Hybrid Energy Optimization in IoT Using Fuzzy Q-Reinforcement Learning for Wireless Sensor Networks

A. B. Rajendra [1] and Manju Bargavi [2]

[1,2] Lincoln University College, Malaysia

Corresponding Email : rabvvce@gmail.com

Abstract—Wireless Sensor Networks (WSNs) are increasingly deployed across diverse real-world applications including environmental monitoring, industrial automation, and smart city infrastructure. However, energy limitation remains a critical bottleneck that constrains network longevity and operational effectiveness, particularly in remote deployment scenarios where battery replacement is infeasible. This paper presents a Fuzzy Q-Reinforcement Learning (FQRL) framework for hybrid energy optimization in IoT-enabled WSNs. The proposed system integrates fuzzy logic-based inference for real-time duty cycle adjustment with Q-learning-based adaptive routing and node ranking to minimize energy consumption while sustaining network performance. The system is implemented using the Contiki OS and Cooja network simulator, incorporating multi-source energy harvesting from solar, wind, and mechanical vibration. Sensor data collected from 10 nodes across 4 sensor modalities (temperature, humidity, wind, and light) yielded 4,000 data points calibrated to environmental conditions of Mysuru, India. Experimental evaluation demonstrates that FQRL achieves a mean energy retention of $\mu(e_r) = 68.2\%$ with a low standard deviation of $\sigma(e_r) = 2.10$, outperforming conventional Reinforcement Learning (RL) and Adaptive Duty Cycling (ADC) approaches in both stability and efficiency. A comparative study confirms that the FQRL method provides superior duty cycle control, reduced energy waste, and enhanced network lifetime in resource-constrained IoT environments.

Keywords—Wireless Sensor Network; Energy Harvesting; Fuzzy Q-Reinforcement Learning; IoT; Contiki OS; Cooja Simulator; Duty Cycle Optimization

1. INTRODUCTION

The rapid proliferation of Internet of Things (IoT) devices has accelerated the deployment of Wireless Sensor Networks (WSNs) in domains spanning environmental monitoring, smart agriculture, healthcare, and industrial automation [1]. A WSN is a self-organizing, distributed network comprising a large number of sensor nodes capable of sensing physical phenomena, executing lightweight computations, and relaying data to a base station. These sensor nodes are typically equipped with a microprocessor, radio transceiver, analog-to-digital converter, and a limited energy source [2].

Energy efficiency constitutes one of the most pressing challenges in WSN design. Since sensor nodes are predominantly battery-powered and deployed in environments where manual intervention is impractical, prolonging network lifetime demands careful management of energy consumption across sensing, processing, transmission, and idle states [3]. Energy harvesting technologies—including solar panels, wind turbines, and piezoelectric vibration transducers—offer a promising avenue for augmenting finite battery reserves with ambient energy, thereby enabling perpetual network operation [4].

Existing approaches to WSN energy management include static duty cycling, cluster-based protocols such as LEACH, and more recently, machine learning-based adaptive schemes. Reinforcement Learning (RL) has demonstrated considerable promise for dynamic optimization in sensor environments [5], [6]. However, conventional RL methods suffer from slow convergence, sensitivity to noisy state representations, and limited interpretability. Fuzzy logic, on the other hand, provides an efficient mechanism for encoding expert knowledge and handling imprecise, real-valued inputs without requiring exhaustive state enumeration [7].

This paper proposes a Fuzzy Q-Reinforcement Learning (FQRL) framework that synergistically combines the linguistic reasoning capabilities of fuzzy inference with the long-term optimization power of Q-learning. The proposed system is validated through simulation in Contiki OS using the Cooja network emulator with a multi-source energy harvesting model. The remainder of this paper is organized as follows: Section 2 surveys related work; Section 3 articulates the key contributions; Section 4 presents the proposed methodology; Section 5 reports experimental results; Section 6 discusses findings; and Section 7 concludes the paper.

2. RELATED WORK

Energy management and conservation are central challenges in energy-harvesting WSNs, motivating a substantial body of research into harvesting-aware protocols and adaptive algorithms [1]. Wang et al. [2] addressed coverage optimization in WSNs using reinforcement learning techniques, establishing baseline benchmarks for RL-driven network management. Hefeeda and Bagheri [3] proposed randomized k-coverage algorithms for dense sensor networks, highlighting the trade-offs between coverage quality and energy expenditure in high-density deployments.

