Kufa Journal of Engineering
Vol. 17, No. 3, July 2026, P.P. 406-424
Article history: Received 1 August 2025, last revised 19 September 2025, accepted 23 September 2025

MITIGATING ENERGY DISSIPATION IN WIRELESS SENSOR NETWORKS VIA REINFORCEMENT LEARNING

Awan Nahel Mahmood[1] and Turkan Ahmed Khaleel[2]

[1] Control System Engineering Department, College of Electronic Engineering, Ninevah University, Iraq, Email: awan.mahmood@uoninevah.edu.iq.

[2] Computer Engineering Department, College of Engineering, University of Mosul, Iraq, Email: turkan@uomosul.edu.iq.

https://doi.org/10.30572/2018/KJE/170323

ABSTRACT

The Wireless Sensor Networks (WSNs) are widely used in environmental monitoring, healthcare, and smart farming. However, energy consumption remains critical since battery-powered sensor nodes directly affect network lifetime. The conventional clustering and multi-hop routing algorithms are prone to collapse when used in dynamic environments, resulting in poor energy consumption and frequent node failures. This paper proposes a novel reinforcement learning (RL) routing algorithm based on Q-learning to enhance energy savings in WSNs. The algorithm is dynamic in assigning routes based on node energy, communication distance, and packet size, and adapts in real-time to network changes. It ensures that data transfer is efficient and the load distribution throughout the network is even by updating routing decisions with Q-learning. These simulation outcomes indicate that the suggested algorithm can save up to 25% of energy per round relative to traditional protocols and increase network life to 50 times that of the LEACH protocol.

KEYWORDS

Energy efficiency, consumption, Routing protocols, Reinforcement Learning, Wireless Sensor Networks.

1. INTRODUCTION

In recent years, Wireless Sensor Networks (WSNs) have attracted much attention because of their extensive use in the environment, healthcare, security, and smart agriculture applications. These networks comprise distributed sensor nodes that gather information and transmit it to sink nodes or processing units. Despite their utility, WSNs have significant limitations, particularly energy-related issues, because most of them are battery-operated and have limited lifetimes (Salih and Sulaiman, 2023; Abdulzahra et al., 2023; Srinivasamurthy and Tabassum, 2024). The significance of energy-efficient routing and clustering methods, localization strategies, has also been discussed in several surveys and comparative studies regarding extending network lifetime (Humady and Mahdi, 2023; Al-Janabi, 2022; Al-Shmary and Al-Doski, 2023). To extend node lifetime and maintain the well-being of the network in general, minimization of energy consumption is crucial. Traditional solutions, including clustering-based and multi-hop routing protocols, offer partial solutions but are limited by the fact that they are not adaptive or dynamic (Raj, 2018; Suresh et al., 2022; Begum and Nandury, 2022; Del-Valle-Soto et al., 2020). Such conventional techniques cannot be efficient under these conditions as the networks grow larger and livelier: the node energy is exhausted, the communication system is jammed, and the topology changes. Recent studies have moved towards incorporating Artificial Intelligence (AI) methods, especially reinforcement learning (RL), in routing policies. With RW, nodes can make adaptive decisions based on past interactions and react to current network conditions. In comparison to the static routing, the RL-based protocols have the feature of optimizing the packet forwarding dynamically according to the residual energy, the hop count, and the quality of communication, so that they are ideal in multi-dynamic environments (Schurgers and Srivastava, 2002; Cengiz and Dag, 2018; Cui et al., 2013; Pradeep et al., 2022; Younus and Kocak, 2022). This paper presents a dynamic routing protocol named Q-learning-based RL-EBRP for WSNs. The protocol optimizes routing decisions based on real-time values like residual energy, communication radius, and packet size. RL-EBRP achieves up to 25% power savings per round of communication and load balancing by nearly 50x longer network lifetime than LEACH. These results indicate that reinforcement learning may tremendously increase the sustainability and efficiency of WSNs, as it provides a viable solution to long-term monitoring applications (Liu et al., 2019; Ghaffari, 2014; Subhashini et al., 2023; Abbas et al., 2023; Bains and Sharma, 2012).

2. LITERATURE REVIEW

Wireless Sensor Networks (WSNs) have received much attention because of their application in environmental monitoring, health and wellness, intelligence farming, and military. One of the key issues of such networks is energy conservation because battery power is limited for most sensor nodes, and their lifetime is relatively short. Over the years, various strategies have been suggested to increase the network lifetime, mainly energy-saving routing and clustering methods. Early researchers like Salih and Sulaiman (2023) had a complete assessment of routing protocols regarding throughput and energy usage. Nevertheless, these works were typically not thoughtful of node mobility and scalability and could not be widely applied in dynamic networks.

