



Two level data centric aggregation scheme for wireless sensor networks

Tahira Batool¹ · Atiq Ahmed² · Dominique Gaiti²

Received: 27 December 2022 / Accepted: 10 August 2023 / Published online: 8 September 2023
© Institut Mines-Télécom and Springer Nature Switzerland AG 2023

Abstract

Wireless sensor networks (WSNs) sense and collect information from a desired phenomenon with the help of sensor nodes that have limited computational power, battery, and memory. Several data aggregation approaches are proposed to make the sensor networks energy-efficient, increasing the network's lifetime by controlling data redundancy at aggregator nodes. Redundant data is suppressed before transmission to the sink. In this work, our aim is to enhance the network lifetime by efficiently utilizing the network's energy through controlled data redundancy and minimizing data transmission to the sink. Data aggregation occurs in two steps: firstly, within clusters where the cluster-head serves as the aggregation point, and secondly, at a central point in the network where the gateway node acts as the aggregation point. Experiments demonstrate that our proposed approach yields better results compared to a benchmark clustering protocol in terms of network stability, the number of data packets transferred to the destination, energy dissipation of nodes, and overall network lifetime.

Keywords Wireless sensor networks · Data aggregation · Network lifetime · Energy efficiency

1 Introduction

WSN is a wireless technology that utilizes sensor nodes to sense various phenomena such as air, water, soil, and temperature. These ad hoc networks have a dynamic topology that constantly changes due to node additions and removals. WSNs are designed for specific applications and operate with limited energy and memory resources. Ensuring energy efficiency is crucial for WSNs, as node replacements in case of failure pose a significant challenge, especially in hostile environments. While WSNs provide significant benefits in various applications, they also have certain limitations. In D. Juneja et al. and A. Kumar et al. [1, 2], the authors discuss different issues related to network performance, design,

and quality of service (QoS). These issues often arise due to the limited power available in WSNs. Given the energy constraints of WSNs, efficient energy utilization becomes crucial to enhance the network's lifetime. Key operations that consume the network's energy include data sensing, processing, and transmission.

Data transmission consumes a significant portion of the energy in WSNs, approximately 70%. Reducing the number of transmissions and utilizing data aggregation can help conserve energy. In WSNs, the sensor nodes are divided into head and member nodes, with the head node responsible for data aggregation. Instead of direct node-to-node transmission, data is sent from the head node to the sink. Figure 1 depicts the basic architecture of WSN clustering, consisting of three vital components: (1) sensor node (SN), (2) central node, and (3) sink or base station (BS). SNs are connected to a cluster-head (CH) or local BS, which serves as the central node. Instead of direct transmission, SNs sense data and connect to the CH for data transmission to the sink. This approach avoids data redundancy when nodes send data directly to the BS using address-centric routing. Clustering solves this issue by having SNs transmit data to the CH after sensing it. To reduce redundancy, the CH aggregates the data received from member nodes within the cluster. The CH then sends the aggregated

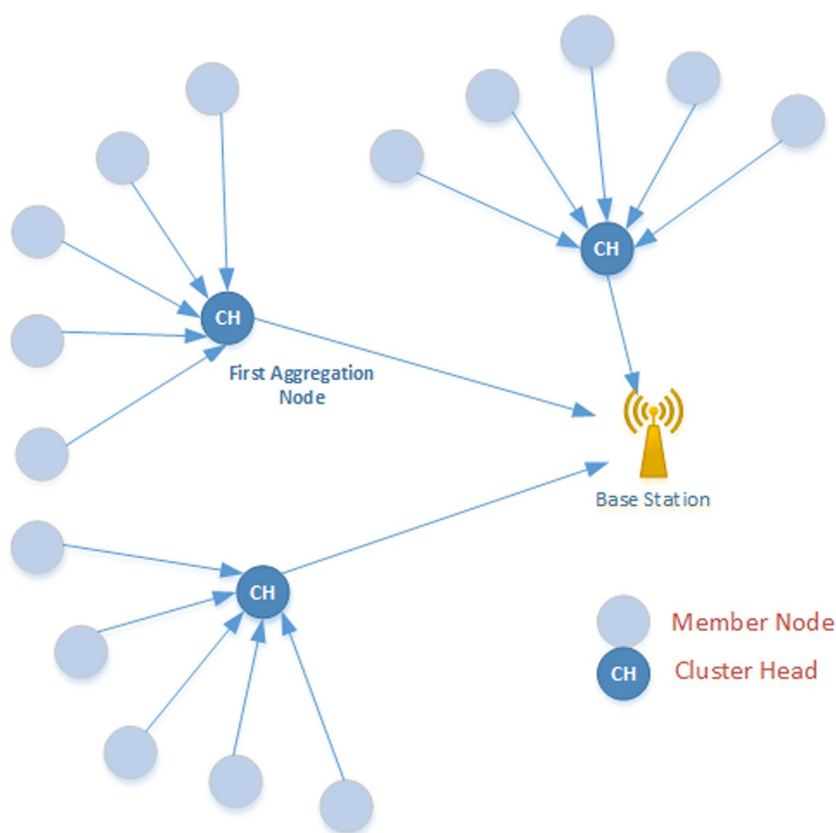
✉ Atiq Ahmed
atiq.ahmed@utt.fr

Tahira Batool
parsabatool@gmail.com

Dominique Gaiti
dominique.gaiti@utt.fr

¹ SBK Women's University, Quetta 87300, Pakistan

² LIST3N, Université de Technologie de Troyes, 12 Rue Marie Curie, 10010 Troyes, France

Fig. 1 Clustering in WSN

data to the BS. Energy usage within each cluster progressively shortens the lifespan of WSNs.

Our work aims to enhance the network lifetime by implementing data aggregation techniques that reduce the number of transmissions to a single long-range transmission to the sink. We introduce two aggregation points, namely the CH within the cluster and the gateway node (GW). These aggregation points effectively reduce data redundancy and the overall number of transmissions from nodes to the sink. By utilizing a single long-range transmission from the GW node to the sink, energy is efficiently utilized. The structure of our work is as follows. Section 2 provides an overview of various existing aggregation schemes, highlighting their advantages, shortcomings, and the primary motivations for proposing a new aggregation method. In Section 3, we present our proposed algorithm, which incorporates a two-level aggregation approach utilizing distinct nodes. This approach aims to minimize energy consumption and increase the sensor network's lifetime. Section 4 comprises the evaluation of our proposed approach, followed by the conclusion in the final section.

