Journal of Intelligent Computing and Networking
https://www.ffspub.com/index.php/jicn/index
ISSN 3079-9228 (print)
ISSN 3107-2798 (online)
E-mail: jicn.office@ffspub.com

Article

A Novel RL-Driven Zone-Based Leader-Aware Energy-Efficient Routing Protocol for Dynamic MANET Environments

Rani Sahu[1,†], Babita Rathore [2]
[1] *Department of Computer Science and Engineering, IES College of Technology, Bhopal 462026, India*
[2] *Govt. MLB Girls P.G. Autonomous College, Indore 452006, India*
[†] *E-mail: rani.princy28@gmail.com*

Received: June 2, 2026 / Revised: June 25, 2026 / Accepted: July 10, 2026 / Published online: July 23, 2026

Abstract: Mobile Ad Hoc Networks (MANETs) face major routing problems because their network structure keeps changing, their energy levels are minimal, their nodes move between locations, and they lack any central network control systems. The traditional routing protocols Ad hoc On-Demand Distance Vector (AODV) and Ad hoc On-Demand Multipath Distance Vector (AOMDV) face major problems in dynamic environments because they use too much energy, their routes fail too often, and their network control demands become too high. The paper introduces RML-ZEREM to solve existing limitations, which functions as a Reinforcement Learning (RL) based Zone-Based Leader-Aware Energy-Efficient Routing Protocol for MANETs. The proposed approach partitions the network into multiple zones and employs energy-aware leader node selection to manage routing operations efficiently. The system uses Q-learning to create an adaptive routing system that chooses the best routing paths according to current network conditions, including residual energy levels, node movement, traffic intensity, and link reliability. The proposed protocol performance assessment uses the NS-2.35 simulator to test different simulation conditions, which include various simulation durations, node mobility rates, network capacity, and simulation area size. The simulation results show that RML-ZEREM achieves better performance than traditional AODV and AOMDV protocols through its ability to increase throughput while decreasing energy usage, improving packet delivery ratio, and reducing routing overhead. The zone-based hierarchical structure enhances network stability and scalability for MANET systems that operate in dynamic environments. The RML-ZEREM protocol functions as an intelligent routing system that adjusts its operations to achieve energy efficiency through its framework, which serves next-generation MANET applications.

Keywords: Reinforcement Learning; Mobile Ad Hoc Networks (MANETs); Energy-Efficient Routing; Q-Learning; Zone-Based Routing; Leader Node Selection
https://doi.org/10.64509/jicn.22.120

1 Introduction

Mobile Ad Hoc Networks (MANETs) operate as decentralized wireless networks without any need for fixed infrastructure because mobile nodes establish connections through multiple transmission hops [1]. The design of MANETs allows each node to act as both a host and a router, which enables users to establish communication links without needing central management. The quick deployment of MANETs and their ability to operate without human guidance make them suitable for various applications, which include disaster recovery operations, military communication systems, emergency response networks, intelligent transportation systems, and Internet of Things (IoT) environments. MANETs experience multiple difficulties, which include dynamic topology changes, limited battery power, constrained bandwidth, and frequent link failures that occur because of node movements [2]. The existing problems result in higher routing overhead and increased packet loss and longer delays, and greater energy usage, which ultimately decrease both network lifespan and communication dependability. The research community requires adaptive, scalable, energy-efficient routing