In the same way, Abuhmida et al. (2015) compared mobile ad hoc routing protocols without considering energy efficiency in the highly dynamic networks. Optimization based on nature, such as bacterial foraging optimization (Abdulzahra et al., 2023) and ant colony optimization (Bains and Sharma, 2012), was also energy-saving, albeit with complicated computations, particularly when used on large scales. Clustering and hierarchical routing protocols are reviewed by Srinivasamurthy and Tabassum (2024) and Al-Janabi (2022) to reduce energy consumption, but most of them do not have a practical implementation in a real network, which restricts the validation of their usefulness. Flat routing methods (Raj, 2018; Suresh et al., 2022) aim at the energy usage on a per-node basis but cannot adjust to changing communication patterns. Redundant transmissions are mitigated by data aggregation techniques (Begum and Nandury, 2022; Abbas et al., 2023), which are yet to be systematically assessed in different network conditions. Sleeping algorithms (Del-Valle-Soto et al., 2020) save on energy, although it does not discuss energy consumption due to routing in dynamic topologies. Multi-hop routing (Cengiz and Dag, 2018) and energy-optimized approaches (Schurgers and Srivastava, 2002; Cui et al., 2013) enhance the effectiveness of packet forwarding but are not flexible to the changes in the network in real time. According to recent studies, Reinforcement Learning (RL) and optimization have been more extensively used in energy-conscious routing. Prithi and Sumathi (2023) put forward frameworks of energy-efficient routing, but they have not been verified practically. Younus and Kocak (2022) mixed grey wolf and dragonfly optimization, which optimizes energy efficiency, and do not compare it to traditional protocols. Liu et al. (2019) and Ghaffari (2014) proposed better routing algorithms, and they need to be evaluated more in realistic network conditions. Routing based on blockchain (Subhashini et al., 2023) has already been implemented to lower energy consumption, but its performance compared to traditional protocols is still to be investigated. The dynamism of routing decisions is made possible by the introduction of Q-learning and RL-based routing protocols (Soltani et al., 2025; Wang et al., 2024; Sachithanandam et al., 2025; Godfrey et al., 2023; Nishat et al., 2025), which allows nodes to dynamically adjust the routing decisions, depending on real-time parameters like residual energy, communication radius, and packet size. Such techniques have demonstrated significant savings in energy and network lifetime and scalability for networks of various sizes. Further examples by Saravanan and Santhosh Babu (2024) and Boonlert et al. (2024) show that further optimization of power consumption and load balancing can be done with the combination of predictive models and RL. Localization and clustering schemes are still crucial in reducing redundant energy consumption and providing network scaling. Combining adaptive routing and energy optimization is necessary in comparative studies of clustering protocols (Humady and Mahdi, 2023; Al-Shamily and Al-Doski, 2023).

In general, the available literature confirms that in the case of wireless sensor networks, there is a plethora of proposed methods to enhance the energy efficiency of this technology; still, several fundamental gaps exist. Most studies are not experimentally verified under real-world networks, whereas comparative analysis with advanced or hybrid protocols is insufficient. In addition, modeling of energy consumption and network adaptivity is also lacking, and issues associated with scalability and mobility in dynamic networks are frequently ignored. These weaknesses highlight the importance of the proposed RL-EBRP protocol, which can overcome such weaknesses by dynamically responding to the real-time network conditions through clustering to reduce energy consumption and perform better than conventional protocols like LEACH and Multi-Hop.

3. CONTRIBUTIONS OF THE CURRENT RESEARCH

In wireless network applications where energy consumption needs to be optimized, this research adds to the creation of reinforcement learning algorithms. It also offers experimental data to advance new strategies in network design and constitutes a viable structure for additional researchers. The findings from this study provide a ground for improving performance in wireless networks, making it easier to extend research in this area.

4. IMPORTANCE OF THE RESEARCH

Wireless Sensor Networks (WSNs) find application in various ways, including environmental, healthcare systems, etc. Such networks comprise sensor nodes that gather and relay information to a central sink node. However, among the most significant issues in WSNs is the short battery life of these sensor nodes. When nodes in the network run out of power, the quality of data transmission within the network is compromised, causing the nodes to collapse earlier than expected and reducing network lifespan. Several traditional routing protocols have been proposed to solve the energy problem in WSNs, but most are highly limited. The disadvantages of conventional techniques are mentioned below:

1- Clustering-Based Routing: In clustering-based protocols, a few sensor nodes are selected as cluster heads, and the remaining sensor nodes form clusters. Although this method decreases the load of communications to individual nodes, the cluster head selection and the asymmetry of energy consumption among the nodes may result in premature depletion of cluster heads and network instability. Moreover, the technique fails to adjust to node energy fluctuations or network variations, resulting in poor energy consumption.

2- Multi-Hop Routing: Multi-hop protocols pass data through various nodes to deliver data to the sink, decreasing the transmission distance in each hop, so the energy allocated to each hop is low. Multi-hop routing is affected by the intermediate nodes' overload since these nodes might be forced to relay information of many nodes, thus disproportionately consuming their energy levels. Moreover, dynamic route selection is not provided, so energy-efficient routes are not always selected, leading to unbalanced energy consumption and congestion in the network.

3- Energy-Aware Routing: Energy-aware routing algorithms attempt to choose routes depending on the node's residual energy. Nevertheless, such protocols continue to use predetermined thresholds and unchangeable decision rules. Their dynamic adaptation in real-time network dynamics is not considered, and they will be unable to adapt to unpredictable situations, such as node failure and network congestion, resulting in sub-optimal energy use and a decreased lifetime.

4- One of the main challenges in Wireless Sensor Networks (WSNs) is the limited energy supply of sensor nodes, which results in early node death and reduced network lifetime. While simple and distributed, conventional protocols such as LEACH often lead to unbalanced energy consumption and do not adapt well to dynamic topology or traffic conditions. To overcome these limitations, we propose a Reinforcement Learning–based Energy Balanced Routing Protocol (RL-EBRP) to optimize routing decisions and dynamically prolong the overall network lifetime.

4.1. System components

The suggested system consists of various essential components that combine to make it an efficient working system. The sensor nodes are widely spread randomly within the monitoring area; they sense the contents of the environment and transmit them to other nodes. Those nodes are selected as cluster heads (CHs) periodically to improve the efficiency of residual energy and distance to the sink. These cluster heads are important in consolidating the data of their member nodes and then transmitting the same onwards. Sensor nodes also communicate with cluster heads; this communication is carried out using routing links, the fundamental data transmission pathways of the network. Lastly, this system has a reinforcement learning (RL) agent (one per node) that actively decides the next-hop forwarding behavior. They make this decision depending on what they have seen at the local point and past experiences learned, so the routing becomes optimum when the network influences them.

4.2. Energy consumption model

A wireless sensor network uses a first-order radio energy model to model the energy consumption of each node in the network (Heinzelman et al., 2002); this has been popular because it is appropriate and straightforward for estimating the energy consumed in the transmission and reception of data packets according to this model:

1- Transmission Energy (E_{tx}): The energy consumed by a node to transmit a k-bit message over a distance d is given by:

$$E_{tx}(k,d) = E_{elec}.k + E_{amp}.k.d^n \quad (1)$$

where:

E_{elec} : is the energy spent per bit in operating the transmitter or receiver circuitry (in nJ/bit),

E_{amp} : is the power lost by the transmit amplifier (in PJ/bit/m^2)

d : is the range of the receiver and the transmitter,

n : is the Path-loss exponent.