2 Related work

Wang et al. [3] propose a scheme that focuses on the computing and transportation capacity of sensors in WSNs. They classify WSNs into two types: random WSNs, which have a fixed monitoring region and expandable scale, and dense WSNs, which have a fixed sensor density but expandable size. The first model determines reliability by adjusting the transmission rate based on the receiver's signal-to-interference and noise ratio (SINR). The second model is a fixed-rate model that allows transmissions only when the data value exceeds a predefined threshold. This scheme provides practical guidelines for designing aggregation schemes and scaling in large networks.

In H. Rahman et al. [4], a QoS-aware hybrid data aggregation scheme for the Internet of Things (IoT) is proposed by combining the features of tree and clustering schemes. This scheme aims to reduce data redundancy, improve energy efficiency, and enhance network lifetime.

The network model consists of statically deployed sensor nodes (SNs). Sink-based CH selection is performed, where SNs send their residual energy and location information to the sink. The results demonstrate that this scheme outperforms Low-energy adaptive clustering hierarchy (LEACH) [5] and LEACH-C in terms of network overhead, latency, and energy dissipation. However, it is applicable only to fixed scenarios.

F. Yuan et al. [6] discuss data density correlation degree and clustering. It introduces the concept of a core SN, which represents data from adjacent nodes. SNs select a core sensor to form clusters, and the maximum correlation value is used to merge clusters and select a representative node. SNs are categorized as isolated SNs, representative SNs, and member SNs, each performing specific tasks of sensing and data transmission. Member SNs only transmit data sent by isolated SNs and representative SNs. The results show that introducing various sensor types improves energy efficiency. However, this scheme is more beneficial in dense areas where the sample data rate changes slowly.

Zhou et al. [7] propose using different data aggregation methods to maximize the lifetime of data gathering trees, which provide both data aggregation and routing for energy efficiency. The methods include: (1) full data aggregation mode, where data is aggregated at each intermediate node; (2) no aggregation mode, where data received at intermediate nodes is only sent with its own data; and (3) hybrid data aggregation, where aggregation is performed if the received data reaches a certain threshold. This approach provides better network lifetime when the transmission range is large but is affected by the number of SNs. Additionally, it lacks scalability for large networks.

To reduce energy consumption, a scheme is proposed in E. Prathima et al. [8] that utilizes a mobile sink for data collection. Mobile sinks generate queries and stay in a region for a predetermined period to receive data. Nodes store the desired data and transmit it to the sink. The results demonstrate a decrease in energy consumption with an increase in the number of mobile sinks. However, the packet delivery ratio decreases as the network size increases.

In L. Villas et al. [9], the authors propose a data aggregation protocol that uses a dynamic and scalable tree structure, reducing the number of messages and improving route selection for high aggregation rates. The scheme achieves better scalability and double the aggregation rate compared to similar protocols by detecting network events for data aggregation. However, this approach needs to address fault tolerance and security issues due to its open nature. Similar methods are presented in N.T. Nguyen et al. and S. Saginbekov and A. Jhumka [10, 11] to maximize WSN lifetime using local tree reconstruction, but they face higher overhead due to the collection of global network information.

In S. Sasirekha and S. Swamynathan [12], the authors combine the features of LEACH and PEGASIS in their proposed approach. The network is divided into clusters, and nodes within each cluster are connected using a chain head that contains aggregated data transferred to the sink. Simulation results show that this approach consumes 12.5% less energy than PEGASIS and Enhanced Coverage Control Protocol (ECCP), and 60% less than LEACH. The proposed scheme also exhibits lower transmission delays compared to LEACH, ECCP, and PEGASIS.

Min et al. [13] propose a secure data aggregation scheme for hierarchical WSNs using renewable hash chain generation, key creation, and data authentication. The scheme provides key updating, authentication of original data, and protection of aggregated data. Data-sending nodes authenticate the data they deliver using one-way hash chains. The receiving node decrypts the data and compares the hash functions to determine data validity. This scheme enhances security against various attacks but has a complex packet design due to its security features.

Hua et al. [14] propose an energy-efficient scheme where the construction of the aggregation tree is initiated from the base station (BS) by sending “Hi” messages to all nodes at one hop distance. Nodes respond with a “Join Request” message, and the BS selects child nodes based on signal strength. Aggregator nodes are selected from leaf nodes using a predefined probability. Nodes communicate at one-hop distance to reduce traffic. The scheme shows reduced traffic, saving energy and extending the network’s lifespan, while privacy remains similar to similar protocols.

Using compressed sensing (CS) data aggregation, Xiang et al. [15] present an energy-efficient and high-fidelity data collection scheme. The proposed model forms a connected graph, achieving high data recovery fidelity at the sink with a high aggregation factor for large networks. Partitioning networks into subnets can save 50–70% of energy costs. Similarly, in X. Xu et al. [16], authors propose a method that combines hierarchical networks and CS, ensuring data accuracy at each step with the CS recovery algorithm. The scheme exhibits a higher signal-to-noise ratio (SNR) compared to other related algorithms, but each level introduces a delay registration overhead.

Authors M. Islam and J. Kim [17] propose a cooperative energy-efficient WSN technique where selected nodes in one cluster transmit data to nodes in other clusters using cooperative multiple input multiple output (MIMO) communication. The scheme optimizes data transmission by employing distributed antennas and considering different channel parameters. The protocol has two functions: (1) it performs well for correlated data by aggregating data in a centralized node before transmission to the sink, and (2) it performs well for uncorrelated data by transmitting data without aggregation at a central node. The impact of correlation on energy

consumption is also investigated, indicating that higher data correlation leads to lower energy consumption. However, the cooperative MIMO scheme focuses only on cluster-to-cluster communication and does not consider node-to-cluster transmission.