[†] Corresponding author: Rani Sahu
[*] Academic Editor: Mengwei Xu

protocols for MANET environments because this area remains unsolved. The routing protocols Ad hoc On-Demand Distance Vector (AODV) and Ad hoc On-Demand Multipath Distance Vector (AOMDV) function as standard methods for discovering routes and maintaining connections in MANETs [3]. The protocols deliver sufficient performance in stable network situations, yet they produce excessive control costs when the network operates under highly changing conditions, which leads to increased route discovery work and extra energy needs. The predefined routing methods that these protocols use stop them from making smart decisions needed to adapt to ongoing changes in their network environments. Reinforcement Learning (RL) serves as an effective solution that enables intelligent adaptive routing for MANET networks through its ability to overcome these system constraints [4]. The network environment provides nodes with continuous training, which enables them to develop optimal routing skills through active learning that takes into account residual energy, mobility, traffic load, and link stability. Q-learning stands out as the most appropriate routing method for MANET networks because it enables routing through its model-free learning method and its ability to make decisions based on changing circumstances. The combination of RL with zone-based routing enables better scalability because it allows routing functions to operate in specific areas while decreasing the overall routing difficulties. RML-ZEREM presents a new routing protocol that uses Reinforcement Learning to establish energy-efficient zone-based routing through leader nodes for dynamic MANET networks. Unlike conventional MANET routing protocols and many existing reinforcement learning-based schemes, RML-ZEREM integrates dynamic zone formation, energy-aware leader selection, and adaptive Q-learning-based route optimization into a unified hierarchical routing framework. Many existing protocols tend to zoom in on either energy-friendly routing or on smarter route learning, but they usually don't coordinate leader-driven zone management very well, and they also miss those adaptive routing choices. In contrast, this proposed protocol deals with those gaps by letting route selection become more intelligent through leader nodes, while at the same time trimming routing overhead, keeping energy usage in check, and boosting scalability, particularly when the network is highly dynamic and constantly shifting.

The proposed protocol combines zone-based network organization with energy-aware leader node selection and Q-learning-based adaptive routing decisions. Intelligent agents known as leader nodes have the capacity to choose the best routing paths that adapt to the existing conditions of the network. The proposed approach aims to reduce energy consumption, improve throughput and packet delivery performance, minimize routing overhead, and enhance scalability in highly dynamic MANET scenarios.

1.1 Challenges in MANET Routing

Routing in MANETs depends on multiple essential elements, which include the following factors [5]:

- Energy Constraints: The available battery power restricts network operations, which in turn affects the stability of routing paths.
- High Mobility: The network experiences route failures because of continuous changes in its topology, which results in unstable communication paths.
- Scalability Issues: The routing system becomes more difficult to maintain because the network size increases, which creates additional routing overhead and network congestion problems. The network experiences continuous changes in its topology, which hinders effective route discovery and maintenance processes.
- Limited Bandwidth: The wireless network has restricted bandwidth, which impairs its ability to transmit data reliably and efficiently [6].

1.2 Limitations of Existing Protocols

The current limitations of MANET routing protocols become evident in their operation within dynamic environments [7]:

- The AODV protocol experiences excessive control overhead together with constant route rediscovery problems during periods of high mobility.
- The AOMDV protocol enables multipath routing to enhance system reliability, yet this improvement results in greater routing difficulties, together with increased energy usage requirements.
- The protocols depend on fixed routing methods, which make it impossible to handle network condition changes that occur throughout time.
- Existing reinforcement learning-based routing protocols do help the system act more adaptively in changing situations, but most of them still run into scalability problems; they also bring a high learning overhead, and they tend to miss efficient zone management mechanisms.
- Existing zone-based routing approaches, on the other hand, can cut down routing overhead, but they usually stick to static route selection strategies and don't really use intelligent learning mechanisms, so the network cannot dynamically adjust as conditions evolve.

1.3 Motivation of the Study

The current routing protocols face limitations because they cannot achieve the four goals of energy efficiency, scalability, adaptability, and reliable communication in dynamic MANET environments. Traditional methods tend to either concentrate on improving routing performance or focus on saving energy, which results in an incorrect balance between those two objectives. The intelligent hybrid routing system needs to develop that process because it requires route changes, which result in increased network costs and energy usage throughout its operations. The research aims to create an energy-efficient adaptive routing system through the combination of reinforcement learning and zone-based routing methods. The research will use these combined methods to analyse their effects on MANET throughput, route stability, and network performance.