2- Reception Energy (E_{rx}): To receive a k-bit message, energy is:

$$E_{rx}(k) = E_{elec}.k \quad (2)$$

In this case, the circuitry energy used is only considered since no amplification is required to receive.

3- Total Node Energy Consumed: A node that is involved in receiving and transmitting a packet consumes the sum of transmission energy plus the reception energy of all packets sent:

$$E_{total} = \sum(E_{tx} + E_{rx} \quad (3)$$

The energy consumption of each node is modeled using the first-order radio energy model. This quantitative model can be used to compare the RL-EBRP with the currently available protocols like LEACH, LEACH-C, and Multi-hop LEACH.

4.3. Reinforcement learning framework

Reinforcement Learning (RL) was chosen as it allows distributed, adaptive, and real-time decision making without global network knowledge. The Q-learning agent on each node in the proposed RL-EBRP has the following components:

State (s): remaining energy of the node, the number of hops until the sink, and the number of neighbors it has.

Action (a): choosing the next-hop node to forward data.

Reward (R): positive reward for energy-efficient forwarding and successful delivery, negative reward for lost packets/too much energy use.

The Q-learning update rule is (Soltani et al., 2025; Wang et al., 2024):

$$Q(s,a) \leftarrow Q(s,a) + \alpha[R + \gamma * max\ a'Q(s',a') - Q(s,a)] \qquad (4)$$

where γ is the discount factor, and α is the learning rate.

Training is online when the network is available.

Time complexity: every decision update is O(1), so it is real-time applicable.

Adaptivity: nodes also report Q-values and hence can be adapted to changing topology or failure of nodes.

Memory usage: a node only holds a small Q-table and the method is therefore practical when only a limited amount of resources is available.

5. METHODOLOGY

The proposed RL-EBRP protocol methodology is shown in Fig. 1, which shows in detail the flow diagram of the operating steps of the algorithm. The first step is network initiation, during which sensor nodes are randomly distributed over the region, each with its own ID, and the distance between nodes is determined. Then every node moves into the data transmission stage, whereas neighbors are chosen depending on the energy levels and the amount of energy needed to transmit. Q-value comparison determines the optimal next-hop node based on the residual energy and the Q-learning update equation. Identification of neighbors and calculating distance is done for all the possible paths, and the Q-table is updated to capture the learned energy-efficient choices. All nodes in each round repeat this neighbor selection/transmission data, rewards calculation, and updating of Q-values cycle. Network statistics, which include alive nodes, consumed energy, and dead nodes, are recorded at the end of each round. Lastly, the results give a broad perspective of the network performance. Such a flow diagram highlights the dynamic, adaptive, and iterative RL-EBRP and lets the network optimize energy usage and make the WSN last longer. The RL-EBRP algorithm has the following steps, which have been drawn in Algorithm 1, in which the energy consumption is defined as in Eqs1 -3 and the Q-learning update rule is used as in Eq.4.

Algorithm: RL-EBRP (Q-learning-based Energy Balanced Routing Protocol)

Step 1. Network Initialization

- Deploy a WSN with n nodes.
- The nodes are all initialized by E0.
- Set Q-values of every node depending on:
    - Residual energy,

- Hop count to the sink,
- Number of neighbors.

Step 2. Set Learning Parameters

- Learning rate (α=1).
- Discount factor (γ=0.8).
- Exploration /exploitation parameters: p = 0.7, q1 = 0.3

Step 3. Iterative Rounds

For each round, do:

Step 3.1. Node Status Update

The parameters of a live node *S* are as follows:

- Residual energy,
- Hop count,
- Packet size.

Step 3.2. Neighbor Discovery

- Determine neighbors in the area of communication R.

Step 3.3. Computation of Energy Consumption.

- The calculation of transmission energy is in Eq.1.
- The reception energy is calculated as shown in Eq.2.
- The overall energy a particular node uses is computed, as shown in Eq.3.

Step 3.4. Reward Calculation

- Calculate the reward r(s, a)
- Favorable incentives for energy-saving forwarding and successful delivery.
- Penalty against wastage of energy or loss of packets.

Step 3.5. Choose a neighbor that has the largest Q-value.

- Q-value Update
- Update the Q-value by using the Q-learning rule in Eq.4.

Step 3.6. Energy Update

- Update the residual energy of each node, after transmission or reception, with the results of Eqs.1-3.

Step 4. Global Metrics Recording, at the end of each round:

- Record the total consumed energy.
- Record the number of alive nodes.
- Record the number of dead nodes.

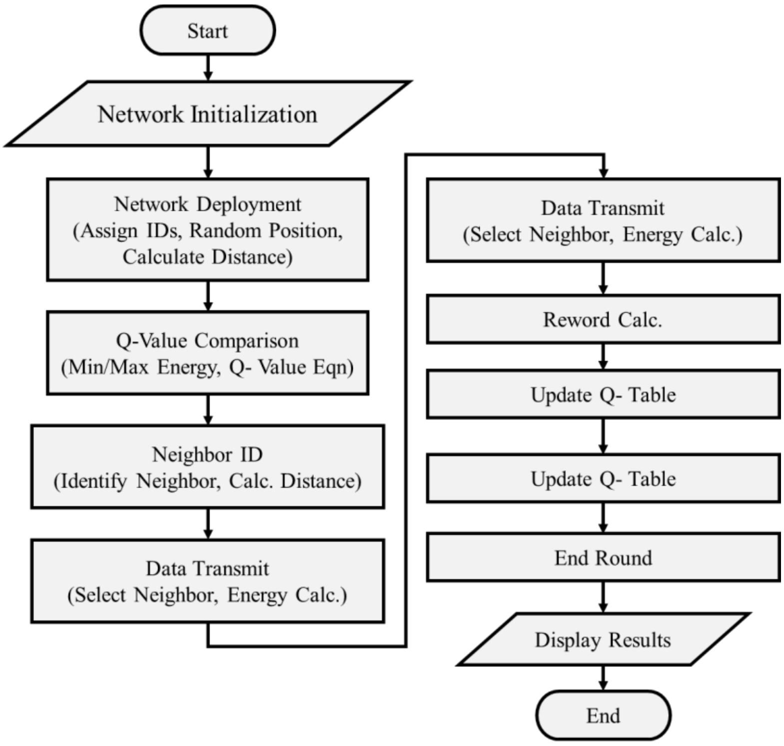

Fig. 1. Methodology block diagram

5.1. Network Setting

The model contains 300 nodes, with the fixed sink node placed at (50, 50) coordinates at 300m x 300m. Every node is assumed to be initialized with an energy of 2 Joules.

