



Energy-Efficient Node Placement and Routing in Dynamic WSNs Using Adaptive Nest Competition and Deep Reinforcement Learning

Monica Gunjal^{1,2*} , Pramodkumar H. Kulkarni² 

¹ Department of Electronics & Telecommunication Engineering, Army Institute of Technology, Pune 41101, India

² Department of Electronics & Telecommunication Engineering, Dr. D. Y. Patil Institute of Technology, Pune 411018, India

Corresponding Author Email: monicagunjal@gmail.com

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.120909>

ABSTRACT

Received: 16 June 2025

Revised: 13 August 2025

Accepted: 18 August 2025

Available online: 30 September 2025

Keywords:

wireless sensor networks, adaptive nest competition algorithm, fuzzy C-means, deep Q-learning, energy efficiency, clustering, routing, node placement

This study aims to improve energy efficiency and routing performance in dynamic Wireless Sensor Networks (WSNs), where node mobility and limited power are major challenges. The objective was to enhance energy efficiency and extend the network lifetime of the dynamic WSN for the implementation of a time-sensitive IoT-based system. We proposed an integrated methodology compounding three key techniques. The Adaptive Nest Competition Algorithm (ANCA) is used for optimal placement of sensor nodes to ensure wide coverage and strong connectivity. Fuzzy C-Means (FCM) clustering groups of nearby nodes to minimize communication within clusters. A Deep Q-Learning (DQL) algorithm that learns and adapts routing decisions based on changing network conditions to ensure efficient data transmission. The proposed framework outperforms traditional methods such as Artificial Bee Colony (ABC), Particle Swarm Optimization (PSO), and Grey Wolf Optimizer (GWO) in simulations. It achieves a Network Lifetime improved by 20–25%, Average Energy Consumption reduced by 15–25%, Packet Delivery Ratio (PDR) increased by 10–18%, End-to-End Delay decreased by 20–30%, and Routing Overhead reduced by 15–28%. This hybrid ANCA-FCM-DQL model provides a robust and adaptive solution for energy-aware node deployment and intelligent routing in dynamic WSNs, making it suitable for real-time, mobile, and energy-constrained applications.

1. INTRODUCTION

Wireless Sensor Networks (WSNs) play a critical role in numerous real-time applications such as smart agriculture, environmental monitoring, and disaster management. Efficient performance of WSNs relies heavily on two interdependent factors: optimal node placement and energy-efficient routing. Node placement influences coverage and connectivity, while routing determines the energy consumption and network lifetime. Achieving a balance between these factors in a dynamic environment, where node mobility and energy constraints are predominant, remains an open challenge. Existing methods that use metaheuristic algorithms such as Particle Swarm Optimization (PSO) or Artificial Bee Colony (ABC) often exhibit limitations in convergence or adaptability to dynamic conditions. To address these limitations, Optimization and intelligent routing in WSNs have attracted extensive research, focusing on metaheuristic algorithms, clustering techniques, and machine learning approaches. Recent advancements in WSNs have explored hybrid optimization, fuzzy logic, and machine learning-based approaches to address challenges in routing, clustering, and energy efficiency. Meshram et al. [1] introduced IBOOST, a lightweight, secure identity-based online/offline signature mechanism employing Fuzzy C-

Means (FCM) for 5G-based WSNs. It ensured provable security for massive device authentication but lacked adaptive routing strategies. Gangwar et al. [2] proposed a Game Theory-Based Fuzzy Routing (GTFR) protocol, which improved routing decisions in dynamic topologies. However, the model's energy efficiency under high mobility scenarios remained a limitation. Sikarwar and Tomar [3] combined Modified FCM with PSO for efficient tree-based routing. Though clustering was optimized, the approach did not account for network reconfiguration under node failure. Khedr et al. [4] presented a fuzzy-based multi-layered clustering model and Ant Colony Optimization (ACO)-driven sink path planning for optimal coverage. While enhancing network longevity, it did not fully address scalability in dense deployments. Cheng et al. [5] developed an FCM and hierarchical voting-based Received Signal Strength Indicator (RSSI) localization algorithm for sensor node positioning. Its accuracy was significant, yet performance degraded with increased environmental noise. Hiyagarajan and Shanmugasundaram [6] evaluated clustering techniques (K-Means, K-Medoids, FCM) for WSNs. Their comparative analysis offered insights into performance trade-offs, but real-time dynamic adaptability was absent. Bensaid and Boujemaa [7] proposed a combined cluster-chain routing protocol to extend network lifespan. Though energy consumption was

reduced, the protocol was not robust under unpredictable node mobility. Sree et al. [8] utilized FCM with Cat Swarm Optimization (CSO) for energy-efficient data gathering. Despite performance gains, it required frequent cluster reformation, increasing computation overhead.