1.3.1 Research Gap

Even if there has been big progress in MANET routing via energy-aware routing, zone-style routing, and reinforcement learning methods, a few issues still stay unresolved. Traditional routing protocols, like AODV and AOMDV, tend to cause lots of route re-discovery and also use up energy quite fast when the network gets noisy and keeps changing. On the other hand, zone-based routing protocols can effectively reduce routing overhead, but they often lack the intelligent decision-making capability required to adapt to changing network conditions. Reinforcement learning-based routing approaches provide greater adaptability in path selection; however, they may suffer from scalability issues and introduce additional computational and learning overhead. What feels less addressed is the integration of dynamic zone management, an energy-conscious leader selection strategy, and an adaptive reinforcement learning mechanism, all inside one routing framework. Because of these open gaps, the proposed RML-ZEREM protocol is motivated and needs to be built.

1.4 Key Contributions

The major contributions of this research are summarized as follows:

- A novel Reinforcement Learning-based Multi-Leader Zone Routing Protocol (RML-ZEREM) is proposed, which integrates dynamic zone formation, energy-aware leader selection, and adaptive Q-learning-based route optimization within a unified framework. It's a little like the protocol tries to learn from feedback over time, by adjusting the routes using Q-learning.
- The energy-aware leader node selection mechanism enables efficient zone management through its developed system.
- The system uses Q-learning to provide an adaptive routing system that enables intelligent path selection.
- A multi-objective reward function is designed by jointly considering residual energy, packet delivery ratio, link stability, and transmission delay, enabling more intelligent and reliable routing decisions.
- The research team assessed their proposed protocol by testing it against AODV and AOMDV through various simulation scenarios.
- A dynamic zone maintenance and cross-zone communication mechanism is brought in to push routing scalability and adaptability forward, especially in highly mobile MANET settings where things change quickly.

There have been quite a few routing protocols proposed in the literature that try to make MANET performance better, often by mixing reinforcement learning, zone based routing ideas, and some energy aware mechanisms, or at least that's the general direction. QLRP uses reinforcement learning for more adaptive route selection, though it doesn't really do zone formation, and it also doesn't include leader based coordination, plus the energy-awareness side is somewhat constrained. Likewise, QMR also relies on reinforcement learning and looks at energy related metrics when choosing routes, but it still misses both the zone structured network organization and any leader node management. The ZBLE protocol, on the other hand, brings in a zone based architecture and energy-aware leader selection, which helps with scalability and energy efficiency, but it does not add reinforcement learning for routing choices that change over time. RL-ZRP blends reinforcement learning with zone-based routing so it can be more adaptive and scalable, however it does not feature an energy-aware leader election mechanism. In contrast, the RML-ZEREM protocol that is being proposed here ties together reinforcement learning, dynamic zone formation, energy-aware leader selection, and adaptive route optimization inside one unified hierarchical framework. Because it jointly accounts for residual energy, link stability, traffic load, plus overall network dynamics during the routing decisions, RML-ZEREM ends up giving a broader approach for energy efficient, scalable, and kind of intelligent routing in very dynamic MANET setups.

The proposed RML-ZEREM protocol is the sole scheme capable of simultaneously supporting energy-aware routing, reinforcement learning adaptation, dynamic zone management, leader-assisted communication, and adaptive route optimization. These bundled characteristics are what make it different from the other routing schemes, and they also become the groundwork for the better performance we see when the network changes often. In order to make sure the proposed protocol actually works well, we run extensive simulations across multiple MANET scenarios. This includes different simulation durations, node mobility levels, network size, and the deployment area, too. The results for RML-ZEREM are judged based on energy consumption, throughput, and packet delivery ratio. Then we compare those outcomes against the benchmark routing protocols AODV and AOMDV.

1.5 Organization of the Paper

The subsequent sections of this paper will follow this specific organizational structure. The literature review of MANET routing and reinforcement learning-based approaches is presented in Section 2. The proposed RML-ZEREM methodology is described in Section 3. The mathematical modelling and Q-learning formulation appear in Section 4. The simulation setup and performance metrics are discussed in Section 5. The simulation results and performance comparison results appear in Section 6. The paper ends with Section 7, which presents research directions for future studies.