5.2. Wireless Sensor Network Formation

In this process, the positions of each node within the specified area are randomly selected, and the functioning status of each node as active or dead. Each node searches for the distance towards the sink and the number of hops needed to reach the sink. The algorithm involves calculating the Q-value at every node so that nodes can decide on energy efficiency using reinforcement learning.

5.3. Algorithm to be used

The node closest to the sink with enough energy to carry out the transmission is chosen in the algorithm. The Q-value is updated using the values from the node's remaining energy, hop count, and distance from the sink. Many rounds are performed to check the maintainability of the network performance when tested under various situations. Thus, the current research offers the concept of an efficient methodology, which, in turn, introduces a model to enhance energy efficiency in wireless networks that has not been proposed before in the field of wireless technology and its applications.

5.4. Improved Basic Reinforcement Learning Algorithm

The proposed reinforcement learning algorithm for energy-efficient routing in wireless sensor networks has many advantages over previous ones. Existing routing algorithms are confined to predetermined rules and manually predefined parameters like the number of hops or energy available at each node, delivering inferior routing selections that fail to change based on the network's dynamicity. On the other hand, the developed algorithm is dynamic in an approach where the reinforcement learning approach is incorporated to update Q-values using feedback from the network in real time. Because of flexibility, the algorithm can make intelligent routing decisions based on energy levels, node capabilities and changes in the physical configuration of the network. In addition to achieving better router operation, the algorithm also increases the level of reliability in the network since it can choose the best routes in case of failures.

Furthermore, while traditional algorithms may prioritize immediate performance metrics, the proposed algorithm integrates a holistic view of the network's state, optimizing for both short-term and long-term energy efficiency. This comprehensive approach addresses existing research gaps by considering the trade-offs between energy consumption and network lifetime, providing a more robust solution to the challenges posed by wireless sensor networks. As a result, the reinforcement learning algorithm contributes to advancing the field by establishing a framework that enhances routing strategies and promotes sustainable network operation.

5.5. Reinforcement Learning Algorithm for Energy-Efficient Routing

The flowchart below is a simplification of the process, which will be used to optimize routing in a Wireless Sensor Network with Q-learning. This flowchart is to plan the necessary steps in the energy-efficient and effective routing of a sensor network using the Q-learning algorithm. The initialization of the network, that is, the establishment of the sensing field, deployment of nodes, and assignment of energy resources, starts the process. After this, each node is assigned Q-values depending on the energy levels, and neighboring nodes are detected. The flowchart is then used to outline how to choose the best data transmission route, where the cost and reward of energy calculations are taken into account. The system repeats several rounds, replenishing the energy of the nodes and saving the results, on which one can analyze further. The above algorithm is a guideline for knowing how Q-learning can improve the routing choice in a wireless sensor network so that the energy used is more efficiently consumed and the network performance is improved.

6. DISCUSSION OF RESULTS

The parameters used to set up the simulation environment are listed in Table 1. Important Construction parameters, that is, the field size, the number of nodes, the initial energy, radio model constants, packet size, and reinforcement learning parameters. These parameters will be the baseline parameters of all three scenarios (100, 150, and 200 nodes).

Table 1: Key Configuration Parameters.

Table 1. Key Configuration Parameters

Parameter	Value
Field Dimensions (xm, ym)	100 meters x 100 meters
Number of Nodes (n)	100, 150, 200
Initial Energy (Eo)	2 Joules
Eelec (Energy Consumption)	50 nJ/bit
Eamp (Transmit Amplifier)	100 pJ/bit/m²
EDA (Data Aggregation)	5 nJ/bit
Packet Size (k)	4000 bits
alpha (Learning Rate)	1
gamma (Discount Factor)	0.8
p, q1 (Probabilistic Parameters)	0.7, 0.3
Communication Range (range_C)	20 meters

In each scenario, the findings are represented in terms of four performance indicators, namely: network topology, dead nodes progression, energy consumption, and operating nodes. To study how the RL-EBRP protocol performs under diverse network densities, the proposed protocol was experimented on three scenarios of deploying it (100, 150, and 200 nodes). The findings offer a thorough insight into the effect of node density on Big Data stability, energy efficiency, and the general lifetime, besides illustrating the soundness of the reinforcement learning-based routing. Fig. 2 demonstrates that the long-term stability of RL-EBRP was highest in the situation of 100 nodes. As illustrated in Fig. 2(a), the topology indicates that there are well-distributed communication paths that are adaptive to reduce localized energy drain. In Fig. 2(b), the mortality curve indicates no failure of nodes before approximately 3600 rounds and complete depletion only after approximately 7200 rounds. The energy consumption was in balance, and the peak was controlled at 4,500-5,200 rounds as illustrated in Fig. 2(c). This trend of the operating nodes depicted in Fig. 2(d) proves that over half of the nodes were operational until approximately the 5200 round, confirming efficient use of energy and long lifetime.

