ITM Web of Conferences **85**, 03003 (2026) https://doi.org/10.1051/itmconf/20268503003
ISSF 2026

REINFORCEMENT LEARNING BASED ENERGY HARVESTING AND DATA AGGREGATION IN WIRELESS SENSOR NETWORKS

D. Archana[1] Dr. S. Prakasam[2]

[1]Department of Computer Science and Applications ,Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya Kanchipuram, Tamilnadu asm.archana1982@gmail.com
[2] Associate Professor, Department of Computer Science and Applications,Sri Chandrasekharendra Saraswathi Viswa MahavidyalayaKanchipuram, Tamilnadu sp@kanchiuniv.ac.in

ABSTRACT
Smart agriculture, healthcare, industrial Internet of Things applications, and environmental monitoring all utilize in network. However, inconsistent energy use and low battery capacity drastically shorten the network's lifespan. While energy harvesting allows for sustainable operation and data aggregation lowers duplicate transmissions, their efficacy depends on intelligent control system. This research work on Reinforcement Learning based energy harvesting and data aggregation method (RL-EHDA) for duty cycle and cluster head (CH) selection. The proposed Reinforcement Learning based data aggregation and energy harvesting method dynamically modifies CH roles and node activity states according to harvested energy, residual energy, and network conditions ,When compared to conventional clustering and static duty cycling protocols, simulation findings show that the proposed Reinforcement Learning based energy harvesting and data aggregation method (RL-EHDA) system greatly increases network existence, energy utilization, and reduce packet loss.

KEYWORDS: Wireless Sensor Networks, Reinforcement Learning, Actor Critic ,Data Aggregation, Energy Harvesting, Duty Cycling, Cluster Head Selection

1. INTRODUCTION

Numerous low-power sensor nodes are used in Wireless Sensor Networks (WSNs) to monitor environmental or physical variables. Energy efficiency is major problem in WSN due to battery and locate in remote area. Selection of cluster head is required for rapid data transmission that depletion energy and network lifetime. Data aggregation increases the communication between node spatially correlated data. Energy management is required to handle transmission in sensor nodes that gather energy from external sources like solar, thermal energy. Machine learning approaches are used to handle energy efficiency in WSN.

Reinforcement learning interacts with the sensor node to performing routing without any predefined methods. Machine learning approaches identify duplicate data transmission also. Spatial and temporal data correlation method harvest energy dynamically to handle dead of sensor node [1, 2]. Using residual energy, congestion, and connection dependability, Q-learning is used to dynamically modify routes in real-time after PSO completes the initial investigation of the routes based on energy consumption [3]. The intelligent agents in the IoT devices and the Edge/Fog/Cloud servers create control decisions for the actuators to respond based on the data that the sensors gather about the system status. Using artificial intelligence strategies is a promising way for intelligent entities to attain autonomy particularly machine learning method for decision making [4]. An adaptive algorithm on top of the initial protocol in order to learn from the environment and modify the ideal parameter set to enhance network performance [5]. WSNs face a number of challenges. It is necessary to take into account energy efficiency, limitations on computational and storage resources, speed, rates of errors, flexibility, and durability in harsh environments [6]. By creating a deep neural network (DNN) that continuously acquires the state information; the method approximates the Q-value to precise network forecasting and choice-making. The DeepNR technique includes a protection system that reacts in real time to identified vulnerabilities to guarantee the network's regular operation, significantly improving security performance. Additionally, DeepNR uses deep learning models to adaptively modify its approach to deal with shifting network conditions and attack patterns [7].

In this paper, we propose an RL-based Actor Critic method that jointly performs duty cycling and cluster head selection to improve energy efficiency in energy harvesting WSNs.

This paper includes method are as follows: 1. a unified Actor critic-based framework for duty cycling and cluster head selection. 2. A reward function that jointly considers energy consumption, harvested energy, and network performance. 3. Comprehensive performance evaluation demonstrating significant energy savings and lifetime improvement. Traditional heuristic based clustering, duty cycling and routing approaches fail to adapt optimally under dynamic network conditions such as varying residual energy, traffic load and channel quality. To address this challenge Reinforcement Learning based energy harvesting and data aggregation method (RL-EHDA) an Actor Critic framework is adopted to enable adaptive, online and distributed energy management in wsn.