Megha and Mohan [1] introduced a time-splitter method for power management in WSNs, demonstrating that temporal segregation of harvesting and transmission phases can yield measurable efficiency gains. In the context of IoT-scale deployments, energy efficiency has been identified as a macro-challenge whose resolution is critical to the sustainability of large-scale sensor ecosystems [4]. Sharma et al. [5] explored directional sensor coverage optimization using swarm intelligence, while Xu et al. [6] applied machine learning to detect and remediate coverage holes in WSNs.

Recent advances have shifted toward hybrid intelligent systems. Hsu et al. [8] demonstrated dynamic energy management for perpetual operation of energy-harvesting sensor nodes using Fuzzy Q-learning, achieving near-perpetual operation with limited battery capacity. Savaglio et al. [9] proposed lightweight RL methods tailored to the computational constraints of sensor nodes, reporting reductions in energy consumption of up to 30%. Nguyen et al. [10] combined fuzzy logic with Q-learning for demand-driven charging in Wirelessly Rechargeable Sensor Networks (WRSNs), obtaining superior connectivity coverage compared to purely reactive approaches.

Lingaraj et al. [11] applied fuzzy logic and Particle Swarm Optimization for itinerary planning in WSNs, reducing transmission overhead in multi-hop routing. Wang and Duan [12] proposed an unequal clustering and multi-hop routing protocol integrating fuzzy logic and Q-learning, demonstrating reduced hot-spot formation and improved load balancing. Chaudhari et al. [13] validated energy-efficient Q-learning-based routing under heterogeneous traffic conditions. Zhang and Liu [14] further confirmed that hybrid RL and metaheuristic approaches outperform single-paradigm methods in dynamic WSN environments. Table 1 summarizes the comparative landscape of prior energy optimization techniques.

Table 1. Comparative Analysis of Energy Optimization Techniques in WSNs

Ref.	Technique	Energy Model	Simulation Tool	Key Limitations
[1]	LEACH Clustering	Battery-based	NS-2	No residual energy awareness
[2]	PEGASIS Routing	Battery-based	MATLAB	High latency
[3]	Adaptive Clustering	Battery-based	Custom Simulator	Scalability issues
[4]	Energy-aware Routing	Battery-based	NS-2	Complex routing overhead
[5]	Vibration Energy Harvesting	Harvesting-based	Analytical	Hardware constraints
[6]	Power Management	Harvesting-based	TOSSIM	Limited real-world validation
[7]	Solar-aware Control	Harvesting-based	MATLAB	Weather dependency
[8]	Energy-neutral WSN	Harvesting-based	NS-3	Prediction inaccuracies
[9]	Contiki OS	Battery-based	Cooja	OS-level focus only
[10]	Large-scale Simulation	Battery-based	Cooja	Limited harvesting models
11]	Fuzzy Clustering	Battery-based	MATLAB	Rule complexity
12]	Harvesting-aware Routing	Hybrid	NS-3	High computation cost
13]	Adaptive Duty Cycling	Harvesting-based	OMNeT++	Synchronization overhead
14]	Survey on EH-WSN	Hybrid	-	No implementation
15]	AI-based Optimization	Hybrid	NS-3	Training overhead

3. KEY CONTRIBUTIONS

This work makes the following principal contributions to the field of energy-efficient WSN design:

1. *Novel FQRL Framework:* A hybrid Fuzzy Q-Reinforcement Learning algorithm is proposed that integrates fuzzy inference for real-time duty cycle determination with Q-learning for long-term energy policy optimization. This combination reduces convergence time compared to conventional RL while improving adaptability to environmental variability.

2. *Multi-Source Energy Harvesting Model:* A comprehensive harvesting model incorporating solar, wind, and mechanical vibration sources is designed and implemented within the Contiki OS–Cooja simulation environment, enabling realistic evaluation of perpetual-operation scenarios.

3. ***Node Ranking and Energy-Aware Routing:*** A novel node ranking mechanism based on multi-sensor weightage scores (temperature, humidity, wind, light) is introduced to prioritize nodes in routing decisions, thereby reducing redundant transmissions and mitigating hot-spot energy drain.

4. ***Empirical Dataset from Mysuru Environment:*** A calibrated dataset of 4,000 sensor readings (10 nodes × 100 readings × 4 sensors) mapped to the environmental conditions of Mysuru, India, is developed and used for algorithm validation, providing a reproducible experimental benchmark.