The results in the 150-node scenario are an indication of the impact of higher communication overhead, as shown in Fig. 3. The topology illustrated in Fig. 3(a) shows that RL-EBRP still shared routing load between several paths, which guaranteed connectivity. However, previous deaths of nodes were observed, starting at round number 2200, and the network was exhausted at round number 5200, while Fig. 3(b) illustrated Dead Nodes per Round. Fig. 3(c) illustrates the Energy Consumption per Round for 150 nodes, showing moderate values of energy consumption during periods of rerouting but without harsh imbalances. The curve of the operating nodes illustrated in Fig. 4, Operating Nodes per Round for 150 nodes, was falling sharply after 2200 rounds, and then it steadily fell till the last rounds. Although the lifetime decreased in comparison with the case of 150 nodes, RL-EBRP still had the capacity to delay the fast collapse, and failures were evenly distributed.

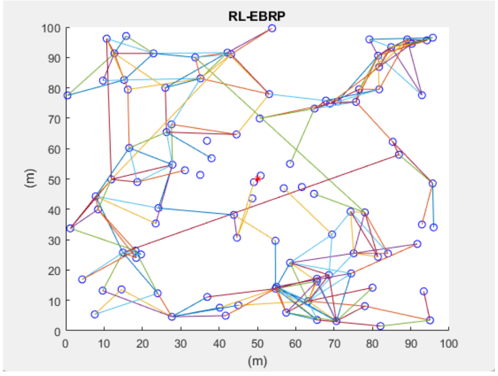

Fig. 2(a). Network Topology with RL-EBRP nodes)

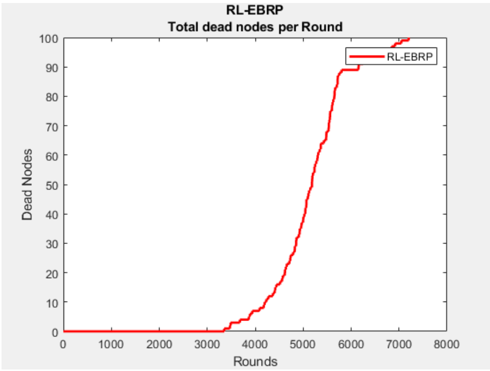

Fig. 2(b) The Dead Nodes in each Round

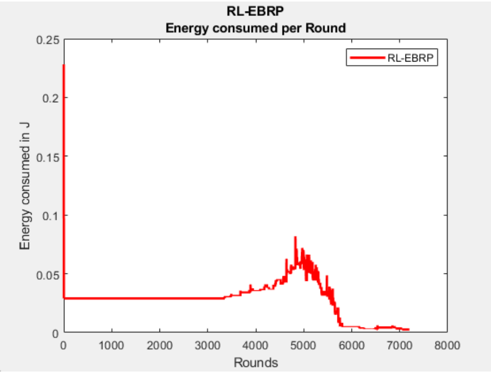

Fig. 2(c) Energy Consumption in each Round

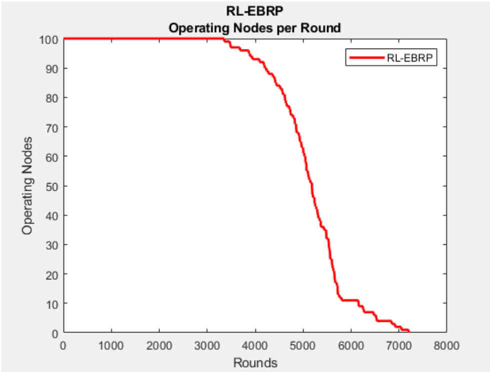

Fig. 2(d) Nodes operating in each Round

Fig. 2. Results for RL-EBRP under 100 Nodes Scenario

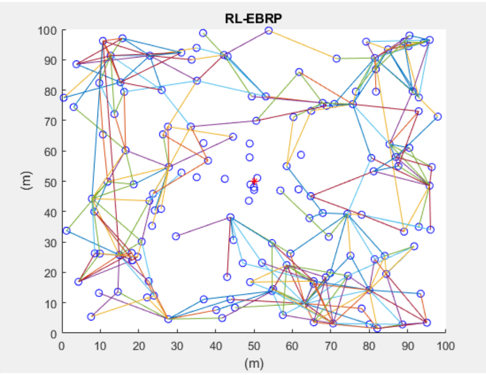

Fig. 3(a) Network Topology with RL-EBRP nodes)