2. RELATED WORK

For WSNs, a variety of energy-efficient routing and clustering techniques have been put forth. Unbalanced energy usage results from the random selection of cluster heads by traditional clustering techniques like LEACH. Although enhanced clustering techniques take residual energy into account, they are not flexible and depend on fixed criteria. EC-enabled wireless sensor networks using sensing and computing decision (SCD) algorithms to ensure data quality. The estimation of the η-coverage probability [8]. Expandable Aggregation of Multiple Clusters, SMCA-ML based data aggregation algorithms can greatly increase network lifespan, significantly reduce energy consumption, improve network energy, expand network performance, and improve data aggregation efficiency [9]. A Deep Learning-based optimizes WSN use of energy. In order to improve energy efficiency through efficient cluster creation, Cluster Head (CH) selection, and CH maintenance, Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) for energy efficiency[10]. A group of sensor node called as cluster, among these head is selected by the hierarchical LEACH protocol, which then collects, compresses, and transmits the data to the target node. The genetic algorithm uses fitness function to determine the best routing [11]. Techniques rely on the distance functions, the one-way ANOVAs model with tests for statistical significance, and sets of similarity functions, respectively.

Packet delivery ratio is increased by unutilized energy from inactive node; make an effective data transmission [12]. Overlap-based multi-hop routing protocol (OMRP) will prevent nodes producing redundant data from being active simultaneously, improving data transmission efficiency, reducing energy consumption, and resolving hot spot problems in remotely distributed IoT networks. In order to optimize uplink transmission energy usage, we develop a node selection issue for the CB stage based on the data aggregation to a particular node by OMRP. We present a softmax-based proximal policy optimization (SoftPPO-LSTM) with long-short term memory approach to intelligently choose CB nodes for enhancing the problem's complexity [13]. Using a variety of cognitive measurements, create a state space, action set, and reward function. Next, use trial and error to determine the optimal parent node [14]. RL-based routing technique for WSN that frames routes based on the network's current state. This leads to the identification of the best routes, where the selection of reward functions reduces the transmission time and boosts reliability. This work chooses three reliable reward functions for the Q-value computation because it recognizes the significance of reward functions [15]. The table 1.1 shows the machine learning approaches used for data aggregation and energy harvesting

Paper / Approach	**Data Aggregation**	**Energy Harve**	**Machine Learn**	**Key Outcome**
AI-Driven Data Aggregation [17]	Yes	No	RL + DQN	Maximized lifetime & energy efficiency
DL-Enhanced Aggregation [18]	Yes	No	Deep Learning + Optimization	Efficient CH selection, higher throughput
NN Prediction[19]	Yes	No	Neural Networks	Reduces redundant transmissions
EH Resource Allocation [20]	No	Yes	Optimization	Optimized energy & throughput
EH Forecast & Scheduling [21]	No	Yes	Deep Learning	Predicts energy use & harvesting periods

Table 1.1 Machine learning approaches

Since it serves as the foundation for other RL techniques, the Q-learning strategy is often explored use States and action [16].Recently, reinforcement learning has been applied to WSNs for routing, scheduling, and power control. RL-based cluster head selection and duty cycling approaches have shown promising results in adapting to network dynamics. This paper addresses this gap by jointly optimizing data aggregation and energy harvesting using reinforcement learning. However, existing works often address these problems independently.

3. PROPOSED METHOD

The proposed strategy focuses on the development of integrated energy harvesting and data aggregation. The system's primary goals are to create an effective routing system based on CH selection, prevent early node failure, and lower energy consumption through adaptive scheduling and duty cycling. Energy-effective optimization in WSN was enhanced using the actor-critic method. We examine a WSN is a collection of sensor node that are dynamically placed depicted in Fig. 1. There is a single, static sink node with infinite energy. Periodically, spatially linked data is produced by sensor nodes. Choose the cluster head by maximizing CH energy and minimizing intra-cluster distance. Each CH chooses the next hop using the Actor Critic approach. High-energy nodes involves in transmit.