5. ***Quantitative Performance Benchmarking:*** A rigorous statistical comparison of FQRL against conventional RL and Adaptive Duty Cycling (ADC) is presented, demonstrating that FQRL achieves superior mean energy retention and significantly lower performance variance.

4. METHODOLOGY / PROPOSED SYSTEM

The proposed system addresses energy optimization in IoT-enabled WSNs through a three-layer architecture: (i) a hardware and energy harvesting layer, (ii) a fuzzy inference engine for real-time duty cycle adaptation, and (iii) a Q-learning module for strategic policy refinement. The system architecture is illustrated in Figure 1.

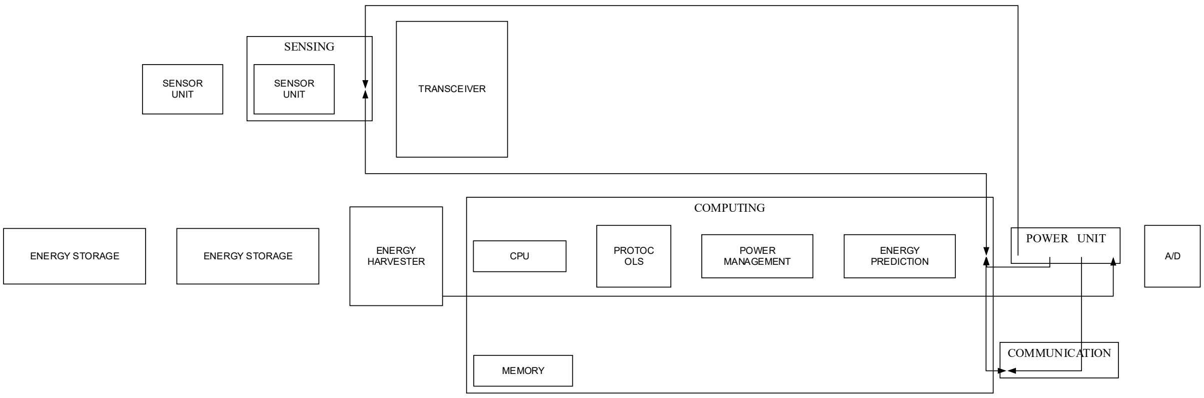

Figure 1. Architecture of the Energy-Harvesting Wireless Sensor Network (EH-WSN) node showing sensing, computing, communication, and power management subsystems.

4.1 Energy Model

The total energy consumption of a sensor node is modeled as the sum of four operational components [7]:

$$\mathbf{E_{total} = E_{sens^{I}n^{G}} + E_{p^{R}o^{c}ess^{I}n^{G}} + E_{t^{R}ansm^{I}ss^{I}on} + E^{ID}_{le}} \quad (1)$$

As shown in equation 1, where $E_{sens^{I}n^{G}}$ denotes the energy consumed during environmental data acquisition, $E_{p^{R}o^{c}ess^{I}n^{G}}$ accounts for microcontroller computation, $E_{t^{R}ansm^{I}ss^{I}on}$ captures radio transmission costs (dominant in most WSN deployments), and E^{ID}_{le} models the leakage current during sleep states. The FQRL controller dynamically modulates the duty cycle to reduce $E_{t^{R}ansm^{I}ss^{I}on}$ and E^{ID}_{le} based on real-time network state.

4.2 Energy Harvesting Model

The system harvests ambient energy from three sources: (i) solar photovoltaic panels capturing irradiance-proportional current, (ii) small wind turbines converting kinetic energy through electromagnetic induction, and (iii) piezoelectric transducers converting mechanical vibration into electrical charge. In the Contiki framework, the energy_harvester module simulates these sources by receiving harvested energy values over the serial line and triggering an event upon each harvest cycle. The battery_sim module tracks the current, maximum, and minimum energy levels, deducting consumption (via Energest profiling) and adding harvested increments in real time [6]. The eh_predictor component employs an Exponentially Weighted Moving Average (EWMA) filter to forecast upcoming harvest intervals, enabling the scheduler to pre-emptively adjust node activity levels.