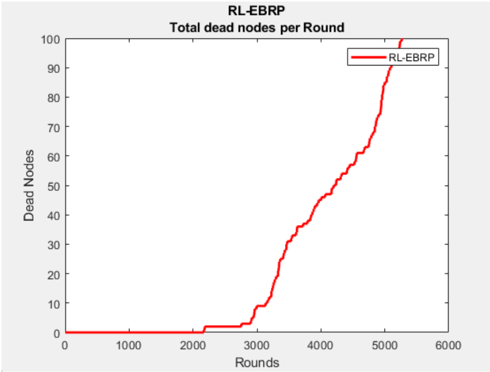

Fig. 3(b) The Dead Nodes in each Round

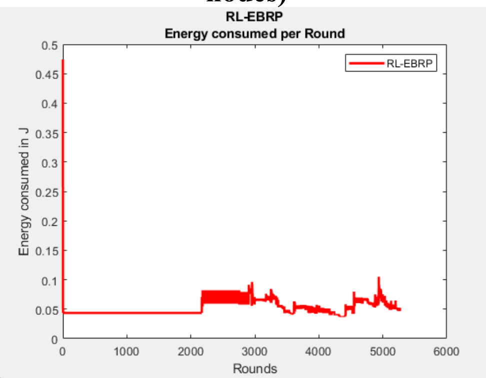

Fig. 3(c) Energy Consumption in each Round

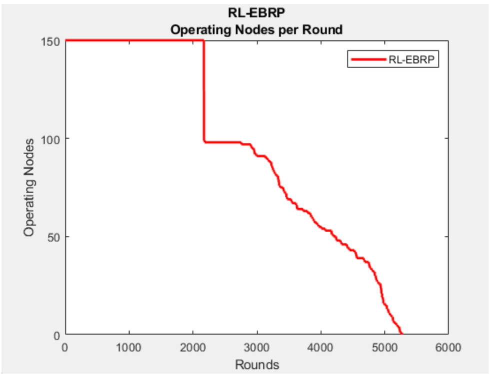

Fig. 3(d) Nodes operating in each Round

Fig. 3. Results for RL-EBRP under 150 Nodes Scenario

The scenario with 200 nodes was the most dense, as shown in Fig. 4. There are dense interconnections, and RL-EBRP is shown to reroute traffic to keep coverage automatically, as shown in Fig. 4(a). The failure of nodes started at approximately 2800 rounds, and the network was fully emptied at approximately 4900 rounds, whereas Fig. 4(b) shows Dead Nodes per Round. The profile of energy consumption by Round, as demonstrated by Fig. 4(c) Energy Consumption per Round, was relatively constant, with small peaks of rerouting during heavy load. Fig. 4(d) Operating Nodes per Round showed that there was an initial steep drop and then a gradual depletion of the operating nodes curve. Although the lifetime was shorter than in the other cases, RL-EBRP managed to prevent catastrophic initial malfunctions and work until the end.

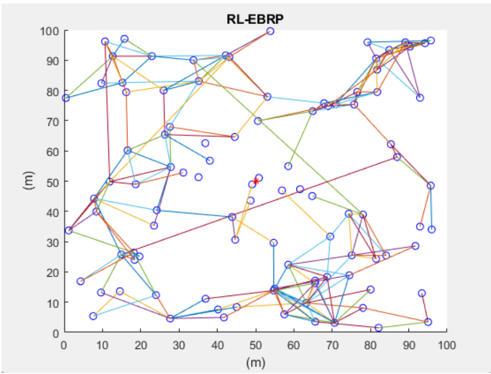

Fig. 4(a) Network Topology with RL-EBRP nodes)

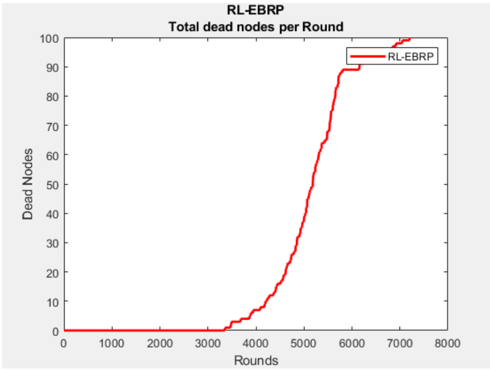

Fig. 4(b) The Dead Nodes in each Round

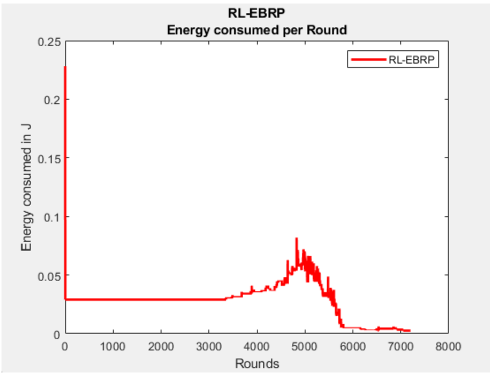

Fig. 4(c) Energy Consumption in each Round

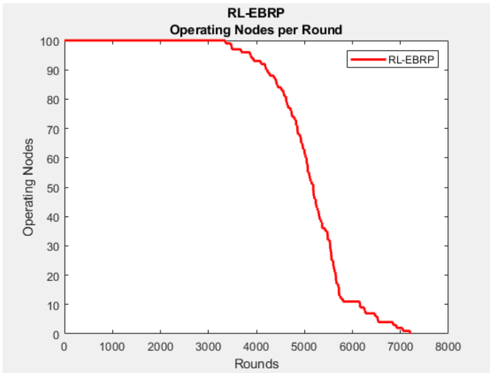

Fig. 4(d) Nodes operating in each Round

Fig. 4. Results for RL-EBRP under 200 Nodes Scenario

In all the scenarios, a common trend is observed whereby the higher the density of nodes, the shorter the stability period and the lifetime, owing to the increased communication requirements. However, in both scenarios, RL-EBRP alleviated energy dissipation through adaptive topology, consumption balancing, and a sudden collapse. The algorithm is thus practical and scalable, and its strengths are adjusted to networks of various sizes.

6.1. Real-Time Data Collection

The proposed protocol dynamically adapts its routing decisions based on real-time network parameters. These parameters are collected as follows:

- Residual energy: Each sensor node updates its remaining energy after every transmission or reception event.
- Packet size: Extracted directly from the packet headers, allowing the algorithm to account for communication costs more precisely.
- Neighbor information: Exchanged periodically through lightweight control packets between neighboring nodes to update local topology and link availability.

This real-time information ensures that routing decisions remain adaptive to the current state of the network, thereby improving energy efficiency and prolonging the system lifetime. The theoretical value of this work is in the use of reinforcement learning, Q-learning, in energy-efficient routing of WSNs. It is a new way to deal with the dynamic and changing nature of real-world sensor networks: by allowing the network to learn and evolve the best routing strategies with time. The study forms a basis for future studies on adaptive, learning-based routing protocols, which would further enhance energy efficiency and robustness in WSNs. To conclude, the suggested approach both makes current protocols better by providing them with the possibility of dynamic optimization and puts a solid theoretical background on the application of reinforcement learning in the field of wireless sensor networks. Further refinement of the reinforcement learning parameters could yield better energy efficiency and performance. According to Table 2 and Table 3, it is possible to understand that the proposed RL-based algorithm will be very efficient, as Table 2 reveals its benefits in terms of energy and network performance over the traditional routing protocols. In contrast, Table 3 demonstrates that the algorithm is more effective than the rest of the existing algorithms in WSNs.