Mohan et al. [9] introduced a Fuzzy Median Graph-Based Energy Efficient Clustering Protocol that minimized communication costs through median-based fuzzy decision-making. Zaier et al. [10] proposed an Interval Type-2 Fuzzy Unequal Clustering and Sleep Scheduling Protocol to handle uncertainty and balance energy consumption effectively in IoT-based WSNs. Rahmani et al. [11] introduced collective Gray Wolf Optimization with Fuzzy Clustering and multi-criteria decision-making approaches, improving throughput and reducing delay through optimized cluster head selection. Shokouhifar et al. [12] reviewed AI-driven clustering and routing protocols, emphasizing fuzzy, metaheuristic, and learning-based methods, and highlighting the need for adaptive and intelligent models in dynamic WSNs. Devika et al. [13] proposed an energy-efficient routing approach using ant-cuckoo hybrid techniques, enhancing data compression and energy savings. However, the scalability under heterogeneous nodes was limited. In a related work, Devika et al. [14] earlier introduced Ant Cuckoo optimized using Energy-Efficient Data aggregation (ACEED), a bio-inspired routing scheme combining ant and cuckoo behaviors. Though it addressed routing complexity, real-time performance under failure scenarios was insufficient. Karthikeyan and Venkatalakshmi [15] optimized clustering using PSO integrated with CS. The method effectively reduced energy use, yet suffered from slow convergence in large networks. Chang et al. [16] focused on recharge scheduling in WSNs via CS, improving node lifespan. Nonetheless, it overlooked optimal path selection during recharge intervals.

Ramadhan et al. [17] proposed an optimized event-based PID control mechanism to improve energy efficiency in Wireless Sensor Networks. Their approach dynamically adjusts control actions based on event triggers, reducing unnecessary energy consumption and extending network lifetime. Taheri et al. [18] introduced Probability Density Based Adaptive Clustering - Low Energy Adaptive Clustering Hierarchy (PDBAC-LEACH) an advanced clustering approach designed to optimize the lifespan of WSNs. The scheme enhances cluster-head selection and load balancing, thereby improving energy efficiency and extending network longevity. Chen et al. [19] proposed a trust-based, self-adaptive coverage model to ensure intrusion tolerance. While robust in hostile environments, the energy model used was static and non-adaptive. With the growing complexity of dynamic IoT networks, researchers have increasingly turned to Deep Reinforcement Learning (DRL) for adaptive and intelligent routing. Song et al. [20] presented High-Efficiency Routing Protocols for Heterogeneous WSNs (HWSNs) using DRL, where a deep Q-network (DQN) optimized routing based on residual energy, relay distance, and transmission delay, achieving superior energy balance and prolonged lifetime. Suresh et al. [21] proposed a Federated DRL-based Intelligent Data Routing Strategy for IoT-enabled WSNs, in which distributed learning among nodes improved scalability, reduced latency, and avoided single points of failure. Shekar et al. [22] Implemented Learning-Based Energy-Efficient Routing Protocols combining adaptive learning and clustering for IoT applications, demonstrating significant gains in energy conservation and load distribution under variable network

conditions. Similarly, Liu et al. [23] introduced Reinforcement Learning-Based Routing for energy-sensitive IoT mesh networks, which effectively balanced exploration and exploitation to achieve stable communication paths and reduced power consumption. Finally, Lingam et al. [24] applied PSO on deep reinforcement learning for spam bot detection in social networks. Though not directly WSN-related, the hybrid model showcased the effectiveness of combining metaheuristics with deep learning. The limitations of existing methods are summarized as follows:

- Metaheuristic-only approaches [2, 3, 5, 6] focus on deployment or clustering but lack adaptability in routing.
- FCM-based clustering techniques [9-12] improve energy balancing but do not dynamically adjust routing paths.
- DQL-based routing models [17-21] excel in path learning but assume static network structures, leading to suboptimal routing in dynamically placed nodes.

The positioning of this article is summarized as follows:

- Adaptive Nest Competition Algorithm (ANCA) is proposed as an improved variant of CS, offering better exploration and exploitation during node deployment.
- FCM is integrated to enable energy-aware clustering based on soft membership, reducing intra-cluster communication costs.
- DQL is incorporated for adaptive routing, leveraging real-time learning of optimal paths based on dynamic WSN parameters.
- Comparative evaluation with ABC, PSO, and Grey Wolf Optimizer (GWO) on metrics including Average Energy Consumption, Packet Delivery Ratio (PDR), End-to-End Delay, Routing Overhead, and Network Lifetime.

2. METHODOLOGY

Figure 1 shows the flow diagram of the proposed framework.

2.1 System overview

The proposed framework for optimizing energy efficiency and routing in dynamic WSNs is depicted in Figure 1. The overall system is structured into a sequential pipeline comprising four major stages: network parameter initialization, node placement, clustering, and deep reinforcement learning-based optimized routing. The system's performance is evaluated through key performance metrics. Network parameter initialization: the process begins with the initialization of essential network parameters, including node density, initial energy, communication range, and mobility patterns. These parameters define the simulation environment and influence all subsequent processes. The second stage is node placement using ANCA: To maximize coverage and maintain connectivity, an ANCA is employed for optimal node placement. ANCA, inspired by the nest competition behavior of birds, strategically distributes sensor nodes across the monitored region to ensure balanced energy consumption and coverage.

