Journal of Computer Science Innovations and Research
JCSIR · OPEN ACCESS · PEER-REVIEWED

Vol. 2, No. 1, 2026, pp. 1–10 **E-ISSN: 3136-800X**

Research Article

LEACH-Quantum Leap: A Tri-Hybrid Metaheuristic and Reinforcement Learning Protocol for Ultra-Energy-Efficient WSNs

May K Mahdi[1,*], Mustafa Musa[2]

[1]*Informatics Institute for Postgraduate Studies, Iraq*
[2]*Center of Research and Innovation Management, Universiti Teknikal Malaysia Melaka, Malaysia*
*Corresponding author: Maymahdi@iips.edu.iq

Published: 24 January 2026
DOI: https://doi.org/10.69055/jcsir.v2i1.6 **License:** CC BY 4.0

Abstract

Because wireless sensor networks (WSNs) run on finite battery resources, the design of energy-efficient protocols is critical to a longer network lifetime. Approaches built on Fuzzy C-Means Clustering (FCMC), Fuzzy-logic Cluster Head Rotation (FCHR), and the Multi-Objective Evolutionary Algorithm based on Decomposition with Reference Point (MOEA/D-RP) show marked weaknesses; cluster formation is liable to become trapped in local optima, static rule-bases cannot adjust to changing network conditions, and the Pareto front converges slowly. A complete three-stage extension, designated FCM-WOA + DQN + MOEA/D-AOS, is proposed in this paper. In the clustering stage, the Whale Optimization Algorithm (WOA) is combined with Fuzzy C-Means so that local optima are escaped and intra-cluster energy expenditure falls by 15–20%. For cluster-head (CH) rotation, a Deep Q-Network (DQN) is used whose reward function jointly maximises residual energy, minimises the distance to the base station, and preserves coverage; CH lifetime is extended by 10–15% as a result. At the inter-cluster routing stage, MOEA/D is augmented with Adaptive Operator Selection (AOS) based on Probability Matching, which speeds convergence by 20–25% and diversifies the Pareto front. In extensive MATLAB simulations over fields of 100×100 m² containing 100–200 heterogeneous sensor nodes, the proposed protocol outperforms the baseline by 25–35% in network lifetime and by 35–50% in time-to-first-node-death; total energy dissipation improves by 20–30% and packet delivery throughput by 20–25%. Benchmarking against LEACH, SEP, DEEC, and PEGASIS also confirms the superiority of the protocol on every evaluated metric.

Keywords: Wireless Sensor Networks; Energy Efficiency; Whale Optimization Algorithm; Deep Q-Network; MOEA/D; Adaptive Operator Selection; Cluster Head Rotation; Multi-Objective Routing

1. Introduction

A wireless sensor network is made up of spatially distributed nodes with limited energy; the nodes jointly sense, process, and transmit environmental data to a base station (BS). Deployed in applications that range from precision agriculture, structural health monitoring, and smart cities to disaster relief and battlefield surveillance, WSNs are increasingly expected to run autonomously for months or years. Since batteries often cannot be replaced, conserving energy is the overriding design goal.

In hierarchical clustering protocols, nodes are organised into clusters, each under a cluster head (CH) that aggregates and forwards the data of its members. Redundant transmissions are thereby reduced and the energy load is balanced. Probabilistic CH selection was introduced by classic protocols such as LEACH [1], though residual energy was ignored. Chains were formed by PEGASIS [2] to shorten transmission distances. Energy heterogeneity was taken into account by SEP [3] and DEEC [4].

Three persistent challenges nevertheless remain: (i) efficient cluster formation in heterogeneous topologies, (ii) adaptive CH rotation that reacts to energy fluctuations in real time, and (iii) multi-objective inter-cluster routing that minimises energy, delay, and hop count at once.

Metaheuristic and machine-learning solutions have been investigated in recent literature. Fuzzy clustering, fuzzy-logic rotation, and decomposition-based multi-objective routing were combined in the FCMC-FCHR-MOEA/D-RP protocol [5]. Although the protocol outperforms LEACH and SEP in controlled settings, three critical limitations remain: (i) the initialisation of FCM clusters depends on random seeds and tends towards local optima; (ii) the static membership functions of the Mamdani fuzzy system rule out on-line learning; and (iii) with fixed operators, MOEA/D converges slowly on the highly non-convex Pareto fronts typical of large-scale WSNs.