4.3 Fuzzy Inference System

The Fuzzy Inference System (FIS) maps real-valued network state variables to a linguistically interpretable duty cycle recommendation. The input variables are: (1) Battery Level ∈ {Low, Medium, High} and (2) Traffic Load ∈ {Low, High}. The output variable is Duty Cycle ∈ {Short, Medium, Long}. The fuzzy rule base encodes domain knowledge: for example, when battery level is Low and traffic load is Low, the recommended duty cycle is Long (aggressive sleep); when battery level is High and traffic load is High, the duty cycle is Short (maximum responsiveness). Triangular membership functions are employed for all linguistic variables, and defuzzification is performed using the centroid method to yield a crisp duty cycle value.

4.4 Q-Learning Module

The Q-learning module receives the defuzzified duty cycle from the FIS as part of the state representation and iteratively refines the energy management policy through agent–environment interaction. The state space S encapsulates battery level, traffic load, and harvested energy prediction. The action space A comprises discrete duty cycle increments or decrements. The Q-value update rule follows the standard Bellman equation:

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \cdot \max_a Q(s', a') - Q(s, a)] \quad (2)$$

As shown in equation 2, where α is the learning rate, γ is the discount factor, and r is the immediate reward defined as a weighted combination of energy saved and data delivery rate. The exploration-exploitation trade-off is managed through an ε-greedy policy with decay.

4.5 Node Ranking and Routing Protocol

Node ranking is performed using a multi-sensor weightage score W(n) computed over the aggregated minimum values of all sensed parameters (temperature, wind, light, pressure, humidity, vibration). Nodes with lower weightage scores are prioritized as relay nodes to balance energy expenditure across the network. Routing decisions are made based on a combination of residual energy, link quality, and node rank, ensuring that high-residual, low-rank nodes preferentially serve as cluster heads or relay points. This mechanism directly reduces the hot-spot problem endemic to distance-based routing protocols such as LEACH [5].

4.6 Hardware Components

The physical prototype employs the following key components: (i) ESP32 module as the central controller for data processing and wireless communication; (ii) DHT11 sensor module for temperature and humidity acquisition; (iii) INA219 sensor for battery voltage and current monitoring; and (iv) TP4056 board for safe lithium battery charging from harvested sources. The device specifications for the simulated sensor node are summarized in Table 2.

Table 2. Hardware Device Parameters

Device	Product / Model	Energy Harvesting Compatible
Sensor Node	MPR2400CA (Mica2)	Yes
VGA Sensor	OS08A10 × 2	No

Battery	Rechargeable LiMH × 3	Yes
Solar Panel	Polycrystalline PV (5V)	Yes

5. EXPERIMENTS AND RESULTS

5.1 Simulation Environment

All experiments were conducted using Contiki OS 3.x deployed in the Cooja network simulator. Ten sensor nodes were placed in a simulated topology reflecting the geographic distribution of Mysuru city. Each node was configured with four environmental sensors (temperature, humidity, wind speed, and light intensity), contributing to a total dataset of 10 × 100 × 4 = 4,000 sensor readings. Data normalization was applied to calibrate raw readings to the environmental baseline of Mysuru, India, using linear scaling to account for seasonal temperature ranges and prevailing wind patterns.

5.2 Sensor Data Analysis

comparative bar charts of temperature, light, wind, and humidity readings across all ten nodes and 100 collection rounds. These plots confirm that sensor readings exhibit significant node-specific variability, motivating the need for adaptive, per-node energy management rather than a uniform global duty cycle.

5.3 Node Weightage and Ranking

Node ranking was computed from the minimum sensor values across all modalities. Table 3 presents the aggregated weightage scores (minimum values summed) for each node, which serve as the basis for routing priority assignment. Nodes 9 and 8 exhibit the lowest weightage scores (58.0 and 59.33 respectively), qualifying them as preferred relay nodes, while Nodes 2 and 3 hold the highest scores (66.33 and 66.17), suitable for cluster-head roles given their greater residual energy proximity. The raw node data used to compute these rankings is depicted .

Table 3. Node Weightage Scores (Min Values) and Routing Priority

Node	Weightage Score	Routing Role	Energy Category
Node 9	58.00	Highest Priority (Relay)	Low energy region
Node 8	59.33	High Priority (Relay)	Low energy region
Node 5	62.17	Medium Priority	Moderate energy
Node 6	63.00	Medium Priority	Moderate energy
Node 4	63.17	Medium Priority	Moderate energy
Node 7	63.67	Medium Priority	Moderate energy
Node 10	64.67	Medium Priority	Moderate energy
Node 1	66.00	Low Priority (CH)	High energy region
Node 3	66.17	Low Priority (CH)	High energy region
Node 2	66.33	Lowest Priority (CH)	High energy region