Clustering using FCM is the third stage after the node placement. Sensor nodes are logically grouped into clusters using the FCM clustering algorithm. FCM allows nodes to

have degrees of membership in multiple clusters, enabling flexible and energy-aware grouping. This step minimizes intra-cluster communication cost and enhances local data aggregation. The next stage is Optimized Clustering & Routing using DQL. The clustered network structure is further optimized using DQL for routing. DQL dynamically learns the best routing paths by interacting with the environment and adapting to changes such as node mobility and energy depletion. The objective is to find energy-efficient routes from

cluster members to the sink while minimizing delay and Routing Overhead. The effectiveness of the proposed ANCA-FCM-DQL framework is measured using the following key performance metrics: Network Lifetime, Average Energy Consumed, PDR, End-to-End Delay, and Routing Overhead.

This integrated model aims to strike a balance between energy efficiency and robust communication in dynamic environments, offering a scalable and adaptive solution for real-world WSN deployments.

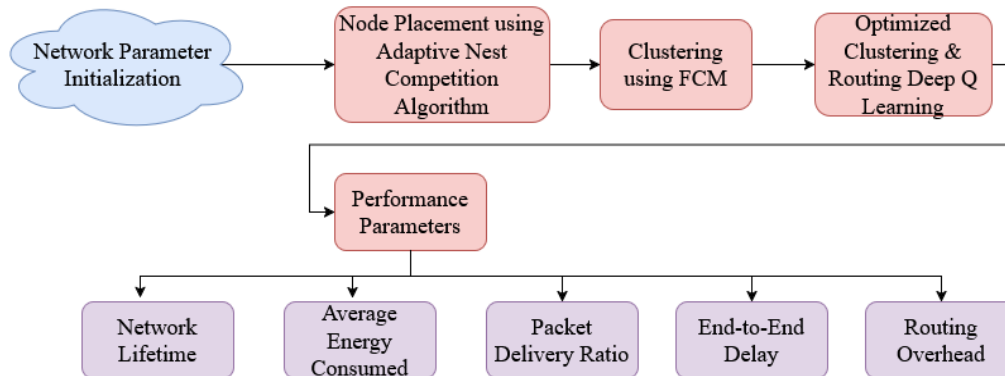


Figure 1. System flow

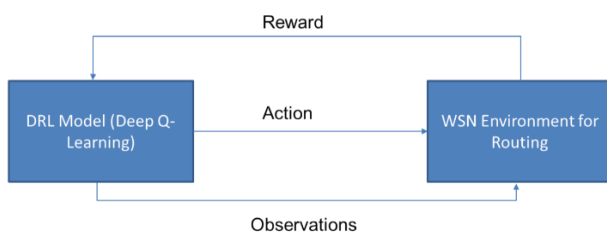


Figure 2. DQL for WSN

Figure 2 illustrates the interaction between the DQL model and the WSN environment. The DRL model (DQL) agent observes the network state, selects an optimal routing action, and receives feedback based on network performance. This feedback is used to update the Q-values, enabling the model to learn better routing policies over time. The loop ensures continuous adaptation to dynamic WSN conditions, enhancing energy efficiency and reliability.

2.2 ANCA for node placement

The ANCA draws inspiration from the natural reproductive strategy of cuckoo birds, particularly their unique approach of laying eggs in the nests of other bird species. In this biological process, if a host bird detects that an egg does not belong to it, it either discards the egg or abandons the nest entirely to construct a new one. Analogously, in the context of optimization, each nest represents a candidate solution, and each cuckoo egg symbolizes a promising or improved solution. The optimization process evolves by refining these solutions iteratively to identify the optimal outcome. In this model, a population of nests, each containing a potential solution, is maintained. The selection of the nest for laying the egg mimics the stochastic behavior of cuckoos and is governed by Levy flight, a random walk strategy that ensures exploration across a wide solution space. To further refine the search capability and enhance exploitation, an adaptive competition-based learning mechanism is introduced. This

strategy incorporates an elite selection and rivalry mechanism, in which members with superior performance are engaged in a competitive learning framework to generate more efficient solutions. This process promotes solution refinement without requiring entirely new individuals, thus improving convergence while preserving diversity.

Key Concepts of ANCA Competitive Strategy:

- (1). Elite selection: Two high-performing candidates are selected from the top-performing subset of the population (top 5% based on fitness).
- (2). Competition and replacement: The two candidates undergo a competitive evaluation, and based on the outcome, the weaker candidate is modified using the traits of the stronger one.
- (3). Fitness evaluation: After modification, both solutions are evaluated. The one with better fitness may replace the current global best if it outperforms it.

2.2.1 Algorithm steps: ANCA

Step 1:

Select two individuals (C1 and C2) at random from the top 5% elite set of the population.

$C1 \leftarrow$ Random selection from elite pool

$C2 \leftarrow$ Another random selection from the elite pool

Step 2:

Perform a competitive learning phase between C1 and C2 to generate modified versions.

$[C1', C2'] \leftarrow$ Competitive_Update (C1, C2)

Step 3:

Evaluate the fitness of updated candidates.

$fitness_C1' \leftarrow$ Evaluate (C1')

$fitness_C2' \leftarrow$ Evaluate (C2')

Step 4:

Update the global best solution if either C1' or C2' has better fitness.