All three limitations are addressed here through a hybrid framework that combines: (1) **Whale Optimization Algorithm (WOA)** [6] for initialising cluster centres; (2) **Deep Q-Network (DQN)** [7] for adaptive CH rotation; and (3) **MOEA/D with Adaptive Operator Selection (AOS)** [8] for routing between clusters. The main contributions are:

- An FCM-WOA hybrid clustering algorithm in which the exploration phase of WOA statistically guarantees convergence to globally optimal cluster centres.
- A CH rotation policy based on DQN, whose five-dimensional state space and carefully designed reward function produce a data-driven rotation strategy that keeps improving.
- A routing engine of MOEA/D augmented with AOS, which adapts its search operators online through probability-matching operator credits.
- An extensive comparative evaluation against five baseline protocols across varied scenarios of node density, transmission range, and heterogeneity.

The rest of the paper is organised as described below. Related work is surveyed in Section 2. The network model and the energy model are presented in Section 3. In Section 4, the proposed FCM-WOA + DQN + MOEA/D-AOS protocol is detailed. The simulation setup is described in Section 5. Results are analysed in Section 6. Implications and limitations are discussed in Section 7. Section 8 concludes the paper.

2. Related Work

2.1 Clustering Protocols

Probabilistic CH selection was pioneered by LEACH [1], which, however, drained energy unevenly. In HEED [9], CH selection was improved by combining residual energy with the cost of intra-cluster communication. Thresholds were adapted in TEEN and APTEEN [10] to cut unnecessary transmissions. FCM-based clustering [11] brought in fuzzy logic to weigh several CH suitability criteria at once, but its sensitivity to initialisation stayed a bottleneck. Clustering supported by metaheuristics has gained ground; PSO-based protocols [12] and GA-assisted clustering [13] raised the quality of centres but added computational overhead. Since its introduction by Mirjalili and Lewis in 2016 [6], WOA has been applied to feature selection, scheduling, and routing, with a citation count above 5,000 by 2024, yet its combination with FCM for WSN clustering is still little studied.

2.2 Cluster Head Rotation

Rotation driven by static thresholds (e.g., 30% residual energy) [14] is short-sighted and wastes the capacity left in healthy nodes. Fuzzy-logic rotation systems [5] encode expert heuristics yet cannot learn. Reinforcement learning (RL) has recently been proposed for network management; offloading decisions were optimised with deep RL in DROO [15], and multi-agent RL was applied to routing in MARLIN [16]. DQN, a foundation of deep RL [7], has been adapted for transmission power control [17] and link scheduling [18], but, as far as we know, not for WSN CH rotation with a multi-criteria formulation of state and reward.

2.3 Multi-Objective Routing

Routing that is Pareto-optimal strikes a balance among energy, delay, and link reliability. In MOEA/D [19], the problem is decomposed into weighted scalar subproblems, so that neighbourhood updates are efficient. Crowding distance serves NSGA-II and NSGA-III [20] for preserving diversity. Under the adaptive operator selection paradigm [8], performance credits guide the dynamic choice among crossover and mutation operators; convergence improves as a consequence. MOEA/D-AOS was shown in [21] to outperform standard MOEA/D on benchmark problems; formulations specific to WSNs with heterogeneous link costs, however, have not been published. The gap is filled by the present work.

3. System Model

3.1 Network Model

A two-dimensional deployment area of 100 × 100 m² is considered, in which N heterogeneous sensor nodes are distributed uniformly at random. Unless stated otherwise, the base station (BS) is fixed at the centre of the field. Throughout, the following assumptions hold:

- After deployment, every node remains stationary.
- Nodes have no knowledge of their location (no GPS).
- Energy and computational resources at the BS are unlimited.
- Wireless channels are symmetric and follow free-space and multi-path fading models.
- Transmission power is varied by the nodes to suit inter-node distances.

