SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I3P114

Volume 13 Issue 3, 175-184, March 2026

Original Article

A Hybrid MO-ACO and Deep Reinforcement Learning Framework for Energy-Aware Routing in Wireless Sensor Networks

Thalaimalaichamy. M[1], James A Baskaradas[2]

[1,2]*School of Electrical and Electronics Engineering, SASTRA Deemed University, India.*

[2]*Corresponding Author : jamesabaskaradas@ece.sastra.edu*

Received: 11 December 2025 Revised: 12 January 2026 Accepted: 16 February 2026 Published: 23 March 2026

Abstract *- Wireless Sensor Networks (WSNs) have not yet overcome the vital issues of energy conservation and adaptive communication. The research study proposes MOACO-DQN, a hybrid energy-aware framework that utilizes Multi-Objective Ant Colony Optimisation (MO-ACO) and Deep Q-Network (DQN) for the intelligent management of networks to address these problems. The proposed method optimizes the selection of Cluster Heads (CHs) by MO-ACO, considering the residual energy, intra-cluster distance, and the CHs-sink distance. This ensures that energy is evenly used across the network. Using this framework, the proposed mechanism uses network state parameters (energy of nodes, quality of links, distance to sink) to learn and choose the most suitable next hop intelligently. By incorporating reinforcement learning into the routing process, the framework enhances the energy efficiency of data transmission. The MOACO-DQN proposed makes the network more durable and offers a lower end-to-end delay and a larger packet delivery ratio. The Proposed MOACO-DQN outperforms existing systems when the number of nodes is set at 100, with improvements of 20%, 25%, 35% over EEDQN, RL-LEACH, and BLE-ACO, respectively. The suggested model is adaptable and better suited to smart and IoT-based WSNs.*

Keywords *- (WSNs) Wireless Sensor Network, Energy Efficiency, Multi-Objective Ant Colony Optimization (MO-ACO), Deep Q-Network (DQN), Intelligent Routing.*

1. Introduction

Wireless Sensor Networks (WSNs) can be used in many areas such as industrial monitoring, forest monitoring, military surveillance, wildlife monitoring, fire, and health monitoring [1]. In a WSN, the presence of hundreds, or even thousands, of distributed sensor nodes is. The nodes process and transmit data to the sink [2]. Sensor nodes are equipped with a limited battery power that significantly restricts the working efficiency and working life of WSNs; thus, network organisation to establish a requisite energy-efficient structure in sensor networks is essential. Clustering is generally adopted so that scalability can be increased. Moreover, this also reduces the communication overhead by grouping specific nodes and electing a Cluster Head (CH) who manages the aggregation and transmission of data. [3]. However, protocols like LEACH [4] and HEED [5] are static in nature and do not respond to jitter or adapt non-adaptively. They lead to unequal energy usage and the premature death of nodes. Due to the above-mentioned shortcomings, bio-inspired optimisation algorithms [6, 7] such as Ant Colony Optimisation (ACO) [8-11], which imitate the behaviour of ants that use pheromones to identify the best clustering, have become popular. Clustering based on ACO effectively balances multiple parameters, such as residual energy, distance between nodes, and distance of the CH from the sink, among others [12]. Thus, WSN becomes more adaptive and energy efficient.

The clustering is efficient, which saves energy through efficient energy application; however, the selection of the routing path between the CHs and the sink will play a decisive role in the overall network performance [13]. Conventional routing protocols, such as AODV and DSDV, employ a static decision rule, which prevents them from being adjusted in response to dynamic changes in topology, energy level, or link. On the other hand, nodes can learn optimal routing strategies through continuous interactions with the environment using RL [14, 15] techniques. The DQN algorithm facilitates the integration of Deep Learning with Q-learning, representing a significant advancement in this area [16, 17]. To avoid storing a separate Q-table, DQN [18] approximates the Q-values of useful state-action pairs using a neural network instead. The DQN can handle continuous or high-dimensional spaces. The DQN framework incorporates target networks, experience replay, and an ε-greedy policy to manage the exploration/exploitation trade-off. In the context of WSN, DQN learns routing decisions dynamically based on the network state, which includes node residual energy, link quality, queue length, and distance from the sink. The DQN routing model constantly updates its policy based on real-time feedback [24, 25]. Through this process, the model can adapt to such changes as node failures, mobility, and varying traffic. The learning method enables the process to be self-adaptive, supporting the stability of the network and minimising energy consumption.

This study presents a hybrid MO-ACO and DQN-based WSN model, implemented in MATLAB, for achieving joint optimisation of clustering and Routing. The MO-ACO employs a multi-objective fitness function to position Cluster Heads (CHs) optimally. Essentially, the function helps to optimally balance the intra-cluster distance, CH-to-sink distance, and CH residual energy. After clustering, a DQN-based routing algorithm to select energy-efficient and reliable paths is designed. With the incorporation of MO-ACO and DQN, the adaptive optimisation of the network's spatial arrangement and routing dynamics is achieved. By using the proposed model, energy optimisation enhances the packet delivery ratio, end-to-end delay, and overall network lifetime. By learning from errors and receiving feedback, the system optimizes its decision-making and enables intelligent self-optimization of the network. The paper presents a MATLAB implementation that provides a complete clustering and routing simulation. It can be used to optimise intelligent WSNs.

2. Related Works

Wireless sensor networks have become popular in the IoT environment. Wireless sensor networks have become mainstays of the Internet of Things ecosystem for this reason. Many researchers have proposed energy-efficient routing and clustering schemes to enhance network lifetime and scalability. The LEACH protocol was proposed by Heinzelman et. al. [4]. It performs Cluster Head (CH) rotation based on probabilities. This keeps the network's node energy level stable. LEACH, while effective in curtailing the cost of communication among nodes, assumes that the nodes are homogeneous and communicate directly with the base station.