If $fitness_C1' > Global_Best_Fitness$:

$Global_Best \leftarrow C1'$

If $fitness_C2' > Global_Best_Fitness$:

Global_Best $\leftarrow$ C2'

Step 5:

Repeat this competitive update across the population until convergence or a termination condition is met.

This approach significantly enhances the exploitation ability of the search process by continuously reusing and refining individuals near the global optimum. It also ensures rapid convergence through focused competition among elite candidates, making it particularly suitable for high-dimensional optimization tasks such as sensor node deployment in dynamic WSNs.

2.3 FCM clustering

In the proposed hybrid architecture for dynamic WSNs, FCM clustering plays a central role in managing energy-efficient data transmission by organizing sensor nodes into flexible, overlapping clusters. Unlike hard clustering methods, FCM minimizes intra-cluster distances while maintaining flexibility in cluster formation, which is critical in dynamic WSN environments. In the proposed system, FCM is employed after optimal node placement to form energy-aware clusters. This soft clustering strategy enhances load balancing and improves local data aggregation, thereby reducing overall energy consumption and communication overhead. This characteristic is advantageous in dynamic or mobile environments where node energy levels and topology change frequently.

2.4 DQL-based routing

To accompany the clustering framework, DQL is employed to dynamically manage routing both within and between clusters. DQL empowers sensor nodes to act as intelligent agents that learn optimal data forwarding paths by interacting with the environment and receiving feedback in the form of rewards or penalties. DQL is a value-based reinforcement learning algorithm that leverages deep neural networks to approximate the optimal action-value function. In the proposed framework, DQL is utilized to optimize the routing process within the clustered Wireless Sensor Network. Each sensor node acts as an agent that learns to select the most energy-efficient routing path based on a reward mechanism that considers factors such as residual energy, hop count, and link reliability.

The DQL agent interacts with the dynamic WSN environment, continuously updating its Q-values to adapt to node failures, mobility, and energy depletion. This learning-based routing approach ensures robust and adaptive communication from cluster members to the base station, effectively minimizing End-to-End Delay, Routing Overhead, and energy consumption. By integrating DQL, the system achieves intelligent decision-making capabilities that enhance the overall network lifetime and performance.

2.5 Mathematical models

(1). FCM clustering

Objective: Partition a set of sensor nodes into c clusters with soft membership, allowing each node to belong to multiple clusters.

Let:

$N = \{n_1, n_2, \dots, n_k\}$: Set of sensor nodes

$C = \{c_1, c_2, \dots, c_m\}$: Set of initial cluster centers from FCM

CH: Set of refined Cluster Heads selected by DQL

BS: Base Station location

$E(n_i)$: Residual energy of node n_i

d_{ij} : Distance between nodes n_i and n_j

D: End-to-End Delay

E_{avg} : Average Energy Consumed

L: Network Lifetime

R: Set of possible routes

(2). Initial clustering using FCM (soft assignment)

The membership matrix $U=[u_{ij}]$ is computed as:

$$u_{ij} = \frac{1}{\sum_{k=1}^m \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (1)$$

where:

x_i : Feature vector of node n_i

c_j : Cluster center

m : Fuzziness factor (typically 2)

Update cluster centers:

$$c_j = \frac{\sum_{i=1}^n u_{ij}^m x_i}{\sum_{i=1}^n u_{ij}^m} \quad (2)$$

(3). DQL state-space definition for clustering and routing

Each state s is defined as:

$$s = \left[E(n_i), d(n_i, BS), PDR(n_i), \text{hop count}, \text{buffer size, cluster assignment} \right] \quad (3)$$

Each action $a \in A$ can be:

Clustering: Elect node n_i as cluster head

Routing: Forward packet to neighbor n_j

The Q-value update is:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t) \right] \quad (4)$$

where:

s_t : Current state

a_t : Action taken

r_t : Reward received

α : Learning rate

γ : Discount factor for future rewards

$\max_{a'} Q(s_{t+1}, a')$: Maximum expected future reward from the next state

(4). Reward function r_t

In the proposed DQL-based routing mechanism, each sensor node (agent) learns to choose the most suitable next-hop node based on the current state of the network. The state-space is defined using key network parameters that reflect the current condition of a node and its neighbors. These parameters include:

- Residual Energy (E) of the node
- PDR of the link
- Distance (D) to the destination or cluster head
- Hop Count (H) from the current node to the sink

The agent evaluates these states to decide the best action, i.e., selecting the next-hop node for forwarding the data packet. The reward function guides the learning process by providing feedback after each action. It is designed to

encourage energy-efficient and reliable routing. The reward at time t is calculated as:

$$r_t = w_1 \cdot \Delta E + w_2 \cdot PDR + w_3 \cdot \frac{1}{D} + w_4 \cdot \frac{1}{H} \tag{5}$$

where:

ΔE : Change in residual energy (preferably low)
 PDR : Packet Delivery Ratio (higher is better)
 D : Distance to destination (shorter is preferred)
 H : Hop count to sink (fewer hops are ideal)
 w_1, w_2, w_3, w_4 : Weight factors controlling the influence of each metric:

$$w_1 + w_2 + w_3 + w_4 = 1$$

$PDR = \text{"Total packets received at destination"} / \text{"Total packets sent by sources"}$

