is the same for all sensor nodes, and all nodes are arranged in a grid form. Parameter values are given in Table-2.

Table-2: Simulation parameters

Variables	Values
Simulation tool	NS2.3 with Sensors Sim
Node energy	1 Joule
Propagation model	Free space
Antenna type	Omni directional
Topology size	200 m x 200 m
Transmitted power	.1 µJ per bit
Receiving power	.05 µJ per bit
Processing power	.1 µJ per second
Data packet size	100 byte
Broadcast packet size	25 byte
Data rate	2.4 kbps
Modulation type	On-off keying
Wireless link type	802.11 (2.4-2.48 Ghz)
Sensor nodes	100
Cluster	3
Simulation time	100 sec
Sleep duration time	8 sec

On the following parameters, as in the above Table 2, the protocol is tested on the basis of various performance parameters. As Figure 1 indicates, the energy consumption of various sensor nodes and the nodes within a single cluster is illustrated. In the event that the ESPDA sensor nodes used more power than the ESRPDA, since the ESPDA nodes went through calculations of the lookup table, which were computed and recomputed after a stipulated amount of time, which was rather consuming. Searching is also participated in the pattern generation of look up table in ESPDA. Thus, the sensor nodes use 5-10 per cent less energy in ESRPDA than ESPDA.

Figure 1: Sensor ID and energy consumption

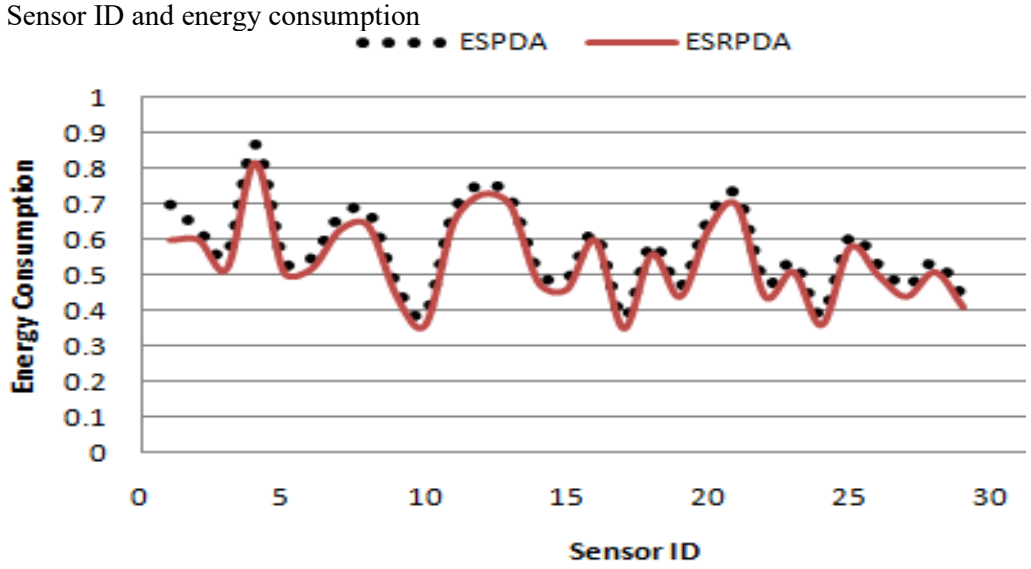
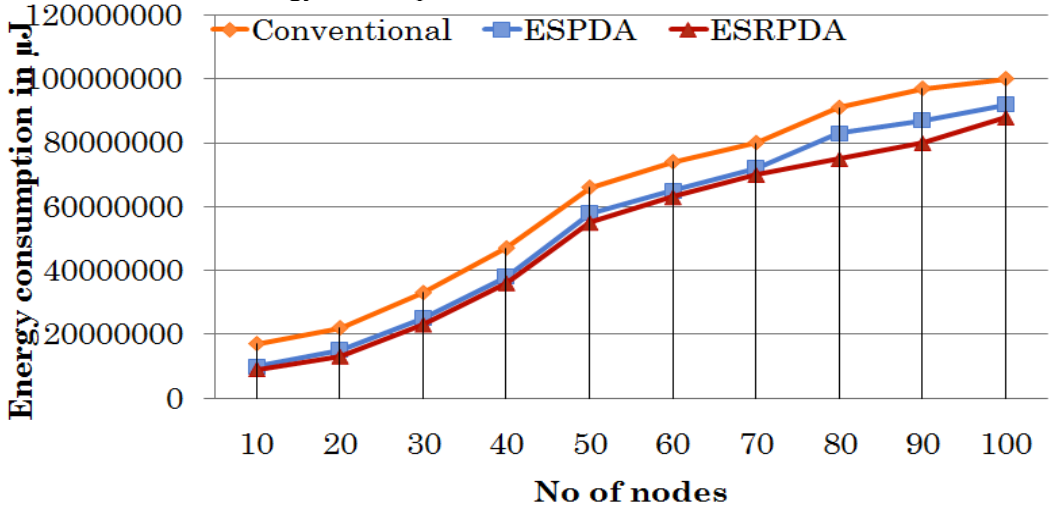