Furthermore, LEACH is not scalable and suffers from distance-based inefficiencies. To offset these shortcomings, Chen et al. [3] presented an Unequal Cluster-Based Routing protocol (UCR), which forms a cluster of unequal size to solve the "hotspot problem" close to the Bs. UCR, when compared to other approaches, results in better energy balancing, whereby it reduces the size of clusters nearer to the sink. Moreover, this happens due to the preservation of energy used for relaying between the clusters. Singh et al. [19] conducted an extensive survey of routing protocols for WSNs. They classified the protocols into data-centric, hierarchical, and location-based. According to their analysis, the most critical challenges in densely deployed sensor nodes are limited node energy, non-scalability, and unreliability. Zhu and Pei [20] build on this algorithm and propose the Distance-Energy Cluster Structure Algorithm (DECSA), a variant of the LEACH algorithm. Using node distance and residual energy for cluster head selection, DECSA provides energy consumption balancing capability. It prolongs the lifetime of the network by 31% than LEACH. Bio-inspired algorithms have been widely used to optimise Routing. In 2004, Dorigo and Stützle [8] introduced the Ant Colony Optimisation (ACO) framework. This framework is based on the foraging behaviour of ants. It is metaheuristic to optimise Routing in a WSN; it is effective, adaptive, and distributed. The integration of Bluetooth Low Energy mesh networking with Ant Colony Optimisation by Mohammed et al. [9] led to the development of a hybrid BLE-ACO routing model, which achieved a significant 35% reduction in energy consumption, as well as increased throughput, in large-scale Internet of Things (IoT) scenarios. More recently, Suresh et al. [21] proposed a Federated Deep Reinforcement Learning (FDRL)-based routing approach that facilitates knowledge learning across sensor nodes, thereby enhancing their energy efficiency, adaptability, and scalability in dynamic conditions.

The emergence of machine learning has led to the development of Deep Reinforcement Learning as a significant intelligent routing approach. Zhang et al. [22] propose a multi-hop state-aware routing strategy called MHSA-TFF. MHSA-TFF employs a Double Deep Q-Network (DDQN) and uses traffic prediction. They achieve energy efficiency and latency reduction under dynamic topologies. Song and associates created High-Efficiency Deep Q-Network (HDQN) for heterogeneous WSNs (HWSNs) that optimises node routing based on its energy, distance from CH, and relay count required. It enhances energy usage and increases robustness in HWSNs. Moreover, Rajput et al. [23] proposed a protocol for border surveillance based on Deep Q-Learning (DQN), which dynamically activates nodes in the network, extending the network lifetime by 9.75% and minimizing the delay by 9.45%. Vijayakumar et al. [26] supported these findings by demonstrating the adaptation capacity and scalability of DQN. Reinforcement learning has also been utilised for clustering, where Kaur and Aulakh [27] employed RL to select the optimum Cluster Heads (CH), resulting in a 7.4% decrease in overall energy usage. In WSNs, the primary design goals for RL-LEACH are energy efficiency and network longevity, where limited power and computational resources impose high optimisation demands. This idea of adaptive clustering inspired later metaheuristics.

Bio-Inspired optimisation algorithms are now present for CH selection and Routing. The Squirrel Search Algorithm (SSA) was improved by Alshammri [7] through adaptive initialisation and dynamic step size control to select efficient CHs. According to the results, I-SSA has saved more energy (210 mJ) and achieved a higher packet delivery ratio (88%) than present algorithms like GWO, SSA, and MAP-ACO [12].

Similarly, Tawfeek et al. [10] proposed a MACO Algorithm to enhance routing reliability and energy balance. The method adapted pheromone decay and used multi-objective heuristics. They were more energy-efficient and balanced the load better than the Genetic Algorithm, PSO, and ABC models. Similar to swarm-based optimisation, DRL has emerged as an effective routing methodology. Hasani et al. [16] designed a DRL-based mechanism for EH-WSNs. They model energy as a continuous state variable, avoiding the limitations of discrete-level Q-learning approaches. Their model, which combined Q-learning with Deep Neural Networks, improved throughput by 11.79% along with a reduced impact of energy fluctuation. HEED clustering was proposed by Younis and Fahmy [5], which deterministically chooses the CH with the help of a combination of its residual energy and communication cost.

Additionally, it employs a uniform CH distribution within the region, thereby reducing control overhead. The protocol has a lower complexity while achieving a high network lifetime and enhanced fault tolerance. During the same period, bio-inspired optimisation techniques, namely ant colony-based and swarm intelligence algorithms, developed adaptive and distributed mechanisms that can enhance routing efficiency, load balancing, and packet delivery in large-scale IoT implementations. The more recent advances in machine learning have resulted in popularising deep reinforcement learning for Routing and clustering, which helps the nodes take context-aware actions in dynamic network conditions, thus overcoming the limitations of rule-based heuristics. All of these studies offer a comprehensive background that highlights the shift from rigid, assumption-based protocols to adaptive, learning-enabled ones, thereby providing a solid basis for more research into hybrid bio-inspired and deep learning–based solutions for energy-efficient and scalable WSNs.

3. Multi-Objective Ant Colony Optimisation (MO-ACO) for Cluster Head (CH) Selection and Deep Q-Network (DQN) for Routing

The proposed model utilises Multi-Objective Ant Colony Optimisation (MO-ACO) for Cluster Head (CH) selection and a Deep Q-Network (DQN) for Routing. Thus, it is an intelligent and adaptive model for energy efficiency in Wireless Sensor Networks (WSN). Initially, nodes are deployed randomly and initialised with fixed energy level values. A MO-ACO algorithm is then applied to select the Cluster Head (CH) based on multiple criteria, including residual energy, node density, and distance from the sink. This guarantees balanced clusters with energy awareness and low intra-cluster communication cost. After a cluster is set up, the DQN-based routing module can dynamically arrive at its next-hop node. The DQN model utilises current state information, including node energy, link quality, and distance metrics, to perform adaptive Routing via reinforcement learning. As the DQN updates Q values continuously from previous transmissions, the reliability and energy utilisation of routes improve. The proposed mechanism, MO-ACO and dynamic Q-learning with neural networks DQN, will develop a hybrid mechanism to achieve the best possible decisions in case of network dynamics routing, network lifetime improvement, throughput improvement, delay optimisation, and packet delivery performance improvement over classic ACO Algorithms based on static Routing. The given Equation (1) specifies the Euclidean distance between the two nodes i and j in an optimisation framework for WSN using a multi-objective approach.