Dias et al. [18] propose a dual prediction scheme for WSNs to reduce transmissions and optimize data transmission. The scheme utilizes historical data to forecast future data values. The selection of a prediction scheme at either the gateway or sensor nodes depends on the transmission of data from sensors to the gateway. Nodes at the sensor level make predictions based on information from the gateway and measured data to choose a prediction model, which is then transmitted by the gateway during the model selection process. Results demonstrate that aggregation with prediction models achieves energy efficiency gains of up to 92% compared to aggregation without prediction models.

Ahmed et al. [19] propose a cluster-based aggregation scheme for coffee plantation pest identification. The scheme employs cluster-based data aggregation and a hybrid model that considers time and events. Clusters with sensors on coffee stems divide the sensing region (the coffee field). Ultrasonic sensors detect changes in sound waves to identify the presence of pests. The proposed scheme outperforms previous methods in terms of aggregation ratio, end-to-end delay, overhead, clustering time, and energy consumption. However, it does not incorporate a specific routing mechanism for efficient data transmission.

In Wang Z. et al. [20], a time synchronization cluster-based scheme for industrial WSNs is proposed. The synchronization process starts at the CH instead of the nodes to save energy. CH-to-CH communication occurs through overlapping nodes, while CH-to-node communication is single-hop. The scheme offers lower convergence time, which is dependent on the number of clusters. Fewer clusters result in reduced convergence time and energy consumption. However, the scheme needs to consider the dynamic topology structure for dynamic communication delays.

From the literature discussed above, it can be inferred that clustering, in addition to controlling data redundancy, plays a crucial role in efficient energy utilization in WSNs. Most energy consumption in WSNs is attributed to data transmission. Individual data transmission from sensor nodes to the BS leads to increased data transmission and energy consumption. On the other hand, clustering minimizes data transmission by allowing only the CHs to transmit data, thereby efficiently utilizing energy. However, when data is aggregated into a single packet for efficient data utilization, the energy consumption among nodes may not be balanced, affecting the network's lifetime. To address these issues, in the next section, we propose a two-level aggregation approach with minimal transmission. This approach not only conserves sensor energy but also improves the network

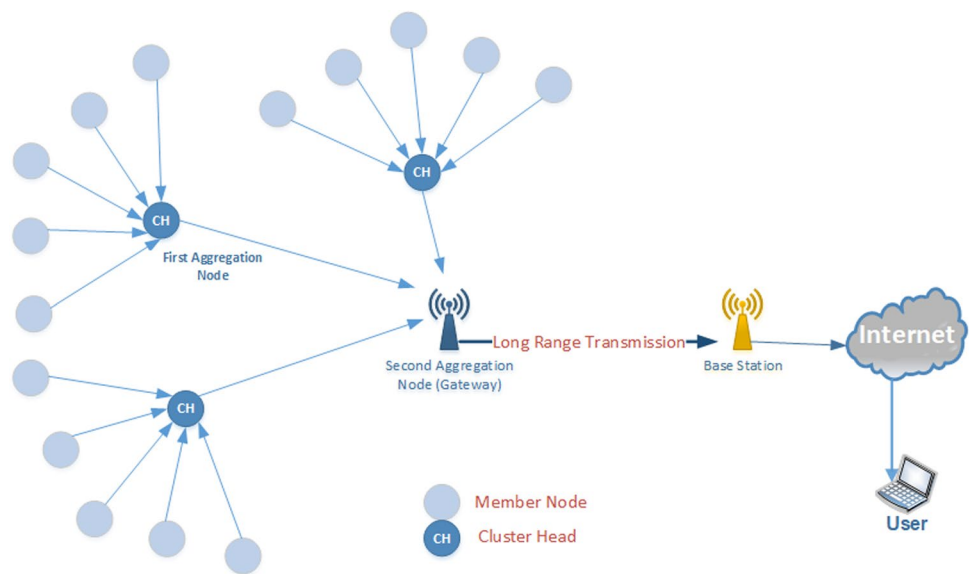
lifetime by reducing the number of communications towards the BS.

3 Dual data aggregation scheme

Data aggregation plays a crucial role in reducing the amount of data that needs to be transmitted to the BS and minimizing data redundancy caused by neighboring nodes sensing the same data. A significant portion of the network's energy is consumed when multiple CHs transmit data to the BS. By introducing CHs in the clustering process, direct transmission from individual nodes to the BS is reduced, resulting in a minimized number of transmissions. In our proposed design, we further enhance this idea by incorporating a GW node positioned at the center of the region. The GW node provides an efficient solution to reduce redundant data and long-distance transmissions. In each cluster, the nodes sense data and transmit it to the respective CH. The collected data from different nodes in each cluster is then aggregated at the GW node. Consequently, data aggregation occurs at two levels, effectively reducing redundancy and the number of transmissions. Only the GW node sends the aggregated data to the BS, optimizing the transmission process. The suggested algorithm is depicted in Fig. 2, illustrating the division of the sensing region into clusters, with member nodes and central nodes in each cluster.

In the proposed scheme, each cluster consists of sensor nodes responsible for data sensing and a central node that serves as the CH. After the sensor nodes in a cluster gather data, they transmit it to the CH. The CH performs data aggregation within the cluster before transmitting the aggregated data to the GW node. This process occurs for each cluster in the network. At the GW node, which acts as the second aggregator node, the received data from all the CHs is aggregated to prevent data redundancy. The aggregated data is then transmitted from the GW node to the BS, which is located near the region of interest. This single long-distance transmission from the GW node to the BS reduces the overall number of transmissions.