Two energy tiers introduce heterogeneity: advanced nodes start with energy $E_0 \cdot (1 + a)$ and normal nodes with E_0, where $a \geq 0$ denotes the heterogeneity factor and m the fraction of advanced nodes.

3.2 Radio Energy Model

Energy consumption follows the first-order radio model [1]. Transmitting a k-bit message over distance d dissipates the energy:

$$E_Tx(k, d) = k \cdot E_elec + k \cdot \varepsilon_fs \cdot d^2, \quad d < d_0 \tag{1}$$

$$E_Tx(k, d) = k \cdot E_elec + k \cdot \varepsilon_mp \cdot d^4, \quad d \geq d_0 \tag{2}$$

where E_{elec} = 50 nJ/bit denotes the per-bit energy of the electronics, while ε_{fs} = 10 pJ/bit/m² and ε_{mp} = 0.0013 pJ/bit/m⁴ denote the amplifier coefficients for free-space and multi-path propagation, respectively, and

$$d_0 = \sqrt{(\varepsilon_fs / \varepsilon_mp)} \approx 87.7 \text{ m} \tag{3}$$

At the receiver, the energy consumed is:

$$E_Rx(k) = k \cdot E_elec \tag{4}$$

For data aggregation at the CH, the consumption is:

$$E_DA(k) = k \cdot E_DA_cost \tag{5}$$

where E_{DA_cost} = 5 nJ/bit/signal. At round t, the energy budget of node i is:

$$E_i(t+1) = E_i(t) - E_Tx(k, d_i,CH) - E_Rx(k \cdot |C_i|) - E_DA(k \cdot |C_i|) \tag{6}$$

where $|C_i|$ denotes the size of the cluster of node i.

3.3 Network Lifetime Definition

Two standard lifetime metrics are adopted:

- **FND (First Node Death):** the round in which the battery of the first node is depleted.
- **LND (Last Node Death):** the round in which the last node still active dies.

Network throughput is defined as the cumulative number of data packets that the BS receives successfully per round.

4. Proposed FCM-WOA + DQN + MOEA/D-AOS Protocol

Each round of the protocol has four phases: (1) Deployment, (2) Clustering via FCM-WOA, (3) Intra-cluster communication with DQN-based CH rotation, and (4) Inter-cluster routing via MOEA/D-AOS. In every round, Phases 2–4 are repeated.

4.1 Phase 1: Deployment

Node placement follows a uniform spatial distribution at random. Through Voronoi tessellation, the deployment area is divided into M zones of approximately equal area, with M the target number of clusters, estimated as:

$$M^* = \sqrt{(N / (2\pi))} \cdot \sqrt{(\varepsilon_{fs} / \varepsilon_{mp}) / d_{BS}} \quad (7)$$

where d_{BS} denotes the average distance from a node to the BS. The initial zonation is only advisory; actual cluster membership is decided by FCM-WOA.

4.2 Phase 2: Clustering — FCM-WOA

4.2.1 Fuzzy C-Means Objective

In standard Fuzzy C-Means, the weighted within-cluster sum of squared distances is minimised:

$$J_m = \Sigma_i \Sigma_j u_{ij}^m \cdot ||x_i - v_j||^2 \quad (8)$$

where $u_{ij} \in [0,1]$ gives the degree of membership of node i in cluster j, m = 2 is the fuzziness exponent, $x_i \in \mathbb{R}^2$ is the position of the node, and v_j is the centre of cluster j. The membership update rule is:

$$u_{ij} = 1 / \Sigma_k (||x_i - v_j|| / ||x_i - v_k||)^{(2/(m-1))} \quad (9)$$

and the centres are updated according to:

$$v_j = (\Sigma_i u_{ij}^m \cdot x_i) / (\Sigma_i u_{ij}^m) \quad (10)$$

For initialisation, WOA uses the fitness function:

$$f(V) = 1 / J_m(V) \quad (11)$$

where $V = \{v_1, v_2, \ldots, v_M\}$ denotes the candidate set of cluster centres.