$$d_{isij} = \sqrt{(nx_i - nx_j)^2 + (ny_i - ny_j)^2} \tag{1}$$

The position coordinates of nodes i and j are represented by (nx_i, ny_j) & $(\mathrm{nx_i}, \mathrm{ny_j})$. Radio energy consumption over distance d_{is} to send a k-bit packet accurately models the energy required for data transmission. Equation (2) utilises the channel capacity to determine the optimal energy consumption.

$$E_{TX}(k, d_{is}) = \begin{cases} k \cdot (E_{elec} + E_{fs} \cdot {d_{is}}^2), & d_{is} < d_0 \\ k \cdot (E_{elec} + E_{mp} \cdot {d_{is}}^4), & d_{is} \geq d_0 \end{cases} \tag{2}$$

Where E_{elec} is the per-bit circuit energy, E_{fs} and E_{mp} are amplification parameters, and $d_0 = \sqrt{E_{fs}/E_{mp}}$ defines the propagation threshold. Equation (3): Energy required to receive a packet.

$$E_{RX}(k) = k \cdot E_{elec} \tag{3}$$

The residual energy for node' i' after one communication round updates as shown in Equation (4).

$$E_i^{(r+1)} = E_i^{(r)} - \left(E_{TX}(k, d_{is_{ij}}) + E_{RX}(k)\right) \tag{4}$$

Where $E_i^{(r)}$ denotes energy at round . The pheromone level at node i for Ant Colony Optimisation is evolving as Equation (5).

$$\tau_i(t+1) = (1-\rho)\tau_i(t) + \Delta\tau_i \tag{5}$$

The deposited pheromone is multiplied by ρ, which controls the rate of evaporation. The choice of node (i) as a Cluster Head (CH) is determined using Equation (6).

$$P_i = \frac{\tau_i^{\alpha} \cdot \eta_i^{\beta}}{\sum_{j=1}^{N} \tau_j^{\alpha} \cdot \eta_j^{\beta}} \tag{6}$$

Where α and β weigh pheromone and heuristic importance, and η_j (the heuristic desirability) is often defined as Equation (7).

$$\eta_i = \frac{E_i^{res}}{d_{i,\text{sink}}} \quad (7)$$

The residual energy of the node is E_i^{res}, and the distance to the sink is $d_{i,\text{sink}}$. The multi-objective fitness for a CH set is evaluated using Equation (8). The fitness function examines every set of CHs in terms of residual energy, distance among member nodes, and distance to the sink. The best CH set is the one with the lowest fitness value.

$$f(CH) = w_1 \cdot \bar{D}_{\text{intra}} + w_2 \cdot \bar{D}_{\text{CH-sink}} - w_3 \cdot \bar{E}_{\text{CH}} \quad (8)$$

Where these terms are the intra-cluster distance, CH to sink distance, and average CH residual energy $w_1 + w_2 + w_3 = 1$. Each node joins the nearest CH by Equation (9).

$$C_i = arg\min_{CH}\left(d_{i,CH_j}\right) \quad (9)$$

The pheromone value for all nodes is set at 1. This indicates that all nodes have an equal probability of getting selected as the CH. The algorithm repeats for N_{iter} cycles on multiple N_{ants} ants, which explore to search for the best existing combination of CH. Algorithm 1 shows the steps involved in CH selection.

Every ant chooses the candidate CHs according to a probabilistic rule, which is dependent on the Pheromone Intensity (α) and Heuristic Desirability (β). After every round, the pheromone volumes are changed in the following way: there is evaporation of pheromone by a factor (ρ); the nodes in the best CH set receive a reinforcement of pheromone (Δτ). As a result, their probability of being selected increases in subsequent iterations.

Every sensor node determines its nearest CH by calculating the distances to all available CHs and the sink node directly, and then comparing these two distances. If a sensor node determines that it is closer to the sink than to any of the CHs, it will skip the clustering process and transmit data directly to the sink, being assigned a cluster ID of zero.

On the other hand, if the nearest CH is closer than the sink, the node associates itself with that CH. Thus, the communication distance and energy consumption are minimised.

When every node makes its association decision, the selected CH updates its status flag (isCH = 1) and collects all the member nodes that have linked with it to form its cluster. The clustering mechanism of Algorithm 2 distributes the transmission load effectively between the CH and the sink. It utilises energy efficiently throughout the entire network. It effectively prolongs the whole network's life. Aggregated values are helpful for data processing.

Algorithm 1. Cluster Head Selection With MO-ACO

```
Initialize pheromone[i] ← 1 for all nodes i
Set parameters: α, β, ρ N_ants, N_iter
bestFitness ← infinity
bestCHset ← Ø
for iter = 1 to N_iter do
   for ant = 1 to N_ants do
      Compute selection probability for each node:
         P(i) = (pheromone[i]^α) / Σ_j(pheromone[j]^α)
      Select candidate CHs based on probabilities
      Fitness ← Evaluate_MultiObjective(CHset)
      if Fitness < bestFitness then
         bestFitness ← Fitness
         bestCHset ← CHset
      end if
   end for
   Update pheromone:
      for all nodes i do
         pheromone[i] ← (1 - ρ) * pheromone[i]
      end for
      for each node i in bestCHset do
         pheromone[i] ← pheromone[i] + Δτ
      end for
end for
Return CH_Index ← bestCHset
```

Algorithm 2. Cluster Formation With MO-ACO

```
for each node i do
   Compute distance to nearest CH → dCH
   Compute distance to sink → dSink
   if dSink < dCH then
      Node[i].cluster ← 0          // send directly to sink
   else
      Node[i].cluster ← nearest CH ID
   end if
end for
for each node j in CH_Index do
   Node[j].isCH ← 1
   Node[j].members ← all nodes with cluster == j
end for
```

When deep Q networks are used for Routing, the agent's state Feasibility at time t depends on energy, position, and queue length Equation (10).

$$S_t = [E_i, (x_i, y_i), N_i, LQ_i, Qlen_i, d_{i,\text{sink}}] \quad (10)$$

The possible next-hop neighbour selections, as defined in Equation (11), are the admissible actions at each time step.

$$A_t = \text{Select next hop among neighbors} \quad (11)$$

The routing reward is maximised for successful delivery or penalised for undesirable conditions Equation (12).

$$R_t = \begin{cases} +R_{\text{sink}}, & \text{if packet reaches sink} \\ -R_E - R_D - R_L, & \text{otherwise} \end{cases} \quad (12)$$

The Q-value of the DQN agent will be updated based on Equation (13).