For example, let us consider nodes A, B, C, D, E, and F as the sensor nodes in "cluster 1" within the sensing region. These nodes sense the phenomenon and transfer the sensed data to the CH of cluster 1, which acts as the first aggregator node. The CH in cluster 1 aggregates the received data from the member sensor nodes and transmits it to the GW node. At the GW node, which serves as the second aggregator node, all the received data is aggregated before being transmitted to the sink. By aggregating the data at the CH and GW nodes, the scheme ensures that only one long-range transmission is required from the GW node to the BS. This approach effectively reduces data redundancy and minimizes

Fig. 2 Data aggregation at two levels

Input: Nodes, energy, probability for CH selection

Output: Clusters, CHs, Dead nodes, transmission of data

- 1) Node initialization, GW and BS placement
 - a) Nodes are initialized
 - b) GW node and BS locations are determined
- 2) CH selection on the basis of probability
- 3) CH nodes are selected which will perform the task of data aggregation after receiving data sensed by the SNs
 - a) Clusters are formed after the association of nodes to the CH
- 4) Nodes to sense phenomenon and transmit data to CH
 - a) All the nodes start task of sensing the desired phenomenon
 - b) After performing this task, data is transmitted to CH
- 5) Data aggregation at CH
- 6) CHs perform task of data aggregation on the data received from the nodes within a cluster
- 7) Transmission of data to GW
 - a) After data aggregation, CHs transmits the aggregated data to the GW node
- 8) Data aggregation at GW
- 9) GW nodes after receiving the data from all the CHs perform task of data aggregation to reduce any data redundancy and have minimum data transmission
- 10) Transmission of aggregated data to BS.
 - a) Final aggregated data is transmitted to BS, after the data aggregation at GW node.

Algorithm 1 Data aggregation strategy

the number of transmissions. Algorithm 1 provides a general description of the proposed scheme.

3.1 Network model

For developing the network model of the proposed scheme, the following assumptions are made:

- Nodes are stationary after deployment and distributed randomly in a square region.
- Sink is stationary and located in a region near the sensing field.
- Gateway node is located in the center of the region.

Our proposed protocol consists of two phases, (1) setup phase and (2) steady state. Cluster formation and CH selection procedures are part of the setup phase, whereas the steady state phase deals with data transmission from nodes to CH, CH to GW node, and from GW to BS.

3.2 Setup phase

In this phase, CHs are selected and clusters are formed. The operation of the proposed scheme is divided into rounds, which serve as time units. For the 1st round, the CHs are selected on a random basis, and for subsequent rounds, CH selection is done by comparing the nodes' energy with a

threshold value. The threshold computation is done using Eq. (1):

$$Th \leq \frac{p}{1 - p \times r \text{mod}(\frac{1}{p})} \quad (1)$$

A node is chosen as a CH if its energy is greater than or equal to the threshold value; otherwise, it functions as a sensor node and collects data. Once the CH is chosen, it broadcasts its status to other nodes in the cluster through an advertisement message and waits for join requests. Join request messages are sent to the CH by nodes with higher RSS from the CH, and the CH subsequently associates with the nodes for the current round.

A specific percentage of nodes to be CHs is taken as input. Random selection of a node as a CH only occurs in the

first round. In subsequent rounds, the node's energy is compared to the threshold value. If the value is greater than or equal to the node's energy level, the node is selected as the CH within each cluster, and the other member nodes associate with it. Otherwise, the nodes function as normal nodes and perform sensing. Therefore, for a node to be selected as a CH, its energy level should be equal to or greater than the threshold value. At the end of this phase, CHs are selected, and clusters are formed. Figure 3 provides a summary of the setup phase.

3.3 Steady state phase

Data transmission takes place during this phase, with nodes sending the data they have sensed to the cluster head using Time Division Multiple Access (TDMA). TDMA scheduling