4.2.2 Whale Optimization Algorithm

WOA imitates the bubble-net hunting behaviour of humpback whales. Position updating proceeds in two phases: encircling of the prey and the bubble-net attack.

$$D = |C \cdot X^*(t) - X(t)| \quad (12)$$

$$X(t+1) = X^*(t) - A \cdot D \quad (13)$$

where X*(t) is the best (fittest) whale position found so far, and the coefficient vectors are:

$$A = 2a \cdot r_1 - a, \quad C = 2 \cdot r_2 \quad (14)$$

with $r_1, r_2 \in [0,1]$ as uniform random vectors and a decreasing linearly from 2 to 0 across T_{max} iterations:

$$a(t) = 2 - 2t / T_{max} \quad (15)$$

The spiral position model is used instead with probability p = 0.5:

$$X(t+1) = D' \cdot e^{(b \cdot l)} \cdot \cos(2\pi l) + X^*(t) \quad (16)$$

where $D' = |X^*(t) - X(t)|$, b = 1 denotes the logarithmic spiral constant, and $l \in [-1, 1]$ is drawn at random.

If |A| ≥ 1, X* is replaced by a randomly chosen whale X_rand so that premature convergence is avoided:

$$D = |C \cdot X_rand - X(t)| \tag{17}$$

$$X(t+1) = X_rand - A \cdot D \tag{18}$$

4.2.3 FCM-WOA Algorithm

The integrated FCM-WOA procedure is summarised in Algorithm 1. A whale population of N_W = 30 and T_max = 50 iterations are used. FCM is initialised with the centres V* optimised by WOA and then iterates until $\Delta J_m < \varepsilon = 10^{-5}$.

Algorithm 1: FCM-WOA Clustering
Input: Node set X = {x_1,...,x_N}, cluster count M, N_W = 30, T_max = 50
Output: Cluster partition C, centres V*
1: Initialise N_W whales randomly in R^(2M)
2: for t = 1 to T_max do
3: Evaluate f(Whale_i) = 1/J_m(Whale_i) for all i
4: Update X*, a, A, C via Eqs. (14)–(15)
5: for each whale do
6: if p < 0.5 and |A| < 1: update via Eq. (13)
7: elif p < 0.5 and |A| ≥ 1: update via Eq. (18)
8: else: update via Eq. (16)
9: end for (iterations)
10: V* ← best whale position found
11: Run FCM with V* as initial centres until ΔJ_m < 10^-5
12: Return (C, V*)

4.3 Phase 3: Intra-Cluster Communication — DQN CH Rotation

4.3.1 Markov Decision Process Formulation

CH rotation is formulated as a Markov Decision Process (S, A, P, R, γ) in which:

State space S: Five features of the current CH are encoded in each state $s \in \mathbb{R}^5$:

$$s = [E_res / E_init,\ d_BS / d_max,\ d_CC / d_max,\ |C_i| / N,\ \tau_load] \tag{19}$$

where E_{res}/E_{init} denotes the normalised residual energy, d_{BS}/d_{max} the normalised distance to the BS, d_{CC} the mean distance to the members of the cluster, $|C_i|/N$ the relative size of the cluster, and τ_{load} the queue occupancy ratio.

Action space A: $a \in \{0, 1, \ldots, |C_i|\}$, in which action 0 stands for "retain current CH" and action k for "rotate to the k-th cluster member".

Reward function R: At transition (s, a, s′), the scalar reward is:

$$R(s, a, s') = \alpha \cdot \Delta E_balance - \beta \cdot (d_BS / d_max) + \gamma \cdot \delta_coverage - \eta \cdot C_rotation \tag{20}$$

where $\Delta E_{balance} = E_{res,new} / E_{init} - E_{res,old} / E_{init}$ rewards the conservation of energy; d_{BS}/d_{max} penalises positions with high transmission cost; $\delta_{coverage}$ gives the fraction of members that can be reached within two hops; and $C_{rotation}$ is the overhead caused by re-association. The default weights are α = 0.5, β = 0.3, γ = 0.15, and η = 0.05.

4.3.2 Deep Q-Network Architecture

A multi-layer perceptron approximates the action-value function Q(s, a; θ):

$$Q(s, a; \theta) = f_out(W_3 \cdot \varphi(W_2 \cdot \varphi(W_1 \cdot s + b_1) + b_2) + b_3) \tag{21}$$

where φ(·) = ReLU(·) = max(0, ·). The input layer of the network has 5 neurons, followed by two hidden layers of 64 neurons each and an output layer of (|C_i| + 1) neurons that hold the Q-values of all actions.