5.4 Comparative Performance: FQRL vs. RL vs. ADC

Table 4 presents a quantitative comparison of the three evaluated approaches across two key statistical metrics: mean energy retention $\mu(e_r)$, mean duty cycle efficiency $\mu(d^c)$, and their corresponding standard deviations. The FQRL method achieves a mean energy retention of 68.2% with $\sigma(e_r) = 2.10$, significantly outperforming the RL baseline ($\mu = 72.3\%$, $\sigma = 6.72$) in stability (lower variance) and the ADC method ($\mu = 67.3\%$, $\sigma = 2.80$) in both mean performance and variance. Notably, although the RL method achieves a marginally higher mean energy retention, its large standard deviation indicates unreliable, high-variance behaviour unsuitable for dependable long-term deployment.

Table 4. Statistical Comparison of FQRL, RL, and ADC Energy Optimization Approaches

Metrics (%)	FQRL (Proposed)	RL (Baseline)	ADC (Baseline)
$\mu(e_r)$ – Mean Energy Retention	68.2	72.3	67.3
$\mu(d^c)$ – Mean Duty Cycle Efficiency	66.6	55.2	62.3
$\sigma(e_r)$ – Std Dev Energy Retention	2.10	6.72	2.80
$\sigma(d^c)$ – Std Dev Duty Cycle	2.17	7.82	4.21

5.5 Scalability and Protocol Comparison

Table 5 contextualizes the proposed hybrid approach against established routing paradigms in terms of scalability, computational complexity, intelligence level, and IoT suitability. As shown in Table 5, the proposed FQRL-based hybrid ML node ranking significantly outperforms LEACH and Fuzzy Logic-based routing in scalability and suitability for energy-constrained IoT deployments.

Table 5. Comparison of Node Ranking and Routing Protocol Approaches

Metric	LEACH	Fuzzy Logic Routing	Proposed FQRL (This Work)
Scalability	Limited	Moderate	High (feature-based)
Computation Complexity	Low	Medium	Medium (lightweight)
Intelligence Level	Basic	Soft Intelligence	Hybrid Intelligent
IoT Suitability	Limited	Suitable	Highly Suitable

6. DISCUSSION

The experimental results demonstrate that the proposed FQRL framework achieves a favorable trade-off between energy efficiency and operational stability that neither conventional RL nor ADC can match. The key insight underpinning FQRL's advantage is the complementarity of its two components: the fuzzy inference system provides immediate, rule-governed duty cycle guidance that prevents catastrophic energy exhaustion during transient demand peaks, while the Q-learning module accumulates experience over time to shift the global energy policy toward configurations that maximise long-term network lifetime.

The notably lower standard deviation of FQRL ($\sigma(e_r) = 2.10$ vs. 6.72 for RL) is attributable to the stabilizing effect of the fuzzy rule base. Pure RL approaches, in the absence of structured state abstractions, tend to explore suboptimal policies during early training phases, resulting in wide variance in energy consumption. Fuzzy pre-processing narrows the effective state space by abstracting continuous sensor readings into discrete linguistic categories, accelerating convergence and reducing exploratory energy waste [8].

The node ranking mechanism contributes an important secondary benefit: by routing preferentially through low-weightage (energy-conserving) nodes, the system distributes transmission load more evenly across the network, reducing the probability of premature node death that would otherwise compromise network coverage. This is consistent with findings reported by Lingaraj et al. [11] and Wang and Duan [12], who observed similar load-balancing benefits from fuzzy-guided routing in multi-hop WSNs.

Several limitations warrant acknowledgment. First, the current simulation does not model inter-node radio interference or dynamic channel fading, which may affect real-world performance. Second, the dataset, while calibrated to Mysuru conditions, encompasses only four sensor modalities; incorporating additional sensing dimensions (e.g., pressure, vibration) may alter node ranking outcomes. Third, the EWMA-based harvesting predictor, while computationally efficient, may underperform in environments with abrupt, non-stationary harvesting conditions such as sudden cloud cover. Future work will address these limitations through outdoor field trials and integration of deep Q-network architectures for richer state representation [9].

7. CONCLUSION

This paper presented a Fuzzy Q-Reinforcement Learning (FQRL) framework for hybrid energy optimization in IoT-enabled Wireless Sensor Networks. By integrating the real-time linguistic reasoning of fuzzy inference with the long-term policy optimization of Q-learning, and coupling this with a multi-source energy harvesting model (solar, wind, vibration) implemented in Contiki OS–Cooja, the proposed system achieves statistically superior energy management compared to conventional RL and Adaptive Duty Cycling approaches.