Table 2. Comparison of Energy Efficiency and Network Performance Between the Proposed RL-Based Algorithm and Traditional Routing Protocols

Criterion	Proposed RL-Based Algorithm	LEACH Protocol	Multi-Hop Routing
Energy Efficiency	25% reduction in energy consumption per round	~10% reduction in energy consumption	Variable energy consumption, less efficient
Network Lifetime	Extended lifetime by up to 50% more rounds	Network operational for 500 rounds	Limited network lifetime, often short due to node overload
Load Balancing	Improved by 15%, more balanced energy consumption across nodes	Poor load balancing, cluster head overloads	Energy distribution is uneven, leading to early node failures
Scalability	Handles up to 200 nodes without a significant performance drop	Performance degradation after 100 nodes	Poor scalability beyond 50 nodes

Criterion	Proposed RL-Based Algorithm	LEACH Protocol	Multi-Hop Routing
Adaptability to Network Changes	Dynamic adaptation to node energy and network conditions in real time	Fixed routing decisions, no adaptation	Fixed routes do not adapt dynamically to changes
Number of Dead Nodes	Fewer dead nodes over time, extended node lifetime	Significant dead nodes within the first 500 rounds	High number of dead nodes due to static routing
Communication Range	Adaptive, adjusts based on node positions and energy levels	Fixed at 20 meters for all nodes	Fixed, may not optimize energy use effectively
Performance on Larger Networks	Can scale well with up to 200 nodes, with no significant drop	Significant performance drop beyond 100 nodes	Performance degrades quickly with more nodes

Table 3. Compares the proposed algorithm with other existing algorithms in wireless sensor networks (WSNs).

Criterion	Energy Efficiency	Network Lifetime	Load Balancing	Scalability	Learning Mechanism	Communication Range
Current Research (Proposed Algorithm)	High (e.g., 25% reduction in energy consumption compared to LEACH)	Extended (Network stays operational for up to 50% more rounds than traditional methods)	Reduced energy hotspots (e.g., 15% improvement in load distribution)	High (Can handle 100 – 200 nodes efficiently, no significant performance drop)	Adaptive Q-learning (Q-values dynamically updated in real-time)	Adaptive (20 meters adjusted according to node positions)
LEACH (Heinzelman et al., 2002)	Medium (Energy consumption reduces by ~10%)	Moderate (50% of nodes die within first 500 rounds)	Low (Cluster heads overloaded early)	Moderate (Performance degradation begins at 100 nodes)	None (Fixed cluster head assignment)	Fixed (Range set to 20 meters for all nodes)
Fixed Routing (Del-Valle-Soto et al. 2020)	Low (Energy consumption increases by ~40%)	Short (Nodes deplete energy within first 300 rounds)	Fixed and inefficient (Leads to load imbalances)	Very limited (Performance drops significantly beyond 50 nodes)	None (No dynamic routing adjustments)	Fixed (Range limited to 20 meters)
Probabilistic Routing (Schurgers and Srivastava, 2002)	Medium (Energy consumption improves by ~15% but still suboptimal)	Extended (Network operational for about 1500 rounds)	Moderate (Load distribution not optimized for energy efficiency)	Medium (Scalability limited by network topology)	None (Stochastic routing, no learning-based adaptation)	Stochastic (Communication range may vary with node density)

Criterion	Energy Efficiency	Network Lifetime	Load Balancing	Scalability	Learning Mechanism	Communication Range
Scalability in WSNs (Srinivasa murthy and Tabassum. 2024)	Low (Energy consumption spikes significantly with network expansion)	Limited (Network fails after 500 nodes due to energy imbalance)	Poor (Load balancing not optimized)	Low (Scales poorly beyond 50 nodes)	None (No dynamic scaling approach)	Fixed (Communication range remains static, limiting network expansion)

7. CONCLUSION

This paper has suggested a new RL-based routing algorithm (RL-EBRP) to counter energy waste in Wireless Sensor Networks (WSNs). The protocol uses Q-learning so that routing decisions are dynamically adjusted to node energy, communication range, and packet size, where both efficiency in the short-term and stability in the long-term are obtained. The results of the simulation showed up to 25 percent energy savings per round and 50 percent network life increase in comparison with LEACH. The technique also reduced the speed of death of nodes, maintained energy balance, and maintained connectivity in different node densities (100, 150, 200 nodes). The results validate RL-EBRP in the context of long-term, energy-intensive applications and emphasize its novelty in both its ability to optimize the short-term and offer the long-term network performance simultaneously. Further developments will be directed to making it more scalable to thousands of nodes, using the algorithm in dynamic and mobile WSNs, and incorporating energy harvesting algorithms. In addition, more hybrid models, which provide reinforcement learning together with deep learning, will be considered to render the WSNs more adaptive, resilient, and efficient in a variety of real-world applications.

8. REFERENCES

Abbas, D.T., Hammood, D.A. and Azemi, S.N. (2023) "Minimizing energy consumption based on clustering & data aggregation technique in WSN (MECCLADA)", Journal of Techniques, 5(2), pp. 10–19.

Abdulzahra, A.A., Khudor, B.A.Q. and Alshawi, I.S. (2023) "Energy-efficient routing protocol in wireless sensor networks based on bacterial foraging optimization", Indonesian Journal of Electrical Engineering and Computer Science, 29(2), pp. 911–920.

Abuhmida, M., Radhakrishnan, K. and Wells, I. (2015) "Performance evaluation of mobile ad hoc routing protocols on wireless sensor networks for environmental monitoring", in 17th UKSim-AMSS International Conference on Modelling and Simulation (UKSim).

Al-Janabi, H.A. (2022) "Energy-efficient clustering techniques in wireless sensor networks: A comparative study", Journal of Kufa for Engineering, 10(3), pp. 101–115.

Al-Shimary, R.H. and Al-Doski, S.H. (2023) "Performance evaluation of clustering protocols for energy efficiency in wireless sensor networks", Journal of Kufa for Engineering, 12(2), pp. 55–67.

Bains, V. and Sharma, K. (2012) "Ant colony based routing in wireless sensor networks", International Journal of Electronics and Computer Science Engineering, 1(4), pp. 2516–2524.

Begum, B.A. and Nandury, S.V. (2022) "A survey of data aggregation protocols for energy conservation in WSN and IoT", Wireless Communications and Mobile Computing, 2022, pp. 1–28.