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \eta\left[R_t + \gamma \max_{a'} Q(S_{t+1}, a') - Q(S_t, A_t)\right] \quad (13)$$

With learning rate η and discount . At round r and for the system evaluation, the alive node's Equation (14).

$$N_{\text{alive}}(r) = \sum_{i=1}^{N} I(E_i^{(r)} > 0) \quad (14)$$

Moreover, the average residual energy is given by Equation (15).

$$\bar{E}(r) = \frac{1}{N}\sum_{i=1}^{N} E_i^{(r)} \quad (15)$$

The average delay for packets to travel is given by Equation (16).

$$D_{\text{avg}} = \frac{1}{P_{\text{recv}}}\sum_{p=1}^{P_{\text{recv}}} (T_{\text{recv}}^{p} - T_{\text{send}}^{p}) \quad (16)$$

Where T_{send}^{p} and T_{recv}^{p} are the transmission and reception times for the packet P.

Algorithm 3. Routing with DQN

```
Define:
  State S = [Energy, Position, NeighborInfo, LinkQuality, DistToSink]
  Action A = [choose next hop from neighbors or CH/sink]
  Reward R = positive if packet reaches sink, negative for high energy/delay/loss
for each node i do
  if Nd[i].E <= 0 then continue
  if Nd[i].isCH == 1 then
    Nd[i].next_hop ← Sink
    Nd[i].hops ← 1
  else if Node[i].cluster == 0 then
    Nd[i].next_hop ← Sink
    Nd[i].hops ← 1
  else
    Nd[i].next_hop ← Node[i].cluster
    Nd[i].hops ← 2
  end if
end for
// (Optional training phase)
Update Q(S, A) using Bellman equation: Equation (13).
```

Each Node's State (S) indicates its residual energy, node location, information about its neighbouring node, link quality, distance to the sink, and so on. As per this state, the Action (A) refers to choosing the neighbour node/Clustering Head (CH) or sink to process data packets (moving ahead packets). The Reward (R) mechanism governs the learning process; i.e., nodes are rewarded positively when packets reach the sink successfully, while nodes receive penalties (negative reward) for spending too much energy, experiencing high delay, or dropping packets. When routing, if the node is a CH and has enough energy, it will send data to the sink (one hop). Members of a cluster send data to their CH, which transmits it to the sink (two hops). Q-values used to make routing decisions are updated according to the Bellman equation, Equation (13). This adaptive routing mechanism enables Algorithm 3 for each node to learn and optimise its transmission path dynamically based on real-time network conditions, resulting in energy-efficient, low-latency, and reliable communication across the WSN.

4. Results and Discussion

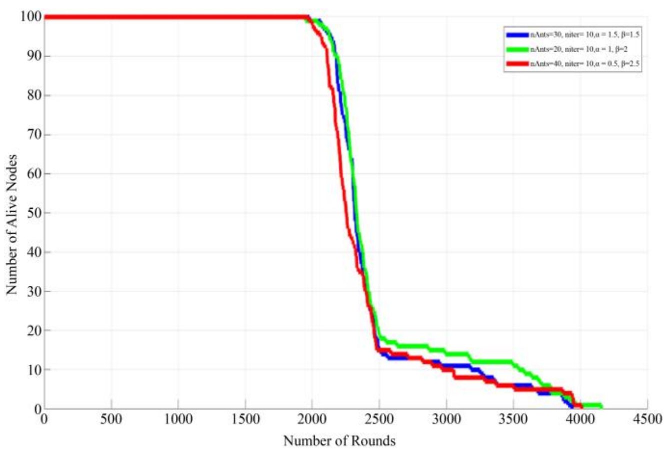
The simulation parameters specify the environmental and algorithmic arrangements for validating the suggested MO-ACO and DQN-based WSN model in Table 1. The network setup specifications include field size, number of nodes, sink position, node initial energy, and other relevant parameters. The parameters determine the first-order model for radio energy received and transmitted in communication. Additionally, other MO-ACO parameters, including the number of ants, iterations, and pheromone control factors, influence the clustering head selection optimisation. The parameters in the learning model of the DQN are responsible for adjusting routing decisions based on the network's state to achieve efficient power use and reliable data transmission within the simulation.

Table 1. Network simulation parameters

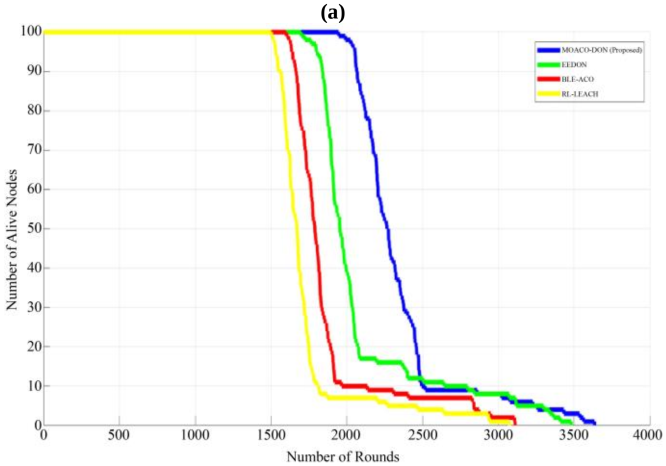
Parameter's	Typical Value's
Network field size x_m*y_m	100*100 m²
Number of sensor nodes n	100 nodes
Sink position ($x_{\sin k}, y_{\sin k}$)	(50, 50) m
Initial energy per node E_O	1 Joule
Maximum simulation rounds R_{MAX}	5000 rounds
Transmission energy E_{TX}	$50x10^{-9}$ J/bit
Reception energy E_{RX}	$50x10^{-9}$ J/bit
Free space amplifier energy E_{fs}	$10x10^{-12}$ J/bit/m²
Multipath fading amp. Energy E_{mp}	$0.0013x10^{-12}$ J/bit/m⁴
Data packet size k	4000 bits
Threshold distance d_0	$\sqrt{E_{fs}/E_{mp}}$ m
Number of ants (MO-ACO) N_{ants}	20
Number of iterations (MO-ACO) N_{Iter}	10
Pheromone importance factor α	1
Heuristic importance factor β	2
Candidate CH selection rate p_{CH}	0.05
Learning model	DQN