Figure 2 illustrates energy consumption across different nodes. ESRPDA used less energy than the other two schemes due to its lower computational requirements, and there is no redundancy wastage at the cluster head level, as it is handled at the node level. In ESPDA, redundant data at the node level is eliminated, but this approach requires more computation to generate lookup tables. Hence, ESRPDA is both energy-efficient and prolongs network lifetimes, more so than the other two.

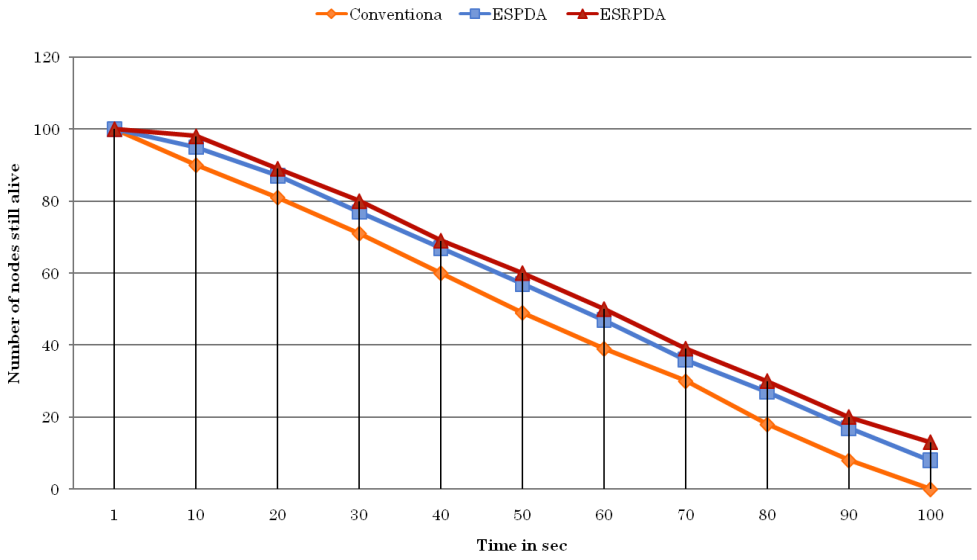
Figure 2: No of nodes and energy consumption



As can be seen in the Figure 3, the available nodes that remain alive and active over time. It is

evident from the graph that the first node dies in conventional after 4 seconds and in ESPDA after 5 seconds; however, in ESRPDA, the first node dies after 7 seconds. Thus, under ESRPDA, more nodes will remain alive, extending the network's lifetime compared to the other two schemes. This is because fewer nodes are alive due to the low energy consumption of conventional and ESPDA schemes.

Figure 3: Time in seconds vs Number of nodes still alive



5. Conclusion and future works

The different data aggregation schemes were studied, and their implications for the lifetime and the energy consumption were reviewed. On this basis, some existing solutions for reducing energy consumption were explained. In our proposed scheme, entitled ESRPDA, special consideration was taken. In that respect, ESRPDA was simulated in NS2.33, and the results were examined.