Boonlert, A., Pornavalai, C., Tuwanut, P. and Tanessakulwattana, S. (2024) "EORL: Energy optimization via reinforcement learning in software-defined wireless sensor networks", in Proc. 9th International Conference on Information Technology (INCIT), pp. 618–623.

Cengiz, K. and Dag, T. (2018) "Energy aware multi-hop routing protocol for WSNs", IEEE Access, 6, pp. 2622–2633.

Cui, R., Qu, Z. and Yin, S. (2013) "Energy-efficient routing protocol for energy harvesting wireless sensor network", in 15th IEEE International Conference on Communication Technology.

Del-Valle-Soto, C., Velázquez, R., Valdivia, L.J., Giannoccaro, N.I. and Visconti, P. (2020) "An energy model using sleeping algorithms for wireless sensor networks under proactive and reactive protocols: A performance evaluation", Energies, 13(11), p. 3024.

Ghaffari, A. (2014) "An energy efficient routing protocol for wireless sensor networks using A-star algorithm", Journal of Applied Research and Technology, 12(4), pp. 815–822.

Godfrey, D., Suh, B., Lim, B.H., Lee, K.C. and Kim, K.I. (2023) "An energy-efficient routing protocol with reinforcement learning in software-defined wireless sensor networks", Sensors, 23(20), p. 8435.

Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2002). An Application-Specific Protocol Architecture for Wireless Microsensor Networks. IEEE Transactions on Wireless Communications, 1(4), 660–670.

Humady, Z.A. and Mahdi, M.A. (2023) "A localization in WSN based on area of field and mobility: A survey", Journal of Kufa for Mathematics and Computer, 10(2), pp. 15–28. Available at: https://journal.uokufa.edu.iq/index.php/jkmc/article/download/10750/10729.

Kumar, A. et al. (2024) "Hardware implementation of a reinforcement learning based energy efficient protocol for wireless sensor network", pp. 1–5. Available at: 10.1109/mac61551.2024.10837438.

Liu, Y., Wu, Q., Zhao, T., Tie, Y., Bai, F. and Jin, M. (2019) "An improved energy-efficient routing protocol for wireless sensor networks", Sensors (Basel), 19(20), p. 4579.

Makimaa, Y.P. and Sudarmani, R. (2024) "Secured routing protocol for improving the energy efficiency in WSN applications", IET Circuits, Devices & Systems, 2024(1).

Nishat, T.A.H., Jeong, J.H., Jo, H., Xia, S. and Liu, J. (2025) "Reinforcement learning for adaptive battery management of structural health monitoring IoT sensor network", Applied Energy, 390, p. 125731.

Pradeep, S., Sharma, Y.K., Verma, C., Dalal, S. and Prasad, C. (2022) "Energy efficient routing protocol in novel schemes for performance evaluation", Applied System Innovation, 5(5), p. 101.

Prithi, S. and Sumathi, S. (2023) "Implementation framework for energy efficient routing in wireless sensor network", International Journal of Engineering, Science and Technology, 15(2), pp. 36–47.

Raj, L.D.A.A. (2018) "Analysis on energy efficiency flat routing protocols for wireless sensor communication network", International Journal of Research in Applied Science and Engineering Technology, 6(4), pp. 526–530.

Sachithanandam, V., Jessintha, D., Balaji, V.S. and Manoharan, M. (2025) "Deep reinforcement learning and enhanced optimization for real-time energy management in wireless sensor networks", Sustainable Computing: Informatics and Systems, 45, p. 101071.

Saleem, M.M. and Alabady, S.A. (2023) "Energy-efficient multipath clustering with load balancing routing protocol for wireless multimedia sensor networks", IET Wireless Sensor Systems, 13(3), pp. 104–114.

Salih, M.A. and Sulaiman, D.R. (2023) "Throughput and energy efficiency evaluation of WSN using efficient routing protocols", University of Thi-Qar Journal for Engineering Sciences, 13, pp. 25–33.

Saravanan, G. and Santhosh Babu, A.V. (2024) "Workload prediction for enhancing power efficiency of cloud data centers using optimized self-attention-based progressive generative adversarial network", International Journal of Communication Systems, 37(1), p. e5634.

Schurgers, C. and Srivastava, M.B. (2002) "Energy efficient routing in wireless sensor networks", in 2001 MILCOM Proceedings Communications for Network-Centric Operations: Creating the Information Force (Cat. No.01CH37277).

Soltani, P., Eskandarpour, M., Ahmadizad, A. and Soleimani, H. (2025) "Energy-efficient routing algorithm for wireless sensor networks: A multi-agent reinforcement learning approach", arXiv preprint arXiv:2508.14679.

Srinivasamurthy, P.K.N. and Tabassum, N. (2024) "An extensive study on energy efficient clustering and routing protocols in wireless sensor networks (WSN)", in International Conference on Electronics, Computing, Communication and Control Technology (ICECCC).

Subhashini, G., Rajalakshmi, R., Shriram, S.H., Krishnakumari, L., Chandralekha, R. and Neerathilingam, E. (2023) "Modern WSN routing protocol using methods for reducing energy consumption based on blockchain technology", in International Conference on Recent Advances in Science and Engineering Technology (ICRASET).

Suresh, S., Prabhu, V., Parthasarathy, V., Boddu, R., Sucharitha, Y. and Teshite, G. (2022) "A novel routing protocol for low-energy wireless sensor networks", Journal of Sensors, 2022, pp. 1–8.

Wang, Y., Deng, K., Ren, Y., Chan, J., Tian, G. and Jensen, C.S. (2024) "Reinforcement learning for tackling energy-saving and energy-balance dilemma of cluster-based routing protocols in WSNs", IEEE Access.

Younus, H.A. and Koçak, C. (2022) "Optimized routing by combining grey wolf and dragonfly optimization for energy efficiency in wireless sensor networks", Applied Sciences (Basel), 12(21), p. 10948.