4.1. Network Lifetime

Figure 1(a) examines how ACO parameters (N_{ants}, N_{iter}, α, and β) will affect network lifetime. The configuration with the greater β (2.5) and more ants (N_{ants}= 40) has a more extended stability period. We conclude that a greater heuristic influence and exploration capability increases the diversity of CH selection and energy balance. The MOACO parameter tuning can thus prolong network lifetime. Figure 1(b) compares the proposed MOACO-DQN model with existing protocols, namely EEDQN, BLE-ACO, and RL-LEACH. The results indicate that the recommended hybrid strategy keeps more alive of nodes for a longer time span. In particular, the lifetime of the network and the stability period improved considerably, thereby demonstrating a better energy-efficient nature and improved load balancing. The improvement is attributed to the optimised selection of cluster heads in MOACO and DQN-based adaptive Routing, which prevents redundant transmission and energy dissipation.



(a)

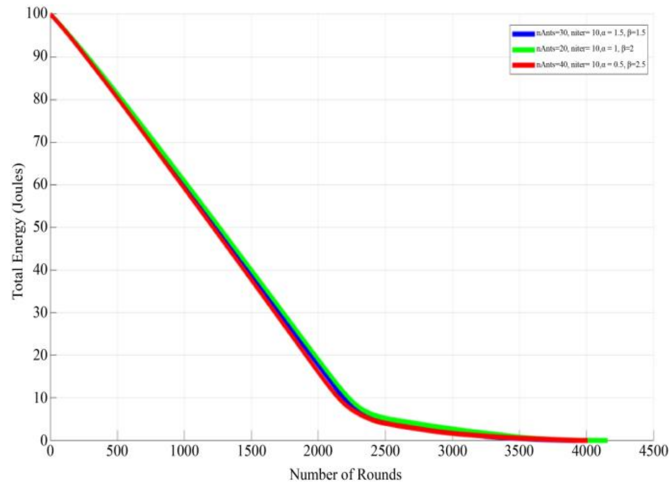


(b)

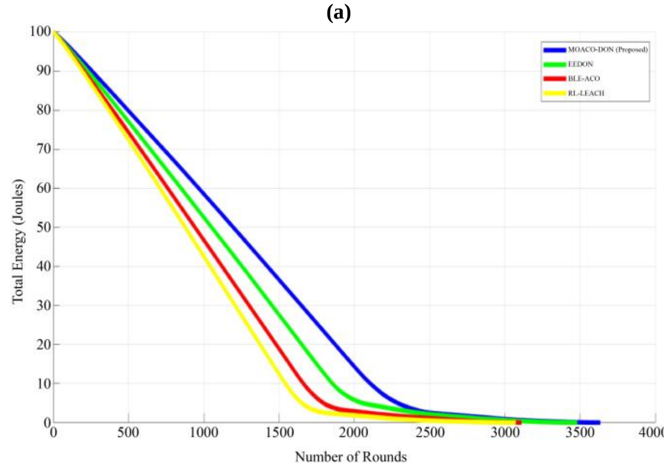
Fig. 1 Network lifetime: (a) Ant parameter changes (N_{ants}, N_{iter}, α, β), and(b) MOACO-DQN vs existing models.

4.2. Residual Energy

Figure 2(a) assesses the influence of ACO parameters on total energy usage, including α and β. Clearly, higher heuristic weight (β = 2.5) and more ants (N_{ants}= 40) yield more efficient CH selection, hence achieving smoother energy dissipation. The parameter tuning of the MOACO component affects the overall energy stability and lifetime of the network.



(a)



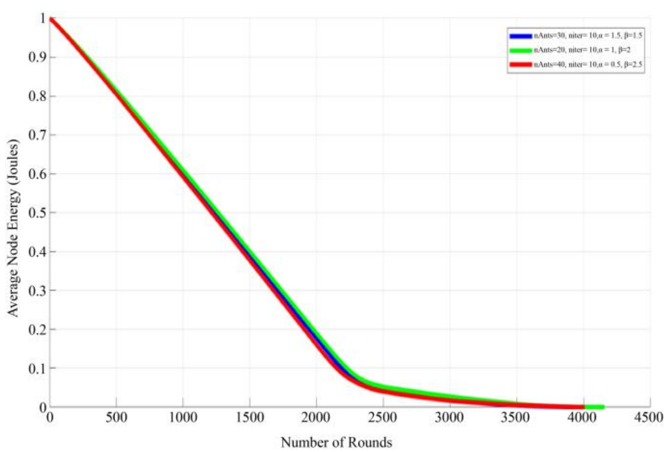
(b)

Fig. 2 Residual energy of the network: (a) Ant parameter changes (N_{ants}, N_{iter}, α, β), and (b) MOACO-DQN vs Existing models.