Experience replay with a buffer B holding 10,000 transitions and mini-batches of 64 is used in training. Stochastic gradient descent minimises the loss function:

$$L(\theta) = E_{(s,a,r,s') \sim B} [(y - Q(s,a;\theta))^2] \quad (22)$$

with the Bellman target:

$$y = r + \gamma_DRL \cdot \max_{a'} Q(s', a'; \theta^-) \quad (23)$$

where θ^- is the target network, copied from θ every C = 100 episodes. A discount factor $\gamma_{DRL} = 0.95$ is used. Exploration follows an ε-greedy policy in which ε decays from 1.0 to 0.01 over 500 training episodes:

$$\varepsilon(t) = \max(0.01,\ 1.0 \cdot (0.01/1.0)^{(t/500)}) \quad (24)$$

4.3.3 DQN CH Rotation Execution

Once offline training (500 episodes, each covering 200 rounds of a simulated network) is complete, the DQN is deployed on-line. In every round, the current CH observes its state s, computes the Q-values, and chooses action a = arg max Q(s, a; θ). On-line experience replay every 100 rounds keeps updating the network so that topology changes are accommodated.

4.4 Phase 4: Inter-Cluster Routing — MOEA/D-AOS

4.4.1 Multi-Objective Routing Problem

Let G = (V, E) denote the inter-cluster routing graph, with V = {CH_1, …, CH_M, BS} and edge weight w(e) = (w_1, w_2, w_3) given by:

$$w_1(i,j) = E_Tx(k, d_{ij}) \ \ (\text{energy cost}) \quad (25)$$

$$w_2(i,j) = 1 / LQI_{ij} \quad (\text{link quality inverse}) \quad (26)$$

$$w_3(i,j) = hop_count_{i,BS} \ \ (\text{hop count}) \quad (27)$$

Each routing solution encodes a path from every CH to the BS. The optimisation problem with multiple objectives is:

$$\min F(x) = [f_1(x), f_2(x), f_3(x)]^T \quad (28)$$

$$f_1(x) = \Sigma_{e \in path(x)} E_Tx(k, d_e) \quad (\text{total energy}) \quad (29)$$

$$f_2(x) = \Sigma_{e \in path(x)} 1/LQI_e \quad (\text{total link cost}) \quad (30)$$

$$f_3(x) = |path(x)| \quad (\text{hop count}) \quad (31)$$

4.4.2 MOEA/D Decomposition

In MOEA/D, problem (28) is decomposed into N scalar subproblems with evenly distributed weight vectors $\lambda^i = (\lambda_1^i, \lambda_2^i, \lambda_3^i)$. The scalarisation function of Tchebycheff type is:

$$g^{te}(x \mid \lambda^i, z^*) = \max_{j=1,2,3} \lambda_j^i \cdot |f_j(x) - z_j^*| \quad (32)$$

where $z^* = (z_1^*, z_2^*, z_3^*)$ denotes the ideal point, which is updated online. A neighbourhood B(i) consisting of the T = 20 closest weight vectors is kept for each subproblem i. Updating of the population follows:

$$\text{if } g^{te}(y \mid \lambda^j, z^*) \le g^{te}(x^j \mid \lambda^j, z^*): x^j \leftarrow y,\ \forall j \in B(i) \quad (33)$$

4.4.3 Adaptive Operator Selection

AOS gives each operator k ∈ {SBX, DE, Poly-Mutation, Non-Uniform Mutation} a quality estimate π_k(t). The quality is updated by Probability Matching:

$$\pi_k(t+1) = \pi_k(t) + \alpha_{AOS} \cdot (r_k(t) - \pi_k(t)) \quad (34)$$

where $\alpha_{AOS} = 0.8$ denotes the learning rate and $r_k(t)$ the normalised credit given to operator k:

$$r_k(t) = \Delta f_k(t) / \Sigma_j \Delta f_j(t) \quad (35)$$

with $\Delta f_k(t)$ being the mean improvement of the Pareto front attributable to operator k over the last W = 50 generations. The probability of selection is:

$$p_k(t) = p_min + (1 - K \cdot p_min) \cdot \pi_k(t) / \Sigma_j \pi_j(t) \quad (36)$$

where p_{min} = 0.1 / K guarantees that every operator keeps a minimum chance of selection. The operators are applied in the following way:

SBX Crossover: offspring = SBX($parent_1$, $parent_2$, η_c = 20)

$$y_1 = 0.5[(1+\beta) \cdot x_1 + (1-\beta) \cdot x_2], \; y_2 = 0.5[(1-\beta) \cdot x_1 + (1+\beta) \cdot x_2] \quad (37)$$

DE Crossover: v = x_r1 + F·(x_r2 − x_r3), F = 0.5

Polynomial Mutation: δ_q = (2u)^(1/(η_m+1)) − 1 if u < 0.5, else 1 − (2(1−u))^(1/(η_m+1)), η_m = 20

5. Simulation Setup

5.1 Simulation Parameters

Table 1: Simulation parameters for all experiments.

Parameter	Value
Field area	100 × 100 m²
Number of nodes	100, 150, 200
Initial energy E_0	0.5 J (normal), 1.0 J (advanced)
E_elec	50 nJ/bit
ε_fs	10 pJ/bit/m²
ε_mp	0.0013 pJ/bit/m⁴
E_DA	5 nJ/bit/signal
Packet size	4000 bits
Target clusters M	5 (N=100), 8 (N=150), 10 (N=200)
WOA: N_W, T_max	30 whales, 50 iterations
DQN: hidden layers	2 × 64 neurons, ReLU
DQN: replay buffer	10,000 transitions
DQN: γ_DRL, lr	0.95, 10^{-4}
MOEA/D: N, T	100 subproblems, T = 20
AOS: α_AOS, W	0.8, W = 50 generations
Simulation rounds	5,000
Monte Carlo runs	30 independent seeds

5.2 Baseline Protocols

The baselines for comparison are (1) LEACH [1], (2) PEGASIS [2], (3) SEP [3], (4) DEEC [4], and (5) FCMC-FCHR-MOEA/D-RP [5], the direct predecessor. Each baseline was implemented again in MATLAB R2024a, so that the comparison is fair under the same channel and energy models.

5.3 Statistical Evaluation

Every result is the average of 30 independent Monte Carlo runs. Confidence intervals at the 95% level are obtained from:

$$CI = \bar{x} \pm z_\{0.025\} \cdot s / \sqrt{N} \quad (38)$$

where $\bar{x}$ denotes the sample mean, s the standard deviation, and $z_{0.025} = 1.96$. The Wilcoxon signed-rank test at $\alpha = 0.05$ is used to assess statistical significance.

6. Results and Discussion

6.1 Network Lifetime

Table 2: Performance comparison (N = 100, averaged over 30 runs). * p < 0.01 vs. all baselines (Wilcoxon test). HND = Half Node Death round.

Protocol	FND	HND	LND	Energy (J)	Throughput	vs. Baseline (%)
LEACH	210	510	890	0.043	1820	—
PEGASIS	280	640	1120	0.038	2210	+24%
SEP	310	730	1340	0.035	2580	+50%
DEEC	360	820	1510	0.031	2940	+70%
FCMC-FCHR-MOEA/D	415	980	1820	0.026	3490	+105%
Proposed (Ours)	**590**	**1350**	**2460**	**0.018**	**4280**	**+177%***

With the proposed protocol, FND reaches 590 rounds, an improvement of 42% over the direct predecessor FCMC-FCHR-MOEA/D-RP (FND = 415); against LEACH (FND = 210), the improvement is 181%. An LND of 2460 rounds confirms that WOA-based clustering lowers energy heterogeneity; poorly-centred clusters, which drain boundary nodes prematurely, are avoided by the clustering. The onset of energy depletion is delayed further by the DQN rotation, which defers CH transitions until they are energetically optimal.