2 Literature Review

The research on Mobile Ad Hoc Network (MANET) routing efficiency has researched various solutions, which include handling dynamic network changes, operational energy restrictions, user movement patterns, and system capacity requirements [8]. The existing routing methods can be divided into five categories, which include traditional routing protocols and energy-efficient techniques, zone-based methods, reinforcement learning-based routing, and hybrid intelligent routing approaches [9–13].

2.1 Traditional Routing Protocols in MANETs

The traditional MANET routing protocols consist of three different types, which are proactive, reactive, and hybrid routing methods [2, 14–16]. The protocols Destination-Sequenced Distance Vector (DSDV), Dynamic Source Routing (DSR), and Ad hoc On-Demand Distance Vector (AODV) serve as popular methods for discovering and maintaining routes [17–19]. AODV establishes routes only when needed to decrease routing overhead, but the protocol experiences frequent route failures and high control overhead during mobile operations [7, 19–21].

Ad hoc On-Demand Multipath Distance Vector (AOMDV) enhances routing reliability by maintaining multiple alternative paths between source and destination nodes [22]. The use of multipath routing increases system reliability during failures, but this method creates additional routing challenges and control packet requirements, and higher energy needs. The conventional routing protocols of this system do not possess smart adaptability features, which leads to performance drops when the network experiences rapid changes.

2.2 Energy-Efficient Routing Techniques

The development of energy-efficient routing protocols aims to enhance network lifetime while decreasing energy waste in MANETs [23]. The routing methodologies use Minimum Total Transmission Power Routing (MTPR) and Minimum Battery Cost Routing (MBCR) to determine the optimal route through transmission power and node energy levels. The system boosts energy efficiency by selecting routes that avoid nodes with limited battery capacity while reducing transmission expenses. The majority of energy-aware routing protocols use unchanging routing metrics, which fail to handle changes in network topology, traffic movement, and user mobility. The performance of these systems declines in actual MANET environments, which experience constant alterations in network conditions.

2.3 Zone-Based Routing Approaches

The zone-based routing protocols create multiple logical regions throughout the network, which helps to enhance scalability while decreasing routing expenses [24, 25]. The Zone Routing Protocol (ZRP) serves as the most popular zone-based solution because it uses proactive routing for zones and reactive routing for inter-zone communication. The structure enables the system to reduce global route discovery expenses while bettering resource usage [26].

Zone-based protocols encounter difficulties that involve creating effective zone structures, choosing leader nodes, and managing inter-zone communication [27]. The existing solutions fail to implement advanced adaptive systems that can react to variations in node movement patterns, energy consumption, and network user density.

2.4 Reinforcement Learning-Based Routing

The solution that uses Reinforcement Learning (RL) for routing purposes provides an effective method to create intelligent routing systems, which can adapt to changes in Mobile Ad Hoc Networks (MANETs) [28]. The network environment allows RL to provide nodes with the possibility to discover the best available routing methods. Q-learning serves as the most commonly used RL method because it enables learning without models and permits dynamic selection of decisions by users. The RL routing protocols use four different parameters, which include delay, residual energy, link stability, and packet delivery ratio, to achieve their best routing results. The solutions enable systems to adapt their routing methods more effectively when network conditions experience changes. The RL-based methods encounter multiple obstacles because they take extended periods to reach their goals while requiring more processing power and making it hard to create working reward systems. The current RL methods of operation fail to connect the network framework with its capacity expansion functions.

2.5 Hybrid Approaches

Recent research has focused on hybrid routing approaches that combine zone-based routing, reinforcement learning, and energy-aware mechanisms to improve overall MANET performance [25, 29]. The methods use leader and cluster nodes as intelligent agents who perform routine tasks together with other decision-making responsibilities. The combination of learning-based routing and hierarchical network structure results in improved scalability, throughput, and energy efficiency for hybrid systems. The current hybrid approaches face several challenges because they need better methods for selecting leader nodes, which groups use to identify network leaders, and they need more effective methods for adjusting reward functions, and they encounter higher computational demands, and their ability to function in fast-changing network environments remains restricted. The system needs additional enhancements to reach its goals of achieving energy-efficient operation, which can expand to larger systems while maintaining dependable routing performance.