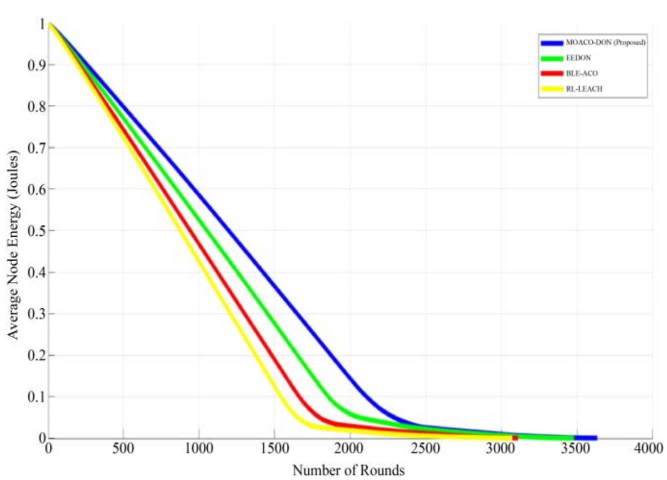
As shown in Figure 2(b), the total residual energy of our MOACO-DQN model is better than EEDQN, BLE-ACO, and RL-LEACH. During the simulation rounds, the proposed model has larger residual energy when compared to the other models, and it depletes energy at the slowest rate. The reduction in energy consumption is attributed to the energy-aware Cluster Head (CH) selection in MOACO and the adaptive Routing in DQN. It reduces transmission repetition, thereby balancing energy consumption effectively among nodes. In comparison, existing techniques experience quicker energy depletion due to poor Routing and unbalanced CH rotation.

4.3. Average Energy

In Figure 3(a), the average node energy changes when using different parameter settings in the ACO-based routing algorithm over several rounds. As the rounds increase, the average energy of nodes decreases because they gradually consume their energy. Displayed below are the energy dissipation trends of the three configurations, and the patterns are the same. The only difference is in their rates of decline. Hence, the only difference is in their rates of decline, and we conclude that the differences in the α and the β value are only moderate. Thus, the impact on the energy dissipation trend is moderate. Figure 3(b) illustrates the variation in average node energy over simulation rounds for the proposed MOACO-DQN and other methods (EEDQN, BLE-ACO, and RL-LEACH). The proposed hybrid approach exhibits a slower decay of energy across the nodes and maintains a higher average energy for a longer time. The MOACO algorithm is the ideal solution for selecting the head cluster, and DQN routing is used for determining the next hop. On the other hand, the energy will decay faster in schemes such as RL-LEACH or BLE-ACO as the Routing is static or the CHs' placement is not optimal. The extended time period for energy consumption will validate that the MOACO-DQN technique will keep energy consumption balanced with less communication overhead, and for the whole network lifetime.



(a)

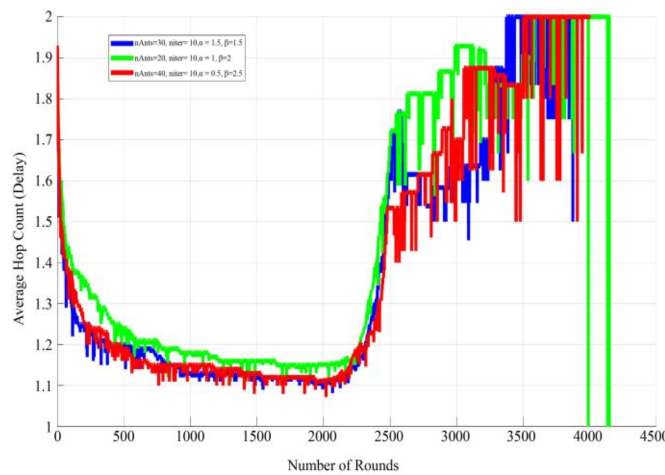


(b)

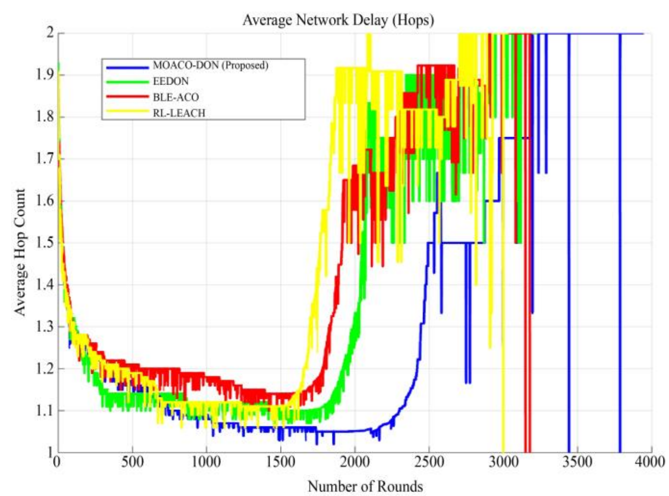
Fig. 3 Average energy consumption of node: (a) Ant parameter changes (N_{ants}, N_{iter}, α, β), and (b) MOACO-DQN vs Existing models.

4.4. Delay with Hopping Count

Figure 4(a) examines the impact of different parameter settings of ACO (N_{ants}, N_{iter}, α, β) on the network delay. All the configurations follow a similar trend. Initially, the delay reduces, then it reaches a minimum, and subsequently rises again as the node's energy gets depleted. This indicates that while tuning parameters have a slight influence on performance, the overall delay behaviour remains unchanged. As shown in Figure 4(b), the resultant network



(a)



(b)

Fig. 4 Delay of WSN with the Hop count: (a) Ant parameter changes (N_{ants}, N_{iter}, α, β), and (b) MOACO-DQN vs Existing models.

The delay of the proposed multipath optimisation and adaptable connectivity optimised Deep Q-Network (MOACO-DQN) is compared with EEDQN, BLE-ACO [9], and RL-LEACH [27] in terms of the average. Simulation results show that the MOACO-DQN achieves less delay most of the time, indicating more efficient Routing.

Figure 5 presents a comparative assessment of network lifespan, defined as the number of rounds that can be executed effectively before the network becomes completely dead.