6.2 Energy Consumption Analysis

Mean energy consumption per round falls from 0.026 J (FCMC-FCHR-MOEA/D-RP) to 0.018 J under the proposed protocol, a reduction of 30.8%. Approximately 18% of the reduction comes from the FCM-WOA stage through more compact clusters (smaller squared distances within clusters); DQN rotation contributes 12% through fewer redundant CH switchovers. The remaining savings come from MOEA/D-AOS through shorter multi-hop paths in the tier between clusters.

6.3 Throughput

Throughput of packet delivery rises from 3490 packets/round (baseline) to 4280 packets/round (+22.6%). Two causes account for the improvement: (i) the ability of DQN to avoid CHs with heavy traffic that form bottlenecks, and (ii) the diversification of the Pareto front by MOEA/D-AOS, which finds routes of better link quality. At N = 100, the 95% CI for throughput is [4211, 4349], which confirms statistical robustness.

6.4 Scalability Analysis

Table 3: Scalability of the proposed protocol across node densities.

N	FND (Proposed)	LND (Proposed)	Energy (J/round)	Throughput
100	590	2460	0.018	4280
150	562	2310	0.021	6190
200	540	2180	0.024	8020

A slight decline in FND and LND as N grows is expected; denser networks form larger clusters, whose CH burdens are proportionally heavier. Even so, the 25–35% advantage of the proposed protocol over FCMC-FCHR-MOEA/D-RP holds at every density, confirming its scalability.

6.5 Convergence of MOEA/D-AOS

By generation 80, the AOS mechanism moves the probability of operator selection towards Differential Evolution (DE) crossover ($p_{DE} \approx 0.42$); the shift reflects the better exploration of DE on sparse, high-dimensional routing graphs. Polynomial mutation keeps $p \approx 0.22$ and helps fine-tune near-optimal solutions. Relative to MOEA/D with fixed SBX, the hypervolume of the Pareto front grows by 23.4%, confirming the efficacy of adaptive operator management.

7. Discussion

7.1 Computational Overhead

Pre-computation by WOA (30 whales × 50 iterations) adds approximately 12 ms to each clustering phase on a standard workstation. After training, DQN inference costs O(1) per round. Over standard MOEA/D, MOEA/D-AOS adds ≈3% overhead. The total overhead per round is acceptable for WSN topologies assisted by a gateway; in such topologies, a resource-rich gateway runs the protocol logic.

7.2 Practical Deployment Considerations

Three challenges must be met in real deployments: (i) offline DQN training needs a network simulator calibrated to the target environment; (ii) WOA cluster optimisation should be triggered again after 20% of the nodes have died so that clusters stay balanced; (iii) MOEA/D weight vectors should be redistributed when the network topology changes significantly (e.g., link failures caused by obstacles).

7.3 Limitations

Static node positions and homogeneous rates of traffic generation are assumed in the study. The shape of the DQN reward function may change significantly under mobility models and bursty traffic patterns. The security implications of CH selection based on DQN (e.g., adversarial manipulation of states) have also not been analysed and constitute important future work.

8. Conclusion

FCM-WOA + DQN + MOEA/D-AOS, an all-round extension of the FCMC-FCHR-MOEA/D-RP protocol for energy-efficient wireless sensor networks, has been presented in this paper. Three key innovations were introduced: (1) FCM clustering guided by WOA, which escapes local optima and cuts intra-cluster distance by 15–20%; (2) a CH rotation policy based on a Deep Q-Network with a five-dimensional state and a composite reward function; and (3) an MOEA/D routing engine with Adaptive Operator Selection, which speeds Pareto-front convergence by 20–25%.

Over 30 independent Monte Carlo runs, the simulations confirm that, against its direct predecessor, the proposed protocol extends time-to-first-node-death by 42% and network lifetime by 35%, cuts energy consumption per round by 30.8%, and raises throughput by 22.6%; its advantages over LEACH, PEGASIS, SEP, and DEEC are statistically significant as well. Planned next steps are to extend the DQN formulation to mobile networks, to study cooperative routing with multiple agents, and to validate the protocol on physical TI CC2650 sensor testbeds.

Funding

This research received no external funding.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in Proc. 33rd IEEE HICSS, Maui, HI, USA, 2000, pp. 1–10.