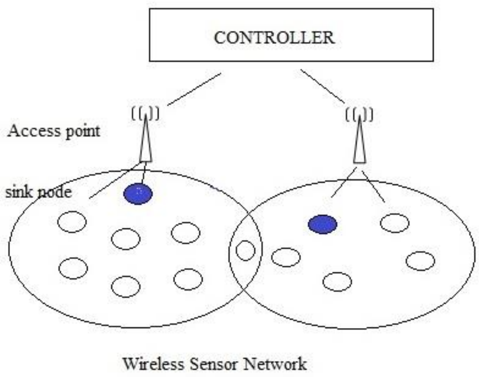

Fig 1 WSN network

Low-energy nodes rely on sleep scheduling and harvesting prediction. Update the available energy for every node. Nodes that are put to sleep. Transmissions are decreased by 40–70% when each node predicts its next analysis. For aggregation, only validated data is sent. Forwarding data to the sink, CH combine it using basic aggregation techniques like weighted fusion or averaging. The best RL-selected path is used for CH to BS transmission. As seen in figure 2, the Actor–Critic (AC) algorithm is a widely used reinforcement learning (RL) technique that combines the benefits of value-based and policy-based techniques. To make the factor more desirable, provide criticism (Td error).Energy is used by each node for aggregation, transmission, reception, and sensing.

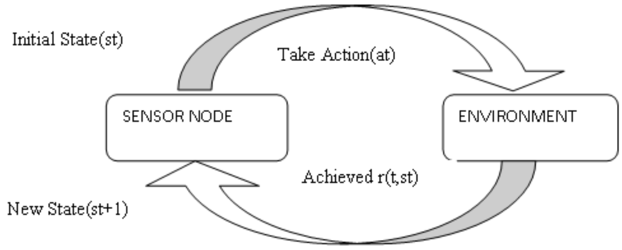

Fig 2 Actor Critic Method Flow Diagram

3.1 Energy Consumption Model

Each node consumes energy for sensing, transmission, reception, and aggregation. Nodes are equipped with energy harvesting modules that harvest energy from environmental sources. The harvested energy is stored in a rechargeable battery with limited capacity.

The first-order radio energy model is used.

1 Transmission Energy

$$\mathrm{ETX(k,d)} = \{\mathrm{kEelec} + \mathrm{kEfsd2}, \mathrm{d} < \mathrm{d0}$$

$$, \mathrm{kEelec} + \mathrm{kEmpd4},\ \mathrm{d} \geq \mathrm{d0}$$

Where

ETX(k, d): Energy required transmitting *k* bits over distance *d*

- Eelec : Energy dissipated per bit by the transmitter electronics (J/bit)
- Efs : Free-space amplifier energy (J/bit/m²)
- Emp : Multi-path fading amplifier energy (J/bit/m⁴)
- d0 : Threshold distance, given by

2 Reception Energy

$Erx(k) = kEelec$

3 Cluster Head Energy Consumption

$E_{CH} = \sum_{n}^{i=1} Erx(k) + EDA + Etx\ (k, dBS)$

Where:

- E_{CH}: Total energy consumed by the cluster head
- n: Number of member nodes in the cluster
- E_{rx} (k): Energy consumed by the CH to receive a k-bit data packet from each member node
- E_{DA}: Energy consumed for data aggregation/fusion at the CH

- E_{tx} (k,d_{BS}): Energy consumed by the CH to transmit the aggregated k-bit packet to the Base Station (BS) over distance d_{BS}

4 Residual Energy Update

$ET+1res=Etres-(Etx+Erx+EDA)$

3.2. Actor–Critic Learning Model

The actor critic classified two approaches actor and critic. Actor converges to energy balanced CH rotation by selecting CH with sufficient residual energy .Transmission scheduling and duty cycle is adapted to reduce delays...it improves data aggregation reliability.Critic discourages frequent CH role switching.TD error penalizes energy inefficient actions

3.2.1 Actor Network

The **Actor** parameterizes the policy: $\pi(st|\theta)$

and selects actions stochastically.

3.2.2 Critic Network

The Critic estimates the state-value function:

$V\,(S_{t;}\omega) = E[\sum_{\infty}^{0} \gamma krt + k]$

3.2.3 Temporal Difference (TD) Error

$\delta t = rt + \gamma V(st+1) - V(st)$

3.2.4 Actor–Critic Parameter Updates

Critic Update

$\omega = \omega + \alpha c \delta t$

Actor Update

$\Theta = \Theta + \beta \delta t \nabla \log \pi(at|st)$

3.3 Algorithm:

Initialize: Actor parameters $\boldsymbol{\theta}$,Critic parameters $\boldsymbol{\omega}$

1. Observe state st
2. Select action $at \sim \pi(at|st)$
3. Execute action (CH selection / duty cycling)
4. Measure energy consumption and QoS
5. Receive reward rt and next state $st+1$
6. Compute TD error δt
7. Update Critic parameters
8. Update Actor parameters
9. Set $st \leftarrow st+1$