To better understand the strengths and limitations of current MANET routing approaches, a comparative look at traditional, reinforcement learning-based, zone-based, and hybrid routing techniques is given in Table 1. This comparison highlights the key characteristics of each approach in terms of routing intelligence, scalability, energy efficiency, and leader-assisted communication. From this, it becomes easier to identify the weak spots of existing solutions and, in turn, provides the rationale for building the RML-ZEREM protocol proposed here.

As shown in Table 1, traditional routing protocols like AODV and AOMDV do not really offer intelligent adaptive routing capabilities. Reinforcement learning-based approaches can enhance routing adaptability; however, they often face scalability challenges and incur additional computational overhead. Zone-based routing protocols tend to improve scalability, though they usually depend on static routing decisions, and they don't include much in the way of learning mechanisms. The existing hybrid methods do try to partially cover these things, but they still give only modest support for coordinated leader selection and adaptive energy-aware routing. These research gaps motivate the development of the proposed RML-ZEREM framework, which integrates reinforcement learning, dynamic zone management, and energy-aware leader selection into a unified routing architecture.

2.6 Research Gaps Identified

The literature review discovered these research gaps, which need to be studied further:

Table 1: Comparative Analysis of Existing MANET Routing Approaches

Author/Protocol	RL	Zone Based	Energy Aware	Leader Selection	Limitation
AODV	No	No	No	No	High routing overhead
AOMDV	No	No	Partial	No	Increased energy consumption
RL-Based Routing	Yes	No	Partial	No	High learning overhead
ZRP	No	Yes	No	No	Static route selection
Existing Hybrid Methods	Yes	Yes	Partial	Partial	Scalability issues
Proposed RML-ZEREM	Yes	Yes	Yes	Yes	–

- Adaptive energy-aware routing systems need development that can operate effectively in dynamic MANET environments.
- The current integration methods between reinforcement learning and scalable network structures show major performance issues.
- The current methods for selecting leader nodes lead to routing problems because they reduce energy efficiency.
- The design of reward functions needs improvement to create a system that optimally balances routing performance.
- Current intelligent routing systems face high routing costs together with operational difficulties.

The RML-ZEREM protocol integrates reinforcement learning with zone-based routing and energy-aware leader selection to create a routing system that adapts to changes while maintaining energy efficiency and network expansion in dynamic MANET environments. Even if a lot of progress has been made in MANET routing research, some important problems still stay unsolved, or at least not fully resolved. Most existing routing protocols focus on routing efficiency, energy savings, or routing adaptability, but rarely on all three. Honestly, very few methods manage to work on intelligent route learning, scalable zone management, and energy-aware leader coordination, all inside one routing idea. And beyond that, reinforcement learning-based solutions usually entail high computational complexity, while the more traditional zone-style approaches lack adaptive decision-making capabilities. So, in the end, there is still a clear need for a routing protocol that can, more or less, "learn" how to adapt to changing network conditions while also staying scalable, energy efficient, and still able to deliver packets reliably.

To respond to these research gaps, this study introduces the Reinforcement Learning-based Multi-Leader Zone Routing for Energy-Efficient MANETs (RML-ZEREM). The proposed protocol merges dynamic zone formation with energy-aware leader selection, then uses adaptive Q-learning routing decisions, plus efficient inter-zone communication. The goal is to boost network lifetime, increase throughput, improve packet delivery performance, and keep routing scalability strong in highly dynamic MANET settings where things change fast and often.