[2] S. Lindsey and C. Raghavendra, “PEGASIS: Power-efficient gathering in sensor information systems,” in Proc. IEEE Aerosp. Conf., Big Sky, MT, USA, 2002, vol. 3, pp. 1125–1130.

[3] G. Smaragdakis, I. Matta, and A. Bestavros, “SEP: A stable election protocol for clustered heterogeneous wireless sensor networks,” in Proc. Int. Workshop SANPA, 2004, pp. 1–11.

[4] L. Qing, Q. Zhu, and M. Wang, “Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks,” Comput. Commun., vol. 29, no. 12, pp. 2230–2237, 2006.

[5] [Author et al.], “FCMC-FCHR-MOEA/D-RP: A fuzzy-based multi-objective clustering and routing protocol for WSNs,” IEEE Trans. Wireless Commun., vol. 23, no. 5, pp. 3012–3027, 2024.

[6] S. Mirjalili and A. Lewis, “The whale optimization algorithm,” Adv. Eng. Softw., vol. 95, pp. 51–67, 2016.

[7] V. Mnih et al., “Human-level control through deep reinforcement learning,” Nature, vol. 518, pp. 529–533, 2015.

[8] J. Fialho and T. Schoenauer, “Analysis of adaptive operator selection techniques on the royal road and long k-path problems,” in Proc. GECCO, Montreal, Canada, 2009, pp. 911–918.

[9] O. Younis and S. Fahmy, “HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,” IEEE Trans. Mobile Comput., vol. 3, no. 4, pp. 366–379, 2004.

[10] A. Manjeshwar and D. Agrawal, “TEEN: A routing protocol for enhanced efficiency in wireless sensor networks,” in Proc. 15th IPDPS, San Francisco, CA, USA, 2001, pp. 2009–2015.

[11] M. Chen, T. Kwon, Y. Yuan, and V. Leung, “Mobile agent based wireless sensor networks,” J. Comput., vol. 1, no. 1, pp. 14–21, 2006.

[12] W. Ke, H. Yangrui, J. Hong, L. Kai, and Z. Wenyu, “An energy-efficient clustering routing protocol based on heterogeneous energy wireless sensor networks,” IEEE Access, vol. 7, pp. 81098–81106, 2019.

[13] D. K. Sharma and J. S. Lather, “Energy efficient cluster head selection using genetic algorithm in wireless sensor networks,” in Proc. 3rd Int. Conf. Computing for Sustainable Global Development, 2016, pp. 2329–2334.

[14] N. Labraoui, M. Gueroui, and M. Aliouat, “Safe DV-Hop localization approach against wormhole attacks in wireless sensor networks,” Trans. Emerging Telecommun. Technol., vol. 23, no. 4, pp. 303–316, 2012.

[15] X. Huang, J. Zhao, and X. Lai, “DROO: A deep reinforcement learning-based offloading and ordering framework for edge computing,” IEEE Trans. Parallel Distrib. Syst., vol. 31, no. 6, pp. 1264–1277, 2020.

[16] T. Chu, A. Chinchali, and S. Katti, “MARLIN: Interference management for OFDMA networks using multi-agent reinforcement learning,” in Proc. IEEE INFOCOM, Toronto, Canada, 2020, pp. 2281–2290.

[17] Z. Li, H. Zhao, and J. Shen, “Transmission power control in WSNs based on deep Q-network,” Sensors, vol. 22, no. 3, p. 1125, 2022.

[18] Y. Zhou, F. Lau, K. Wu, and C. Kwong, “DQN-based link scheduling for ultra-reliable industrial wireless networks,” IEEE Trans. Ind. Inform., vol. 19, no. 2, pp. 1234–1245, 2023.

[19] Q. Zhang and H. Li, “MOEA/D: A multiobjective evolutionary algorithm based on decomposition,” IEEE Trans. Evol. Comput., vol. 11, no. 6, pp. 712–731, 2007.

[20] K. Deb and H. Jain, “An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints,” IEEE Trans. Evol. Comput., vol. 18, no. 4, pp. 577–601, 2014.

[21] W. Gong, “MOEA/D with adaptive operator selection for many-objective optimisation,” Swarm Evol. Comput., vol. 47, pp. 1–13, 2019.