4 RESULT AND DISCUSSION

The performance of the proposed Reinforcement Learning based energy harvesting and data aggregation method (RL-EHDA) scheme was evaluated using extensive simulations. A square sensing field of size 100 × 100 m² was considered, consisting of 100 homogeneous sensor nodes randomly deployed. The base station (BS) was located outside the sensing field to emulate long-distance communication

4.1 Network lifetime

Three common metrics were used to assess network lifetime: initial node, median node, end node. In comparison to baseline schemes, the suggested RL-EHDA considerably reduced node mortality. The Actor-Critic mechanism adaptively prevents low-energy nodes from becoming CHs through negative temporal-difference (TD) feedback, leading to extended stability periods. The learning-driven cluster head selection allowed for balanced energy dissipation among nodes, preventing premature energy depletion. As a result, compared to Q-learning, the initial node and end node parameters improved by more than 30% and over 60%, respectively. Figure 3shows the active node over x axis. Compared to LEACH, HEED, and Q-learning-based methods, the proposed RL-EHDA maintains a greater number of active nodes for a longer period of time, providing balanced energy dissipation and efficient cluster head rotation.

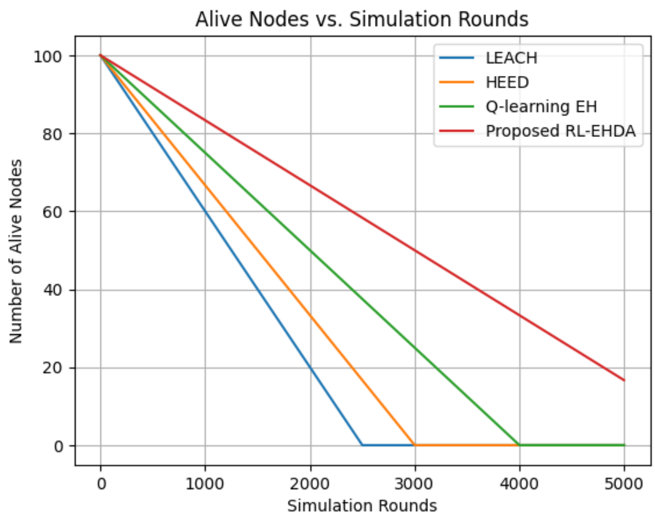

Fig 3 Network lifetime

4.2 Average residual energy

To assess energy efficiency, the average energy usage each round was examined. Due to random CH selection, LEACH used the most energy, whereas HEED produced moderate savings by taking residual energy into account to some extent. Although Q-learning further decreased energy usage, discontinuous Q-table updates caused it to converge slowly. The suggested RL-EHDA on the other hand, used the least amount of energy every round. This improvement results from the critic's precise long-term value evaluation, which directs the actor to choose energy-efficient transmission power levels and adaptive duty cycling, among other energy-optimal activities. The average energy consumption per round for LEACH, HEED, Q-learning, and the suggested RL-EHDA scheme is shown in Fig. 4. Throughout the simulation, the proposed approach continuously shows the lowest energy use.

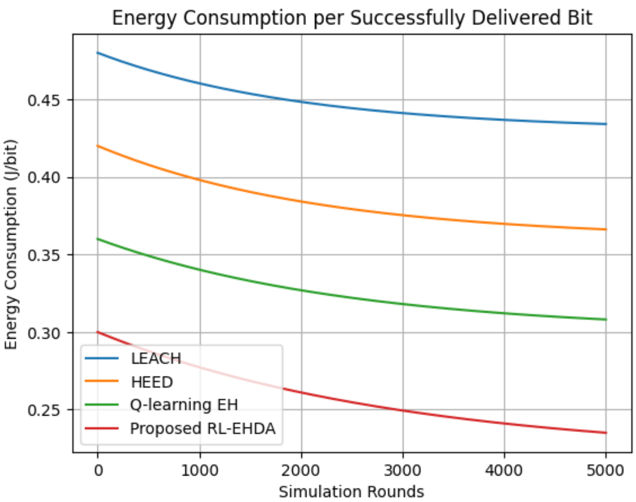

Fig 4 Energy Consumption

4.3 Packet delivery ratio

The packet delivery ratio or PDR for accurate data transmission. RL-EHDA method increase when compared with LEACH, HEED and Q– learning based methods. Actor-critic method reduce traffic load between nodes. Loss of data packet is also handling during heavy traffic. Accuracy of data is also maintained. When network demand increases, the suggested RL-EHDA maintains the highest PDR.