Let:
 P_{recv} : Total successfully received packets.
 P_{sent} : Total packets sent.
 So:

$$PDR = \frac{P_{recv}}{P_{sent}} \tag{6}$$

$$\text{End-to-End Delay } D = \frac{\sum_{i=1}^N (t_{recv,i} - t_{sent,i})}{P_{recv}} \tag{7}$$

where:

$t_{recv,i}$: Time packet i was received.
 $t_{sent,i}$: Time packet i was sent.
 P_{recv} : Number of packets successfully received.

$$\text{Hop Count } H = \frac{\sum_{i=1}^{P_{recv}} h_i}{P_{recv}} \tag{8}$$

where:

h_i : Number of hops taken by packet i .
 P_{recv} : Total packets received.

3. RESULT AND DISCUSSION

This section presents the performance evaluation of the proposed ANC-DQL approach in comparison with existing optimization and routing schemes, including Artificial Bee Colony (ABC), Particle Swarm Optimization (PSO), and Grey Wolf Optimizer (GWO). The simulations were conducted using MATLAB with identical network settings to ensure fairness. The network and radio model simulation parameters are described in Table 1 and Table 2, respectively. Table 3 consists of performance parameters.

The proposed method consistently outperformed baseline algorithms across all evaluated metrics, as shown in Figures 3 to 7.

- Network Lifetime improved by 20–35%.
- Average Energy Consumption reduced by 15–25%.
- PDR increased by 10–18%.
- End-to-End Delay decreased by 20–30%.
- Routing Overhead reduced by 15–28%.

These improvements are attributed to the synergistic design of ANCA (for optimized node placement), FCM (for flexible and energy-aware clustering), and DQL (for adaptive,

intelligent routing). The model shows strong scalability with increasing node density and is especially effective in both static and mobile node environments, suggesting suitability for real-world applications like smart agriculture, disaster response, and industrial monitoring.

Table 1. Network simulation parameters

System Parameter	Specification
Base Station Position	Center
Simulation area	200 m × 200 m
Initial energy (Eo)	0.1-0.5 J
Number of Nodes	100, 200, 300, ...500
Node Position	Fixed and Mobile
Traffic Patterns	CBR

Table 2. Radio model parameters

System Parameter	Specification
Threshold Distance (do)	$\sqrt{E_{fs}/E_{mps}}$
Energy consumed per bit (Eelec)	50 nJ /bit
Receiver Power Consumption (ERX)	50 nJ/bit
Transmission Power Consumption (ETX)	50 nJ /bit
Multipath Amplification Factor (Emp)	0.0013 pJ/bit/m ⁴
Free Space Amplification Factor (Efs)	10 pJ/bit/m ²
Message bits (K)	2000 bits

Table 3. Performance metrics

Parameter	Meaning
Network Lifetime	Number of rounds until the first and last node dies.
Energy Consumption	Total residual energy of the network over time.
PDR	Ratio of successfully received packets to total sent packets.
End-to-End Delay	Average time taken for packets to reach the sink.
Routing Overhead	Number of control packets generated per data packet delivered.
Network Lifetime	Number of rounds until the first and last node dies.
Energy Consumption	Total residual energy of the network over time.