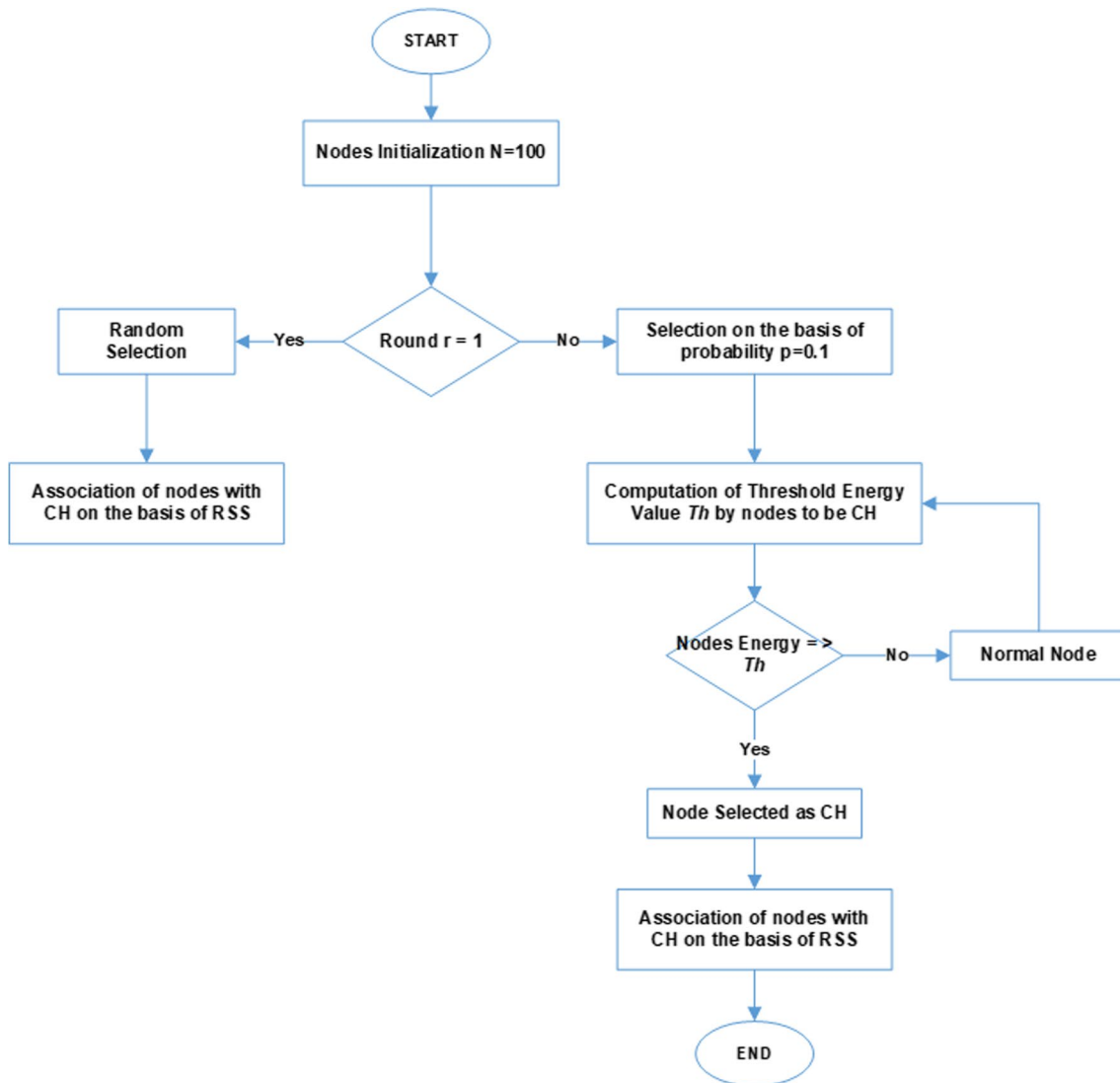


Fig. 3 Setup phase

is necessary to avoid data collision or interference that may occur if all nodes in a cluster transmit data simultaneously. Hence, CHs allocate TDMA schedules for inter-cluster data transfer after their selection in the setup phase. Nodes detect events and transmit data to the CH based on the designated TDMA schedule. The CH receives data from each node, aggregates the data upon reception, and sends it to the gateway, which is located at the center of the sensing field.

After completing the data aggregation process, the CHs transmit the aggregated data to the GW. At the GW, the received data from the CHs undergoes further aggregation to minimize the possibility of receiving duplicate data from other CHs in the vicinity. This dual aggregation approach reduces the likelihood of duplicate data being sent to the base station. The steady-state phase is summarized in Fig. 4. In the network, which utilizes a shared transmission channel, if a node is currently serving as a CH, it creates a TDMA

schedule for its member nodes. These member nodes then use this schedule to transmit data to the CH.

After collecting data from the nodes, the data is aggregated at the first aggregation point, the CHs. From each CH within its respective cluster, the data is transmitted to the second aggregation point, the gateway node. At the GW, the received data from all the CHs is aggregated once again to further minimize data duplication. Subsequently, the data is transmitted from the GW to the destination point, the BS. This approach ensures only one long-range transmission of data from the GW to the BS.

The inclusion of the GW node in the proposed scheme optimizes data aggregation and reduces the number of long-range data transmissions to the BS. Instead of each CH transmitting data directly to the BS, there is a single long-distance transmission from the GW to the BS. As a result, the proposed scheme significantly reduces the energy consumed by each CH during data transmission. The main focus of the proposed scheme is to minimize long-range data transmission within the network, thereby enhancing stability, prolonging network lifetime, and improving energy consumption. In the next section, we present the evaluation of our proposed scheme.

4 Performance evaluation

In the simulation of the proposed scheme, a $100\text{ m} \times 100\text{ m}$ square region is considered, with 100 nodes randomly deployed within it. The gateway node is positioned at the center of the region ($50\text{ m} \times 50\text{ m}$), while the base station is located near the region ($120\text{ m} \times 120\text{ m}$). For this simulation, ideal conditions are assumed, with minimal channel interference, relaxed bandwidth limits, and no data loss. However, future simulations will include these parameters to provide a more realistic assessment.

- Figure 5a illustrates the positions of the nodes, gateway node, and sink.
- Figure 5b showcases the network model with all nodes alive.
- Figure 5c presents the network model with all nodes inactive.

4.1 Simulation parameters

The parameters utilized for the simulation are outlined in Table 1.

4.2 Radio energy dissipation model

The radio energy dissipation model consists of several main components, including the transmitter, amplifier, and