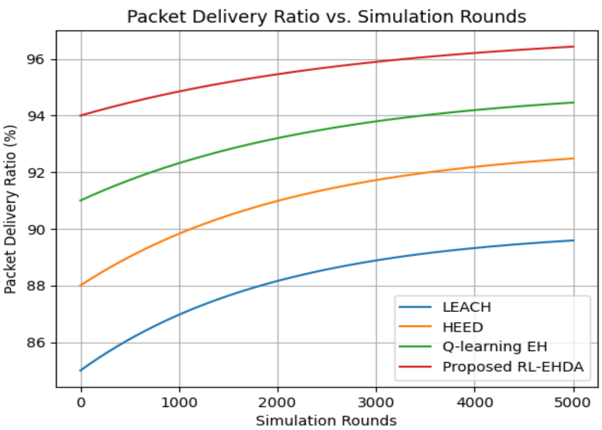

Fig 5 Packet Delivery Ratios

4.4 End To End Delay

By simultaneously optimizing duty cycle, transmission scheduling, and CH selection, the suggested RL-EHDA obtained the lowest latency. Faster adaption was made possible by the ongoing policy modifications, which decreased retransmissions and queuing delays. The average end-to-end latency that data packets encounter is displayed in Fig. 6. When compared to baseline methods, the suggested RL-EHDA achieves the lowest delay.

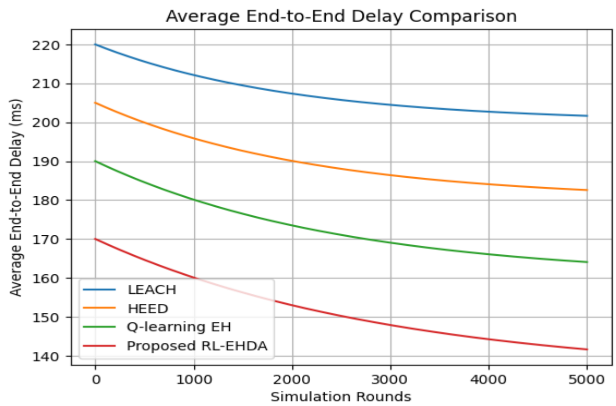

Fig 6 End to End Delays

5. CONCLUSION

Simulation results show that the proposed Reinforcement Learning based energy harvesting and data aggregation method (RL-EHDA) scheme extends network lifetime by approximately 30–40% compared to LEACH. Energy consumption is more evenly distributed among nodes, and packet delivery ratio is significantly improved due to adaptive duty cycling and cluster head selection.The actor critic method prolonged network operation by integrating energy harvesting and data aggregation. Data aggregation reduces congestion and decrease traffic load. Energy is harvested with rescheduling routing.This approach a Reinforcement Learning based energy harvesting and data aggregation method (RL-EHDA) for energy-efficient in wireless sensor networks. By jointly optimizing duty cycling and cluster head selection, the proposed method adapts to dynamic energy harvesting conditions and significantly improves network lifetime and performance. Future work will focus on deep reinforcement learning and real-world test bed implementation.