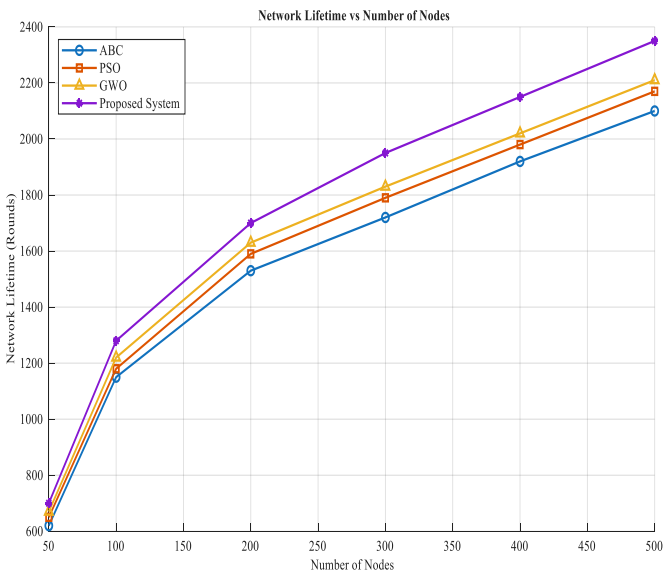


Figure 3. Network lifetime vs number of nodes

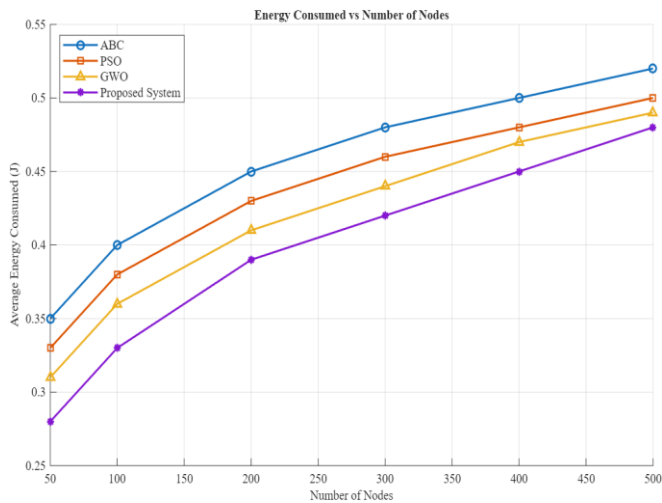


Figure 4. Average energy consumed vs number of nodes

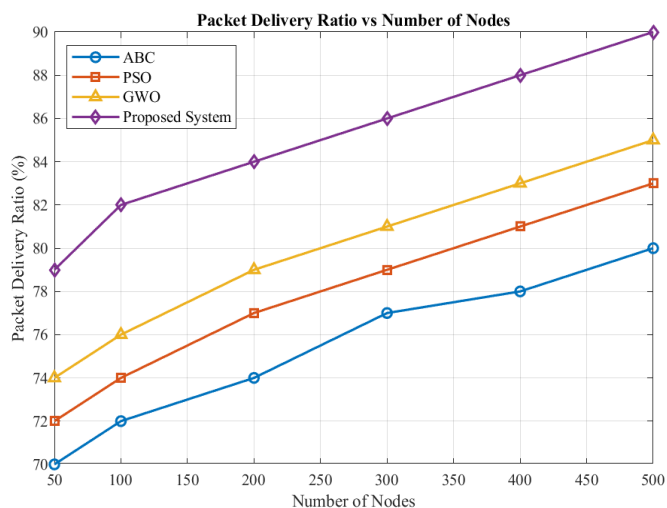


Figure 5. Packet Delivery Ratio vs. number of nodes

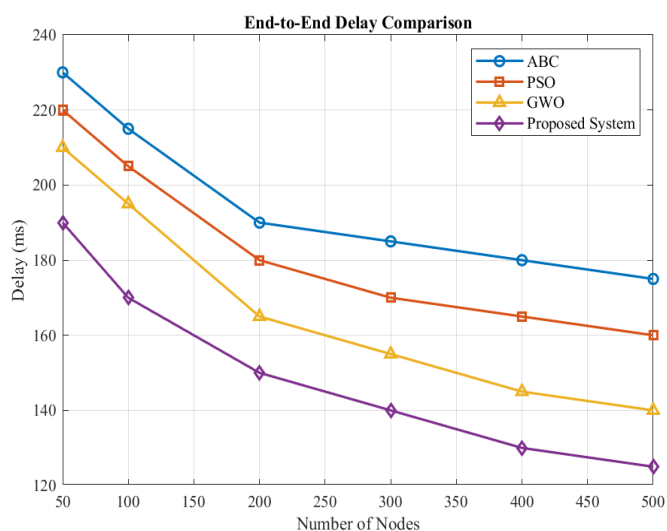


Figure 6. End-to-End Delay vs. number of nodes

Figure 3 shows network lifetime across varying node densities (50–500 nodes). The ANCA-FCM-DQL approach extends lifetime by up to 35%, due to balanced energy distribution achieved through adaptive nest-based placement and energy-aware clustering. Unlike PSO, GWO, and ABC, which cause energy holes around the sink, ANCA dynamically

adjusts node engagement to avoid early node death.

Figure 4 illustrates Average Energy Consumption. ANCA-FCM-DQL shows 15–25% lower consumption, thanks to FCM’s compact clusters and DQL’s ability to avoid long or congested routes through real-time learning. PSO, GWO, and ABC lack this adaptability and frequently trigger redundant transmissions.

Figure 5 displays the PDR, where our method maintains the highest delivery rates, even as node count increases. This is due to DQL’s policy learning for selecting reliable, stable paths, while the proposed system ensures fewer re-clustering events. Compared to traditional algorithms, the improvement in PDR is 10–18% on average.

Figure 6 shows the End-to-End Delay, which is 20–30% lower in our proposed method. DQL selects low-hop, high-quality routes and avoids route rediscovery, resulting in faster data transmission ideal for time-sensitive applications like surveillance and healthcare.

Figure 7 compares Routing Overhead. ANCA-FCM-DQL exhibits the lowest control packet overhead due to the stable, learned routing policies and minimal re-clustering. This reduces bandwidth consumption, making the protocol more scalable for dense or large-area WSN deployments.