3 Proposed Methodology: RML-ZEREM Protocol

This section introduces the RML-ZEREM protocol, which uses reinforcement learning to create a zone-based energy-efficient routing system for Mobile Ad Hoc Networks. The proposed framework uses zone-based hierarchical routing and reinforcement learning to create an energy-efficient communication system that adapts to changing network conditions in highly active environments. The protocol uses intelligent routing decisions together with energy-aware network organization to solve the problems found in traditional MANET routing methods.

3.1 Overview of the Proposed Framework

The proposed RML-ZEREM framework divides the network into several logical zones which each contain a leader node that operates as a reinforcement learning agent. The routing process uses Q-learning-based adaptive decision-making to allow the protocol to choose optimal routing paths that match the present network conditions. The proposed protocol shows its complete workflow through its three main operational stages, which include network initialization and zone formation, energy-aware leader node selection and Q-learning-based route discovery and data forwarding, and continuous learning and route adaptation. The proposed framework requires leader nodes to conduct ongoing surveillance of network parameters, which include residual energy, node mobility, traffic load, and link quality. The Q-learning mechanism selects next-hop nodes for packet forwarding based on their current status, which includes network parameters. A multi-parameter reward function evaluates routing performance using metrics such as energy consumption, delay, packet delivery ratio, and link stability. The routing agent develops optimal routing strategies through its active network environment interactions, which result in improved network performance that builds up over time. The RML-ZEREM framework shown in Figure 1 displays its ability to conduct hierarchical routing in MANETs through its three components, which include zone creation, selection of energy-efficient leader nodes, and data transmission optimization using reinforcement learning methods.

3.2 Network Model and Assumptions

The MANET is represented as a graph $G = (N, L)$ which uses N to define the mobile node set and L to show node communication links. The network allows nodes to communicate through multi-hop wireless transmission, which operates in a decentralized mode without any fixed infrastructure.

The proposed network model uses the following assumptions for its implementation:

- The network area contains randomly distributed nodes that spread throughout its entire space.