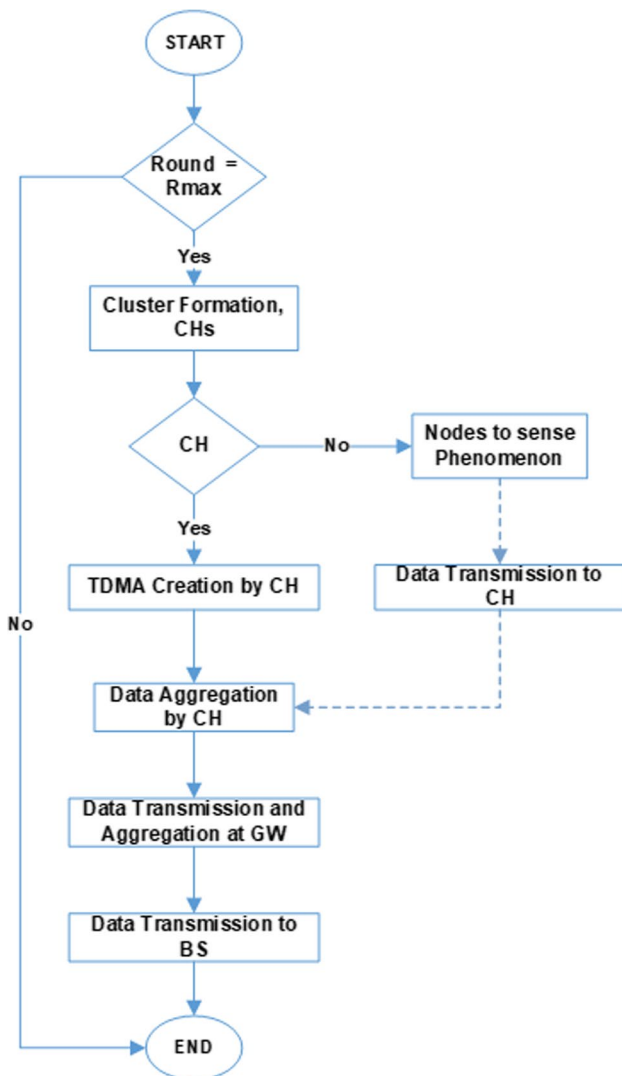


Fig. 4 Steady state phase

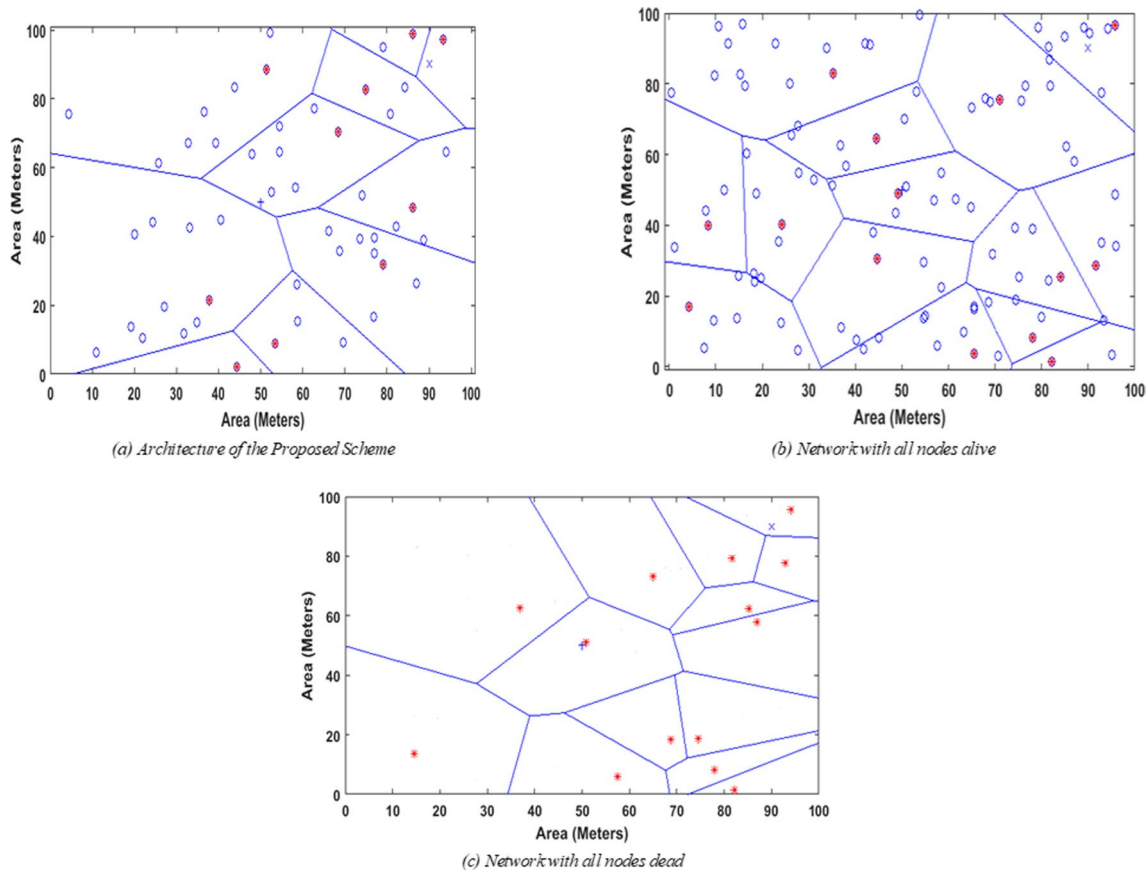


Fig. 5 Network architecture with alive and dead nodes

Table 1 Simulation parameters

Parameters	Value
Number of Nodes (n)	100 (300 for Fig. 10)
Probability to be a CH (P)	0.1
Nodes energy E_o	0.5 J
Transmitter’s energy E_{Tx}	50×10^{-10} J
Receiver’s energy E_{Rx}	50×10^{-10} J
Energy (free space) E_{fs}	10×10^{-11} J
Energy (Multipath fading) E_{mp}	0.0013×10^{-12} J
Energy (Data aggregation) EDA	5×10^{-9} J
Maximum rounds R_{max}	5000
Frequency	25 kHz

receiver. The transmitter generates and transmits signals, while the amplifier increases the power of the signal. On the other hand, the receiver receives the signals and converts them into a usable form. To operate, the transmitter requires energy for running the amplifier and radio electronics, while the receiver requires energy for powering its own radio electronics. In the proposed model, a radio energy dissipation model [21] is utilized, which quantifies the energy

needed to transmit an “L” bit message packet over a distance “d.” Additionally, it considers the energy consumed by the receiver electronics during the reception of the “L” bit message packet. The energy required to transmit the “L” bits message packet over a distance “d” is represented by Eq. (2):

$$d = \begin{cases} E_{Tx}(L, d)L \times E_{elec} + L \times E_{fs} \times d^2 & \text{when}(d < d_0) \\ L \times E_{elec} + L \times E_{mp} \times d^4 & \text{when}(d \geq d_0) \end{cases} \quad (2)$$

The energy dissipated to run transceiver circuitry is represented by E_{elec} , multi-path fading (E_{mp}) and free space (E_{fs}) are the parameters which represent the dissipation of energy per bit depending upon the distance d_o where the threshold value for the distance is d_o and its value is computed in Eq. (3):

$$d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}} \quad (3)$$

Two channels are used in the simulation: (1) free space channel, representing ideal propagation conditions with a clear line-of-sight path between the transmitter and receiver, and (2) multipath fading channel, accounting for scenarios

with multiple propagation paths. The selection of these channels depends on the distance between the transmitter and receiver.

For distances below the computed threshold value (d_0), the free space model (fs) is employed. When the distance exceeds the threshold, the multipath fading model is used. The energy required by the receiver to receive an “L” bit message packet, denoted as E_{Rx} , is defined by Eq. (4).

$$E_{Rx}(L) = L \times E_{elec} \tag{4}$$

4.3 Results and discussion

We compared our proposed protocol with the benchmark clustering protocol, LEACH, through simulations conducted in MATLAB. The results obtained are as follows: Fig. 6 illustrates the energy dissipated by the nodes in both protocols. In LEACH, the nodes dissipate their energy at nearly 500 rounds, whereas in our proposed scheme, nodes dissipate energy at nearly 1500 rounds. It shows that in the proposed scheme energy is dissipated after 1000 rounds compared to LEACH due to minimum long-range transmission. The aggregation of data at CH and gateway node minimizes the transmission of redundant data. Thus, the energy dissipation of nodes in the proposed scheme is better than LEACH.