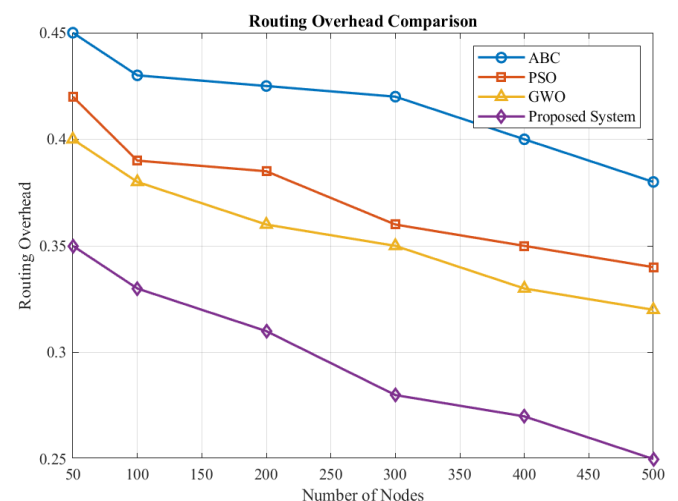


Figure 7. Routing Overhead vs. number of nodes

4. CONCLUSION

In this research, we proposed a novel, energy-efficient, and intelligent framework for WSNs by integrating the ANCA for optimal node placement, FCM for clustering, and DQL for dynamic routing. The hybrid ANCA-FCM-DQL approach effectively addresses challenges such as uneven energy depletion, suboptimal cluster formation, and inefficient routing under dynamic network conditions.

Comprehensive simulation experiments demonstrated that the proposed method significantly outperforms benchmark algorithms like ABC, PSO, and GWO in multiple performance metrics:

- Up to 20–35% improvement in network lifetime.
- 15–25% reduction in Average Energy Consumption.
- Enhanced PDR by 10–18%.
- Lower End-to-End Delay and reduced Routing Overhead.

These improvements are attributed to the synergistic combination of ANCA’s global optimization for node placement, FCM’s adaptive clustering capabilities, and DQL’s

intelligent, feedback-driven routing decisions.

Implications

- The framework demonstrates strong adaptability in dynamic WSN environments, with performance scaling effectively as node density increases.
- The integration of machine learning and bio-inspired algorithms opens new avenues for real-time decision-making in distributed sensor deployments.
- The proposed system architecture is generalizable and can be adapted for IoT-based monitoring, smart agriculture, and disaster-response systems, where energy efficiency and reliability are critical.

Future scope

- Amalgamation with real-time hardware testbeds (e.g., Arduino, Raspberry Pi with XBee modules) will be pursued to validate practical performance under real deployment conditions.
- Exploration of transfer learning or meta-reinforcement learning can reduce DQL training time in dynamic environments.

REFERENCES

- [1] Meshram, C., Imoize, A.L., Elhassouny, A., Aljaedi, A., Alharbi, A.R., Jamal, S.S. (2021). IBOOST: A lightweight provably secure identity-based online/offline signature technique based on FCM for massive devices in 5G wireless sensor networks. *IEEE Access*, 9: 131336-131347. <https://doi.org/10.1109/ACCESS.2021.3114287>
- [2] Gangwar, S., Prasad, I.B., Yadav, S.S., Pal, V., Kumar, N. (2023). GTFR: A game theory-based fuzzy routing protocol for WSNs. *IEEE Sensors Journal*, 24(6): 8972-8981. <https://doi.org/10.1109/JSEN.2023.3248226>
- [3] Sikarwar, N., Tomar, R.S. (2023). A hybrid MFCM-PSO approach for tree-based multi-hop routing using modified fuzzy C-means in wireless sensor network. *IEEE Access*, 11: 128745-128761. <https://doi.org/10.1109/ACCESS.2023.3331312>
- [4] Khedr, A.M., Al Aghbari, Z., Khalifa, B.E. (2022). Fuzzy-based multi-layered clustering and ACO-based multiple mobile sinks path planning for optimal coverage in WSNs. *IEEE Sensors Journal*, 22(7): 7277-7287. <https://doi.org/10.1109/JSEN.2022.3150065>
- [5] Cheng, L., Hang, J., Wang, Y., Bi, Y. (2019). A fuzzy C-means and hierarchical voting based RSSI quantify localization method for wireless sensor network. *IEEE Access*, 7: 47411-47422. <https://doi.org/10.1109/ACCESS.2019.2909974>
- [6] Thiagarajan, N., Shanmugasundaram, N. (2024). Accessing the performance of K-Medoid, K-Means and FCM clustering techniques for wireless sensor networks. In 2024 IEEE 5th India Council International Subsections Conference (INDISCON), Chandigarh, India, pp. 1-5. <https://doi.org/10.1109/INDISCON62179.2024.10744273>
- [7] Bensaid, R., Boujemaa, H. (2022). A combined cluster-chain based routing protocol for lifetime improvement in WSN. In 2022 International Wireless Communications and Mobile Computing (IWCMC), Dubrovnik, Croatia, pp. 542-547. <https://doi.org/10.1109/IWCMC55113.2022.9824293>
- [8] Sree, R.N., Ananth, A.G., Reddy, L.S. (2018). An energy efficient FCM and CAT swarm optimization based data gathering in WSN. In 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECOT), Msyuru, India, pp. 735-740. <https://doi.org/10.1109/ICEECOT43722.2018.9001375>
- [9] Mohan, A., Thalapala, V.S., Guravaiah, K., Dhanyamol, M.V. (2025). Fuzzy median graph based energy efficient clustering protocol in WSN. *Information Sciences*, 721: 122573. <https://doi.org/10.1016/j.ins.2025.122573>
- [10] Zaier, A., Lahmar, I., Yahia, M., Lloret, J. (2025). Interval type 2 fuzzy unequal clustering and sleep scheduling for IoT-based WSNs. *Ad Hoc Networks*, 175: 103867. <https://doi.org/10.1016/j.adhoc.2025.103867>
- [11] Rahmani, A.M., Haider, A., Ali, S., Mohammadi, M., Mehranzadeh, A., Khoshvaght, P., Hosseinzadeh, M. (2025). A routing approach based on combination of gray wolf clustering and fuzzy clustering and using multi-criteria decision making approaches for WSN-IoT. *Computers and Electrical Engineering*, 122: 109946. <https://doi.org/10.1016/j.compeleceng.2024.109946>
- [12] Shokouhifar, M., Fanian, F., Rafsanjani, M.K., Hosseinzadeh, M., Mirjalili, S. (2024). AI-driven cluster-based routing protocols in WSNs: A survey of fuzzy heuristics, metaheuristics, and machine learning models. *Computer Science Review*, 54: 100684. <https://doi.org/10.1016/j.cosrev.2024.100684>
- [13] Devika, G., Ramesh, D., Karegowda, A.G. (2019). An energy efficient routing and compression based data collection applying bio-inspired ant-cuckoo technique for wireless sensor network. In 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), Bengaluru, India, pp. 1-8. <https://doi.org/10.1109/CSITSS47250.2019.9031048>
- [14] Devika, G., Karegowad, A.G., Ramesh, D. (2017). Bio-inspired ant-cuckoo energy efficient data aggregation algorithm: A solution for routing problem of wireless sensor networks [ACEED]. In 2017 2nd International Conference on Emerging Computation and Information Technologies (ICECIT), Tumakuru, India, pp. 1-8. <https://doi.org/10.1109/ICECIT.2017.8453438>
- [15] Karthikeyan, M., Venkatalakshmi, K. (2012). Energy conscious clustering of wireless sensor network using PSO incorporated cuckoo search. In 2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12), Coimbatore, India, pp. 1-7. <https://doi.org/10.1109/ICCCNT.2012.6395963>
- [16] Chang, H., Feng, J., Duan, C., Xu, Z., Yin, M. (2018). Research of recharging scheduling scheme for wireless sensor networks based on cuckoo search. In 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, pp. 1-7. <https://doi.org/10.1109/IJCNN.2018.8489112>
- [17] Ramadhan, N.M., Raafat, S.M., Mahmood, A.M. (2024). Optimized event-based PID control for energy-efficient wireless sensor networks. *Mathematical Modelling of Engineering Problems*, 11(1): 63-74. <https://doi.org/10.18280/mmep.110106>
- [18] Taheri, S.M., Alalwany, W.S.H., Yonan, J.F. (2024). Optimizing wireless sensor network lifespan through