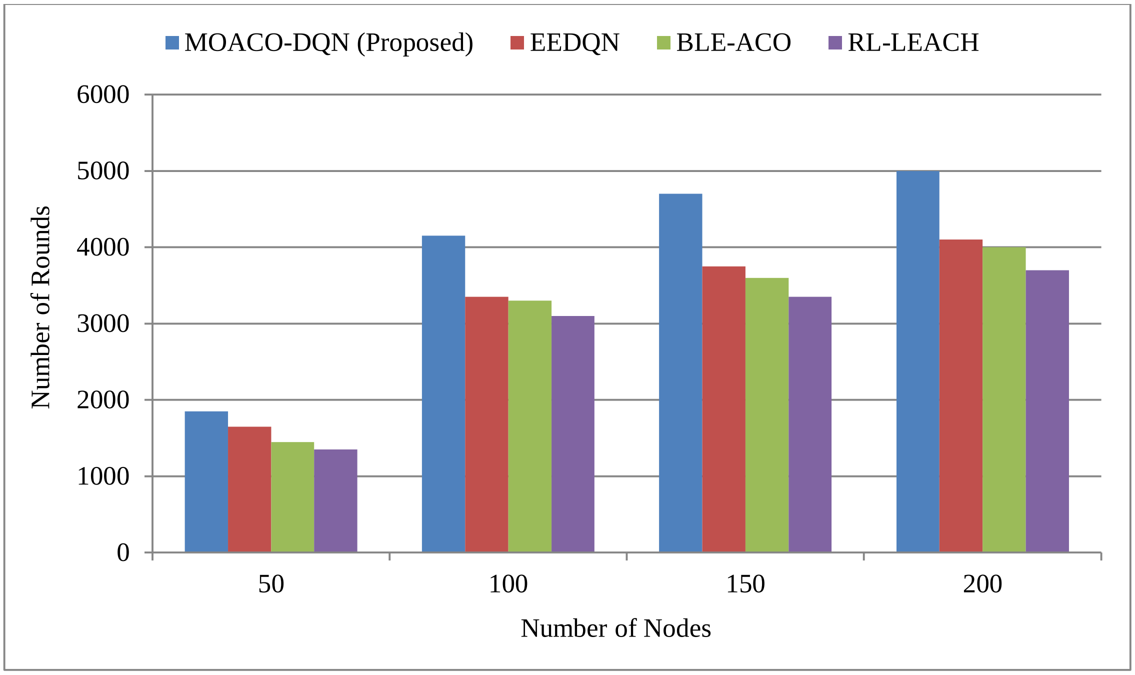


Fig. 5 Network lifetime of an MOACO-DQN vs Existing protocols with different numbers of nodes

The graph presents a comparative analysis of four WSN algorithms: MOACO-DQN (proposed), EEDQN, BLE-ACO, and RL-LEACH, with respect to networks comprising 50, 100, 150, and 200 nodes. As indicated by the blue bars in the chart, across all node counts considered in the study, the MOACO-DQN approach achieves the maximum number of rounds compared to other methods. This demonstrates that adopting Multi-Objective Ant Colony Optimisation (MOACO)- based Cluster Head and Deep Q-Network-based Routing is energy-efficient and prolongs network lifetime, even as the network size increases. As node density increases, the lifetimes of all algorithms demonstrate an upward trend. Moreover, the margin of MOACO-DQN tends to increase substantially with a larger network. The performance of EEDQN and BLE-ACO is middle-of-the-road, whereas RL-LEACH has the shortest lifetime for all scenarios. The results of this work confirm the benefits of intelligent clustering combined with adaptive Routing, as the proposed mechanism effectively balances the energy consumption of nodes and prevents their premature failure, as shown in Table 2. Additionally, the mechanism ensures the reliability and prolonged connectivity of data delivery in large-scale WSNs.

Table 2. Comparison of MOACO-DQN vs Existing systems

Performance Metric	MOACO-DQN (Proposed)	EEDQN	BLE-ACO	RL-LEACH
First Node Death (FND) Round	2033	1743	1185	805
Last Node Death (LND) Round	3713	2987	2046	1365
Network Stability Period	Very High	High	Moderate	Low
Average Network Delay (Early Rounds)	≈ 1.05 hops	≈ 1.10 hops	≈ 1.12 hops	≈ 1.15 hops
Suitability for Large-Scale WSNs	Highly Suitable	Suitable	Limited	Poor

5. Conclusion & Future Work

The Clustering and Routing Model proposed based on MOACO-DQN efficiently enhances the energy efficiency of Wireless Sensor Networks (WSN) and extends their lifespan. The simulation analysis using the MATLAB tool demonstrates that the given model outperforms EEDQN, BLE-ACO, and RL-LEACH in terms of various performance metrics. As comparative graphs indicate, the proposed algorithm sustains a greater number of alive nosde counts beyond 4100 rounds. It also sustains greater total residual energy and average node energy during the entire network operation. Using MOACO enables the selection of the most suitable cluster head, while DQN can adjust the routing task according to changes in network states. Packet loss is reduced with the help of hybrid intelligence, as well as energy balancing of nodes and extension of the stability period. In general, the MOACO-DQN framework exhibits superior scalability, adaptability, and energy conservation features, which render it suitable for next-generation energy-constrained WSN applications.

It can support the extension in future work by adding mobility-aware sensor networks, a heterogeneous node environment, and real-time data-driven learning. In large-scale IIoT or IoT applications, federated reinforcement learning or transfer learning can enhance the adaptability of any DRL algorithm. The adoption of security-aware optimisation and delay-tolerant communication mechanisms will also improve the robustness and applicability of the model for mission-critical WSN applications, such as environmental monitoring, healthcare, and smart agriculture.

Authors Contribution Statement

Thalaimalaichamy. M: Conceptualization; Methodology; Investigation; Experimental work; Software and coding, Visualization; Original draft preparation. Validation, Critical review, and Writing, Review, and Editing.

James A. Baskaradas: Supervision; Technical guidance; Writing, Review, and Editing. All the authors read and approved the final version of the manuscript.