Figure 7 shows the data packets sent to the aggregation point, CH, by nodes in both protocols. It depicts that in the proposed scheme, more packets are sent to the CH compared to LEACH. In the proposed protocol, nodes do not send data directly to the BS; instead, they forward data packets to the CH for aggregation before transmission to the BS. In LEACH, the packets are sent to the CH until 1000 rounds, whereas in the proposed scheme, the nodes send packets to the CH until 3000 rounds, which demonstrates a better result.

The number of packets sent to the BS is shown in Fig. 8, where a comparison is made between the two protocols. In

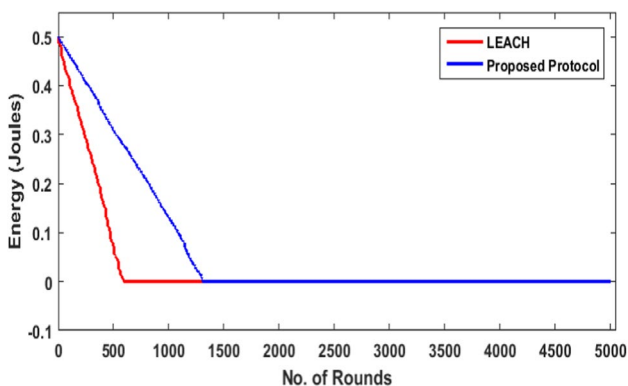


Fig. 6 Energy dissipation of nodes

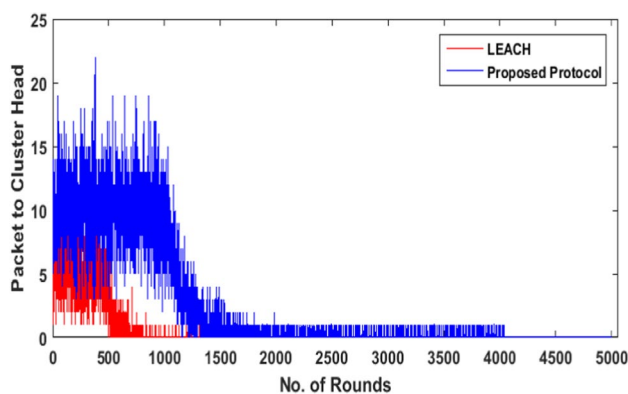


Fig. 7 Transmission of Data Packets from nodes to CH

LEACH, data packets are sent from CH to BS until 1500 rounds, while in the proposed scheme, data packets are sent to BS until 3000 rounds. The proposed scheme efficiently utilizes the network’s energy by incorporating a gateway node for data aggregation, resulting in a single long-range transmission to the BS. This indicates that the proposed scheme allows for a longer duration of data transmission to the BS compared to LEACH.

Figure 9 illustrates the stability period and network lifetime. In LEACH, the first node dies at 500 rounds, and all nodes are dead by 1500 rounds. In our scheme, the network becomes unstable after 790 rounds when the first node dies. The network lifetime of our scheme is also superior to the LEACH protocol, as the last node’s death occurs at 3000 rounds. Through simulation analysis and comparison of various parameters, it is evident that the proposed protocol outperforms LEACH in terms of data packets sent to the BS, energy dissipation of nodes, stability period, and overall network lifetime. The results indicate that employing minimum transmission to the sink is a more effective approach for homogeneous networks. The

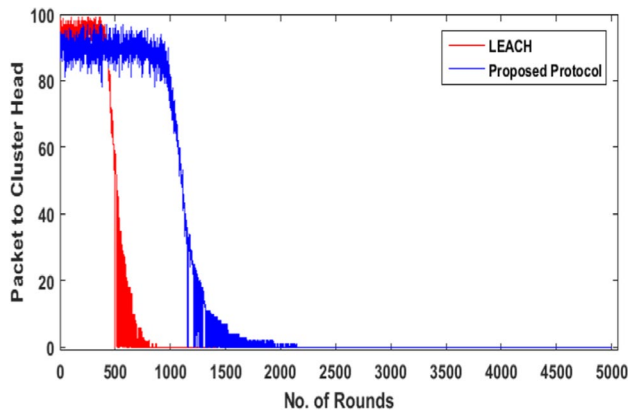


Fig. 8 Transmission of data packets to BS

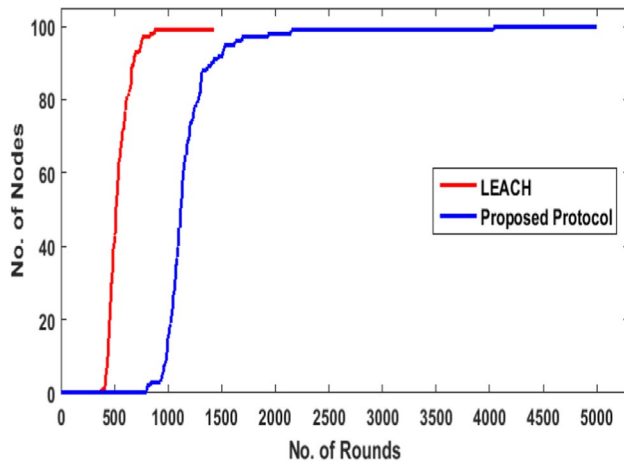


Fig. 9 Stability and network lifetime

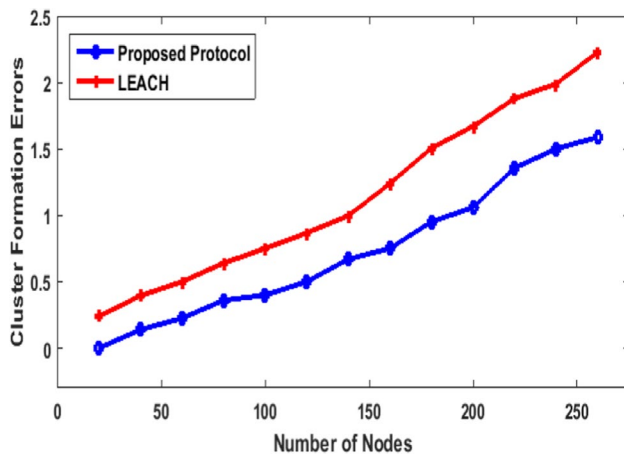


Fig. 10 Cluster formation errors

utilization of data aggregation at two points significantly reduces the number of long-distance transmissions to the BS, resulting in more efficient energy usage by consolidating multiple transmissions into a single long transmission after aggregation at the gateway node.