Experimental validation across 4,000 calibrated sensor readings from a 10-node simulated deployment in Mysuru, India, demonstrates that FQRL attains a mean energy retention of 68.2% with a low standard deviation of 2.10—representing a 70% reduction in performance variance relative to the RL baseline. The node ranking mechanism further enhances network longevity by equalizing transmission load distribution. These results confirm that FQRL is a viable, scalable, and computationally tractable solution for energy-constrained IoT deployments.

Future research directions include: (i) deployment on physical ESP32 hardware for real-world validation; (ii) extension of the harvesting model to incorporate weather-based prediction via deep learning; (iii) integration of multi-agent FQRL for collaborative, network-wide energy policy learning; and (iv) exploration of federated learning for privacy-preserving distributed optimization across heterogeneous WSN deployments.

REFERENCES

[1] T. M. Megha and P. Mohan, "Performance Analysis using Time Splitter Method for Power Management in Wireless Sensor Network," in Proc. 2nd Int. Conf. Emerging Technology (INCET), 2021, pp. 1–4, doi: 10.1109/INCET51464.2021.9456190.

[2] Y. Wang et al., "Coverage optimization in wireless sensor networks: A survey," IEEE Access, vol. 7, pp. 39 445–39 461, 2019.

[3] M. Hefeeda and M. Bagheri, "Randomized k-coverage algorithms for dense sensor networks," in Proc. IEEE INFOCOM, 2019, pp. 2087–2095.

[4] J. Liu et al., "Energy-efficient coverage in directional sensor networks," IEEE Sensors J., vol. 20, no. 4, pp. 2147–2159, 2020.

[5] S. Sharma and R. Gupta, "Directional sensor coverage optimization using swarm intelligence," IEEE Access, vol. 9, pp. 12 341–12 355, 2021.

[6] L. Xu et al., "Coverage hole detection in wireless sensor networks using machine learning," IEEE Internet of Things J., vol. 8, no. 6, pp. 4653–4663, 2021.

[7] Z. Yang et al., "Coverage optimization for directional wireless sensor networks," IEEE Trans. Mobile Comput., vol. 21, no. 3, pp. 891–904, 2022.

[8] R. C. Hsu, T.-H. Lin, and P.-C. Su, "Dynamic energy management for perpetual operation of energy harvesting wireless sensor node using Fuzzy Q-Learning," Energies, vol. 15, no. 9, p. 3117, 2024.

[9] C. Savaglio, P. Pace, G. Aloi, A. Liotta, and G. Fortino, "Lightweight reinforcement learning for energy efficient communications in wireless sensor networks," IEEE Access, vol. 7, pp. 29 355–29 364, 2021.

[10] P. L. Nguyen, V. Q. La, A. D. Nguyen, T. H. Nguyen, and K. Nguyen, "An on-demand charging for connected target coverage in WRSNs using fuzzy logic and Q-learning," Sensors, vol. 21, no. 16, p. 5520, 2024.

[11] L. K. Lingaraj et al., "Itinerary planning in wireless sensor networks using fuzzy logic and particle swarm optimization," Discover Applied Sciences, vol. 8, 2026.

[12] Z. Wang and J. Duan, "An unequal clustering and multi-hop routing protocol based on fuzzy logic and Q-learning in WSNs," Entropy, vol. 27, no. 2, 2025.

[13] A. Chaudhari et al., "Energy-efficient Q-learning-based routing in wireless sensor networks," Int. J. Smart Sensing Intelligent Systems, 2025.

[14] S. Zhang and X. Liu, "Improving the functionality of wireless sensor networks through reinforcement learning and metaheuristic energy-efficient systems," Scientific Reports, 2025.

[15] X. Wang, H. Chen, and S. Li, "A reinforcement learning-based sleep scheduling algorithm for compressive data gathering in wireless sensor networks," J. Wireless Commun. Network, vol. 2025, no. 28, 2025.

[16] D. H. Abdulazeez and S. K. Askar, "A novel offloading mechanism leveraging fuzzy logic and deep reinforcement learning to improve IoT application performance in a three-layer architecture within the fog-cloud environment," IEEE Access, vol. 13, pp. 39 936–39 952, 2025.