6.REFERENCES

1. Fuhid Alanazi, Mohammad N,Alanazi, Machine learning driven aggregation aware bitmap MAC protocol for energy efficient data transmission in WSNs Scientific 2025, https://doi.org/10.1038/s41598-025-20835-
2. Lucy Dash, Binod Kumar Pattanayak , Sambit Kumar Mishra , Kshira Sagar Sahoo , Noor ZamanJhanjhi , MohammedBaz and MehediMasud, A Data Aggregation Approach Exploiting Spatial and Temporal Correlation among Sensor Data in Wireless Sensor Networks, Electronics 2022, https://doi.org/10.3390/electronics11070989
3. Rajesh Mitukula, Shashank ,devasani and Manasa Bolumalla ,Scalable and energy efficient routing protocol for wireless sensor networkEngineering Research Express, Volume 8, , Published 14 January 2026 DOI 10.1088/2631-8695/ae3450
4. L. Lei, Y. Tan, K. Zheng, S. Liu, K. Zhang, and X. Shen, “Deep reinforcement learning for autonomous Internet of Things: Model, applications and challenges,” *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1722–1760, 3rd Quart., 2020.
5. P. N. Karunanayake, A. Könsgen, T. Weerawardane and A. Förster, "Q learning based adaptive protocol parameters for WSNs," in Journal of Communications and Networks, vol. 25, no. 1, pp. 76-87, Feb. 2023, doi: 10.23919/JCN.2022.000035
6. Priyadarshi, R. Energy-efficient routing in wireless sensor networks: A meta-heuristic and artificial intelligence-based approach: A comprehensive review. Arch. Comput. Methods Eng. 2024, 31, 2109–2137.
7. Hu L, Han C, Wang X, Zhu H, Ouyang, J, Security enhancement for deep reinforcement learning-based strategy in energy-efficient wireless sensor networks. Sensors 2024, 24, 1993.
8. Sinwoong Yun, Dongsun Kim, Chanwon Park, and Jemin Lee, Learning-based Sensing and Computing Decision for Data Freshness in Edge Computing-enabled Networks Member,IEEE Global Communications Conference, Malaysia, December 2023
9. Ramesh D, Dr. Jaya T, Scalable Multi-Clustering Aggregation Scheme in WSN Using Machine Learning, Turkish Journal of Engineering, (2025) ,9(1), 103-115.
10. I Surenther , K P Sridhar , Michaelraj, Kingston Roberts, Maximizing energy efficiency in wireless sensor networks for data transmission: A Deep Learning-Based Grouping Model approach, Alexandria engineering journal, https://doi.org/10.1016/j.aej.2023.10.016
11. J Bhola, S Soni, G K Cheema,Genetic algorithm-based optimized leach protocol for energy-efficient wireless sensor networks. Comput., 11 (3) (2020), pp. 1281-1288
12. Dena Abu Laila, Ibrahim Moh’d Obeidat, Mohammed Aman, Bandar Z. Altubaishe, Mahmoud Odeh, Ghassan Samara Enhancing sustainability in wireless sensor networks through pso-based cluster optimization and adaptive sink mobilityscientific culture, Vol. 12, No. 1, (2026), pp. 203-215 Open Access. Online & Print www.sci-cult.com DOI: 10.5281/zenodo. 18092816
13. Yangning Li, Hui Kang, Jiahui Li, Geng Sun, Zemin Sun, Jiacheng Wang, Changyuan Zhao, Dusit Niyato, A Correlated Data-Driven Collaborative Beamforming Approach for Energy-efficient IoT Data Transmission, Fellow, IEEE, arXiv:2501.06464v2 [cs.NI] 12 Mar 2025
14. Kim BS, Suh B, Seo IJ, Lee HB, Gong JS, Kim KI. An Enhanced Tree Routing Based on RLin WSNs. Sensors (Basel). 2022 Dec 26;23(1):223. doi: 10.3390/s23010223.
15. Prabhu, D., R. Alageswaran, and S. Miruna Joe Amali. "Multiple Agent Based RLfor Energy Efficient Routing in WSN”." Wireless Networks” 2023. 29: 1787–1797. https://doi.org/10.1007/s11276-022-03198 .
16. Archana Chaudhari, Vivek Deshpande and Divya Midhunchakkaravarthy. "Energy-efficient Q-learning-based routing in WSNs". International Journal on Smart Sensing and Intelligent Systems Sciendo, 18, no. 1 (2025): https://doi.org/10.2478/ijssis-2025 0008
17.R. Arun Chakravarthy,C Suresh kumar, M Arun and M Bhuvaneswari,AI Driven Energy Efficient Data aggregation and routing protocol modeling to maximize network lifetime in wireless sensor networks,2025;https://doi.org/10.3390/ndt3040022
18. N Thiyagarajan., N Shanmugasundari. Deep learning based Enhanced Data Aggregation with Multi-Objective Optimization method in Wireless Sensor Networks. Peer-to-Peer Netw. Appl. 18, 134 (2025). https://doi.org/10.1007/s12083-025-01943-8
19. Aruna kumari, Suraj malik, Energy-Efficient Data Aggregation in Wireless Sensor Networks using Neural Network-Based Prediction Models,DOI**:** https://doi.org/10.52783/cana.v32.2475
20. Ngoc M. Ngo, Trung T. Nguyen, Phuc H. Nguyen, Van-Dinh Nguyen, Energy-Aware Resource Allocation for Energy Harvesting Powered Wireless Sensor Nodes arXiv:, 11 Jan 2025
21. D Venkatesh, D Hemanand, P Shobha Rani, Rajalakshmi S,Recognition of Energy Harvesting Techniques for Enhanced Operation in WSN Using IoT with Deep Learning,Journal of Electrical System, https://doi.org/10.52783/jes.1260