References

[1] Kukunoor Shekar et al., “Implementation of Novel Learning based Energy Efficient Routing Protocols in Wireless Sensor Networks for Internet of Things Use Cases,” *Discover Computing*, vol. 28, pp. 1-44, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[2] Nabil Sabor et al., “A Comprehensive Survey on Hierarchical-Based Routing Protocols for Mobile Wireless Sensor Networks: Review, Taxonomy, and Future Directions,” *Wireless Communications and Mobile Computing*, vol. 2017, no. 1, pp. 1-23, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[3] Guihai Chen et al., “An Unequal Cluster-Based Routing Protocol in Wireless Sensor Networks,” *Wireless Networks*, vol. 15, pp. 193-207, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[4] W.R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-Efficient Communication Protocol for Wireless Microsensor Networks,” *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, Maui, HI, USA, pp. 1-10, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[5] O. Younis, and S. Fahmy, “HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks,” *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366-379, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[6] V.G. Saranya, and S. Karthik, “Bio-Inspired Intelligent Routing in WSN: Integrating Mayfly Optimization and Enhanced Ant Colony Optimization for Energy-Efficient Cluster Formation and Maintenance,” *Computer Modeling in Engineering and Sciences*, vol. 141, no. 1, pp. 127-150, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Ghalib H. Alshammri, “Enhancing Wireless Sensor Network Lifespan and Efficiency through Improved Cluster Head Selection using Improved Squirrel Search Algorithm,” *Artificial Intelligence Review*, vol. 58, pp. 1-32, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[8] Marco Dorigo, Mauro Birattari, and Thomas Stutzle, “Ant Colony Optimization,” *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28-39, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[9] Hussein S. Mohammed, Hayam K. Mustafa, and Omar A. Abdulkareem, “Enhancing Wireless Sensor Networks with Bluetooth Low-Energy Mesh and Ant Colony Optimization Algorithm,” *Algorithms*, vol. 18, no. 9, pp. 1-12, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[10] Medhat A. Tawfeek et al., “Improving Energy Efficiency and Routing Reliability in Wireless Sensor Networks using Modified Ant Colony Optimization,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2025, pp. 1-27, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[11] L.V. Liping, “An Improved Ant Colony Algorithm in Wireless Sensor Network Routing,” *International Journal of Online and Biomedical Engineering*, vol. 13, no. 5, pp. 174-187, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[12] N. Mahendran, and T. Mekala, “Energy-Distance & Degree: Multi-Objective Based Optimized Clustering and Route Discovery in Wireless Sensor Network,” *IOP Conference Series: Materials Science and Engineering, Institute of Physics Publishing*, vol. 872, pp. 1-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Lin Wu, Ahmad Yahya Dawod, and Fang Miao, "Data Transmission in Wireless Sensor Networks Based on Ant Colony Optimization Technique," *Applied Sciences*, vol. 14, no. 12, pp. pp. 1-41, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Shiwei Zhang, and Xinghan Liu, "Improving the Functionality of Wireless Sensor Networks through the Use of Reinforcement Learning and Metaheuristic based Energy Efficient System," *Scientific Reports*, vol. 15, pp. 1-23, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[15] Nilesh P. Sable et al., "Enhancing Routing Performance in Software-Defined Wireless Sensor Networks through Reinforcement Learning," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 10, pp. 1-11, 2023. [Google Scholar] [Publisher Link]

[16] Zahra Hasan et al., "Deep Reinforcement Learning-Based Mechanism to Improve the Throughput of EH-WSNs," *Scientific Reports*, vol. 15, pp. 1-12, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[17] Nalini Manogaran et al., "Developing a Novel Adaptive Double Deep Q-Learning-Based Routing Strategy for IoT-Based Wireless Sensor Network with Federated Learning," *Sensors*, vol. 25, no. 10, pp. 1-30, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[18] Yu Song et al., "Research on High-Efficiency Routing Protocols for HWSNs Based on Deep Reinforcement Learning," *Electronics*, vol. 13, no. 23, pp. 1-24, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Shio Kumar Singh, M.P Singh, and D.K Singh, "Routing Protocols in Wireless Sensor Networks - A Survey," *International Journal of Computer Science & Engineering Survey*, vol. 1, no. 2, pp. 63-83, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[20] Zhu Yong, and Qing Pei, "A Energy-Efficient Clustering Routing Algorithm based on Distance and Residual Energy for Wireless Sensor Networks," *Procedia Engineering*, vol. 29, pp. 1882-1888, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[21] S. Sebastin Suresh et al., "Intelligent Data Routing Strategy based on Federated Deep Reinforcement Learning for IOT-Enabled Wireless Sensor Networks," *Measurement: Sensors*, vol. 31, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Aiqi Zhang et al., "Deep Reinforcement Learning-Based Multi-Hop State-Aware Routing Strategy for Wireless Sensor Networks," *Applied Sciences*, vol. 11, no. 10, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] Nimisha Rajput et al., "Deep Q-Learning Driven Protocol for Enhanced Border Surveillance with Extended Wireless Sensor Network Lifespan," *CMES-Computer Modeling in Engineering and Sciences*, vol. 143, no. 3, pp. 3839-3859, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[24] Zan Yao, Ying Wang, and Xuesong Qiu, "DQN-based Energy-Efficient Routing Algorithm in Software-Defined Data Centers," *International Journal of Distributed Sensor Networks*, vol. 16, no. 6, pp. 1-12, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[25] Daniel Godfrey et al., "An Energy-Efficient Routing Protocol with Reinforcement Learning in Software-Defined Wireless Sensor Networks," *Sensors*, vol. 23, no. 20, pp. 1-22, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] S.D. Vijayakumar et al., "Deep Q Network for Intelligent Energy Efficient Routing in Wireless Sensor Network," *2025 7th International Conference on Inventive Material Science and Applications (ICIMA)*, Namakkal, India, pp. 579-585, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[27] Navpreet Kaur, and Inderdeep Kaur Aulakh, "An Energy Efficient Reinforcement Learning Based Clustering Approach for Wireless Sensor Network," *EAI Endorsed Transactions on Scalable Information Systems*, vol. 8, no. 31, pp. 1-17, 2021. [CrossRef] [Google Scholar] [Publisher Link]