Figure 10 displays the preliminary results regarding cluster formation errors in our proposed scheme. The graph shows an almost linear function, indicating that as the number of nodes in the network increases, the stability of our algorithm remains consistent. This observation strongly supports the efficiency of our approach. It suggests that employing a smaller number of cluster heads can also lead to energy reduction. However, further investigation is required to accurately measure the computational overhead and end-to-end delay that may arise as a result of dual aggregation in our proposed algorithm.

5 Conclusive remarks and perspectives

This work introduces a novel approach to efficiently utilize the energy in wireless sensor networks through data aggregation. By aggregating redundant data at aggregator nodes, the number of transmissions is reduced, resulting in lower energy consumption. The proposed approach includes the introduction of a gateway node that performs a second aggregation of data received from cluster heads and transmits it to the sink as a single long-range transmission. This approach significantly improves network lifetime by conserving energy.

Comparative analysis with the LEACH protocol demonstrates that our proposed approach outperforms LEACH in terms of stability, energy consumption, and network lifetime. However, it is important to note that this work specifically focuses on homogeneous networks. In future research, we plan to extend the concept of dual data aggregation to heterogeneous networks and consider additional performance parameters such as packet loss, delay, and overhead.

Declarations

Conflicts of interest The authors declare no competing interests.

References

1. Juneja D, Sharma A, Sharma AK (2011) Wireless sensor network security research and challenges: a backdrop. In: Mantri A, Nandi S, Kumar G, Kumar S (eds) High Performance Architecture and Grid Computing. HPAGC 2011. Communications in Computer and Information Science, vol 169. Springer, Berlin, Heidelberg
2. Kumar A et al (2014) Data aggregation in wireless sensor networks Proc. of Intl J. of Science & Research, vol. 3(3), pp. 249–251
3. Wang C, Jiang C, Liu Y, Li X-Y, Tang S, Ma H (2011) Aggregation capacity of wireless sensor networks: extended network case. In: 2011 Proceedings IEEE INFOCOM, Shanghai, pp 1701–1709. <https://doi.org/10.1109/INFCOM.2011.5934966>
4. Rahman H, Ahmed N, Hussain I (2018) A QoS-aware hybrid data aggregation scheme for Internet of Things. Annals of Telecom 73(7–8):475–486
5. Xiangning F, Yulin S (2007) Improvement on LEACH protocol of wireless sensor network. In: 2007 International Conference on Sensor Technologies and Applications (SENSORCOMM 2007), Valencia, pp 260–264. <https://doi.org/10.1109/SENSORCOMM.2007.4394931>
6. Yuan F et al (2014) Data density correlation degree clustering method for data aggregation in wsn. IEEE Sens J 14(4):1089–1098
7. Zhou F et al (2016) Maximizing lifetime of data-gathering trees with different aggregation modes in wsns. IEEE Sens J 16(22):8167–8177

8. Prathima EG, Laxmikant H, Naveen SA, Venugopal KR, Iyengar SS, Patnaik LM (2017) DAMS: data aggregation using mobile sink in wireless sensor networks. In: Proceedings of the 5th International Conference on Communications and Broadband Networking (ICCBN '17). Association for Computing Machinery, New York, pp 6–11. <https://doi.org/10.1145/3057109.3057118>
9. Villas LA, Guidoni DL, Araújo RB, Boukerche A, Loureiro AAF (2010) A scalable and dynamic data aggregation aware routing protocol for wireless sensor networks. In: Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems (MSWIM'10). Association for Computing Machinery, New York, pp 110–117. <https://doi.org/10.1145/1868521.1868540>
10. Nguyen NT et al (2016) On maximizing the lifetime for data aggregation in wireless sensor networks using virtual data aggregation trees. *Comp Nets* 105:99–110
11. Saginbekov S, Jhumka A (2017) Many-to-many data aggregation scheduling in wireless sensor networks with two sinks. *Comp Nets* 123:184–199
12. Sasirekha S, Swamynathan S (2017) Cluster-chain mobile agent routing algorithm for efficient data aggregation in wireless sensor network. *J of Comm & Net* 19:392–401
13. Min W, Ruixiang C, Shunbin H (2016) A secure data aggregation approach in hierarchical wireless sensor networks. In: Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication (IMCOM'16). Association for Computing Machinery, New York, pp 1–7. <https://doi.org/10.1145/2857546.2857637>
14. Hua P et al (2018) Energy-efficient adaptive slice-based secure data aggregation scheme in wsn. *Procedia Computer Science* 129:188–193
15. Xiang L et al (2013) Compressed data aggregation: energy-efficient and high-fidelity data collection. *IEEE/ACM Tran on Net* 21(6):1722–1735
16. Xu X et al (2015) Hierarchical data aggregation using compressive sensing (hdacs) in wsns, *ACM Trans. Sen. Net.*, vol. 11, pp. 45:1–45:25
17. Islam M, Kim J (2010) On the cooperative mimo communication for energy-efficient cluster-to-cluster transmission at wireless sensor network, *annals of telecom*, vol. 65, pp. 325–340
18. Dias G et al (2017) The impact of dual prediction schemes on the reduction of the number of transmissions in sensor networks. *Comput Commun* 112:58–72
19. Ahmed R et al (2016) Cluster-based data aggregation for pest identification in coffee plantations using wireless sensor networks. *Comput Electr Eng* 56:591–607
20. Wang Z et al (2017) Cluster-based maximum consensus time synchronization for industrial wireless sensor networks. *Sensors* 17(1):141
21. Heinzelman WR, Chandrakasan A, Balakrishnan H (200) Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, vol 2, pp 10. <https://doi.org/10.1109/HICSS.2000.926982>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.