Compared to other protocols, ESRPDA is more computationally and energy-efficient, though it has limitations that will form the basis of future work. The former includes that this protocol does not permit the intermediary nodes to aggregate data. Data aggregation occurs only at the cluster head, thereby undermining the aggregation advantage. Thus, there is a requirement to use the ESRPDA schemes at intermediate nodes rather than at the cluster head. There is also more future in remaining with a larger sensor network, and in taking into consideration the network's complexity in the PC generation algorithm.

Declaration of conflict of interest

The author(s) declared no potential conflicts of interest(s) with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD:

Tahir Saleem https://orcid.org/0000-0001-7828-5382

Khadija Sarwar https://orcid.org/0009-0003-9461-2613

Muhammad Shaheer https://orcid.org/0000-0003-1928-5926

Inamur Rehman Rao https://orcid.org/0000-0002-9180-9758

Publisher's Note

IDEA Publishers Group (AJSET IDEA-PG) stands neutral with regard to the jurisdictional claims in the published maps and the institutional affiliations.

References

Chan, H., Perrig, A., & Song, D. (2006, October). Secure hierarchical in-network aggregation in sensor networks. In *Proceedings of the 13th ACM conference on Computer and Communications Security* (pp. 278–287). https://doi.org/10.1145/1180405.1180440

Çam, H., Özdemir, S., Nair, P., Muthuavinashiappan, D., & Sanli, H. O. (2006). Energy-efficient secure pattern-based data aggregation for wireless sensor networks. *Computer Communications*, *29*(4), 446–455. https://doi.org/10.1016/j.comcom.2004.12.029

Heinzelman, W. B. (2000). *Application-specific protocol architectures for wireless networks*. Doctoral dissertation, Massachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/26881

Hu, L. & Evans, D. (2003). Secure data aggregation for wireless network. In *Symposium on Applications and the Internet (SAINT) Workshops, IEEE Computer Society* (pp. 384–394).

Khedhiri, K., Ben Omrane, I., Djabour, D., & Cherif, A. (2025, May). Clustering for Lifetime Enhancement in Wireless Sensor Networks. *Telecom, 6*(2), 30. https://doi.org/10.3390/telecom6020030

Murthy, C. S. R., & Manoj, B. S. (2004). *Ad hoc wireless networks: Architectures and protocols*. Pearson Education.

Ozdemir, S., & Xiao, Y. (2009). Secure data aggregation in wireless sensor networks: A comprehensive overview. *Computer Networks*, *53*(12), 2022-2037. https://doi.org/10.1016/j.comnet.2009.02.023

OzgurSanli, H., Ozdemir, S., & Cam, H. (2004, September). SRDA: Secure Reference-Based Data Aggregation Protocol for Wireless Sensor Networks, in *IEEE 60th Conference on Vehicular Technology*, VTC 2004-Fall, Volume 7, pp. 4650–4654, 26–29.

Reddy, M. R, Ravi Chandra, M. L., Venkatramana, P., & Dilli, R. (2023). Energy-efficient cluster head selection in wireless sensor networks using an improved grey wolf optimisation algorithm. *Computers*, *12*(2), 35. https://doi.org/10.3390/computers12020035

Saadallah, N. R., & Alabady, S. A. (2024). An energy-efficient and scalable WSN with enhanced data aggregation accuracy. *Journal of Telecommunications and Information Technology*, (2), 48–57. https://doi.org/10.26636/jtit.2024.2.1510

Sharmin, S., Ahmedy, I., & Md Noor, R. (2023). An energy-efficient data aggregation clustering algorithm for wireless sensor Networks using hybrid PSO. *Energies*, *16*(5), 2487. https://doi.org/10.3390/en16052487

Yuan, H., & Gao, C. (2025). Minimising redundancy in wireless sensor networks using sparse vectors. *Sensors*, *25*(5), 1557. https://doi.org/10.3390/s25051557