- advanced clustering in PDBAC-LEACH. *Mathematical Modelling of Engineering Problems*, 11(11): 3047-3060. <https://doi.org/10.18280/mmep.111117>
- [19] Chen, Z., Li, X., Yang, B., Zhang, Q. (2015). A self-adaptive wireless sensor network coverage method for intrusion tolerance based on trust value. *Journal of Sensors*, 2015(1): 430456. <https://doi.org/10.1155/2015/430456>
- [20] Song, Y., Liu, Z., Li, K., He, X., Zhu, W. (2024). Research on high-efficiency routing protocols for HWSNs based on deep reinforcement learning. *Electronics*, 13(23): 4746. <https://doi.org/10.3390/electronics13234746>
- [21] Suresh, S.S., Prabhu, V., Parthasarathy, V., Senthilkumar, G., Gundu, V. (2024). Intelligent data routing strategy based on federated deep reinforcement learning for IOT-enabled wireless sensor networks. *Measurement: Sensors*, 31: 101012. <https://doi.org/10.1016/j.measen.2023.101012>
- [22] Shekar, K., Reddy, N.R., Arvind, S., Kumar, T.V.S., Kodukula, S., Varahagiri, G. (2025). Implementation of novel learning based energy efficient routing protocols in wireless sensor networks for internet of things use cases. *Discover Computing*, 28: 190. <https://doi.org/10.1007/s10791-025-09718-8>
- [23] Liu, Y., Tong, K.F., Wong, K.K. (2019). Reinforcement learning based routing for energy sensitive wireless mesh IoT networks. *Electronics Letters*, 55(17): 966-968. <https://doi.org/10.1049/el.2019.1864>
- [24] Lingam, G., Rout, R.R., Somayajulu, D.V., Ghosh, S.K. (2020). Particle swarm optimization on deep reinforcement learning for detecting social spam bots and spam-influential users in twitter network. *IEEE Systems Journal*, 15(2): 2281-2292. <https://doi.org/10.1109/JSYST.2020.3034416>

NOMENCLATURE

CH	Cluster Heads
DQL	Deep-Q-Learning
BS	Base Station location
$E(n_i)$	Residual energy of node n_i
d_{ij}	Distance between nodes n_i and n_j
PDR	Packet Delivery Ratio
D	End-to-End Delay
E_{avg}	Average Energy Consumed
L	Network Lifetime
R	Set of possible routes