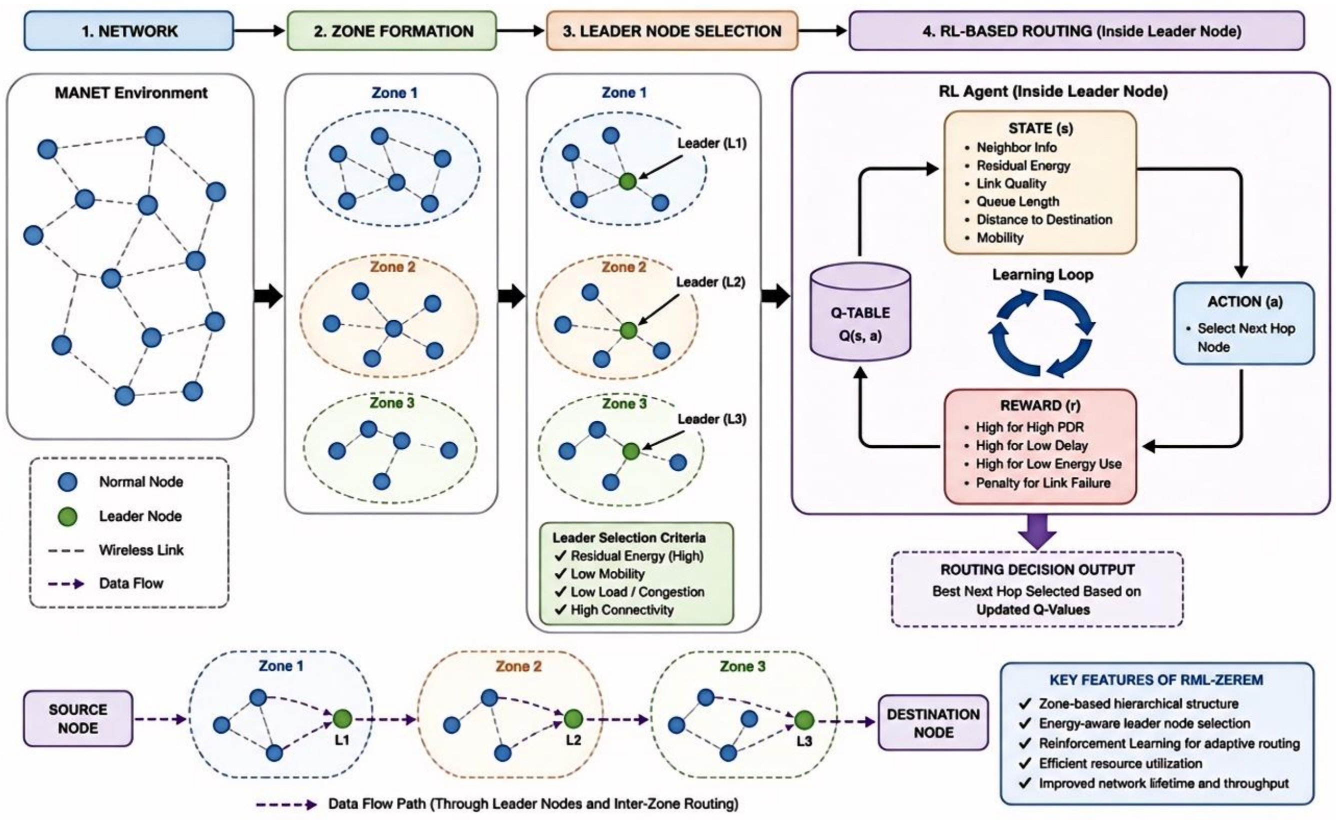

Figure 1: Overall Architecture and Operational Workflow of RML-ZEREM

- Each node uses its available battery power to operate its functions.
- The nodes in the system move around while they change their locations.
- All nodes share the same abilities to transmit data and their communication distance.
- Each node keeps track of its adjacent nodes and their remaining energy.
- Wireless links become active when two nodes establish a connection within their operating range.

The assumptions create an authentic MANET environment, which enables testing of the routing protocol under different network conditions.

3.3 Zone Formation Strategy

The proposed protocol divides the network into multiple logical zones, which use node proximity as the basis for their partitioning to achieve better scalability and decreased routing costs. The system creates zones that combine all nodes that fall within the designated transmission range. The system updates zone structures whenever nodes move or network topology changes occur.

The zone-based organization delivers multiple benefits, which include localized routing operations and decreased control packet creation, lower route discovery costs, and better resource utilization. The system conducts routing operations inside specific zones, which use leader nodes to handle all network routing procedures instead of searching for routes throughout the entire network. The hierarchical system provides major benefits for scalability, together with routing performance improvements that apply to large mobile ad hoc networks that experience constant changes. Figure 2 shows how the proximity-based zone formation mechanism creates organized clusters from scattered individual nodes through its different stages, which start at Zone 1 and end at Zone 5, where certain nodes function as leader nodes and zone heads.

3.3.1 Dynamic Zone Maintenance

To maintain routing efficiency under varying network conditions, the proposed RML-ZEREM protocol employs a dynamic zone maintenance mechanism. A zone update kicks in whenever any one of these happens: (i) a node drifts out of the communication range of its existing zone, (ii) a new node shows up and joins the network, (iii) the remaining energy of the current leader drops under a set threshold, or (iv) major shifts in the topology are noticed. During this zone upkeep, nodes regularly pass around nearby-node details, then zone membership gets adjusted based on that info. If the leader node turns bad, because of mobility or just energy depletion, then a fresh leader gets picked through the energy-aware leader selection approach.

In order to keep the routing performance efficient in those dynamic MANET environments, the proposed RML-ZEREM protocol brings in a dynamic zone maintenance idea, so nodes don't just sit there and hope for the best. Each node keeps watching its relative location compared to the zone leader. If the distance between a node and its present leader turns out to be larger than that predefined zone radius, then the node gets reassigned toward a more fitting neighbouring zone. Also, when the residual energy of the current zone leader

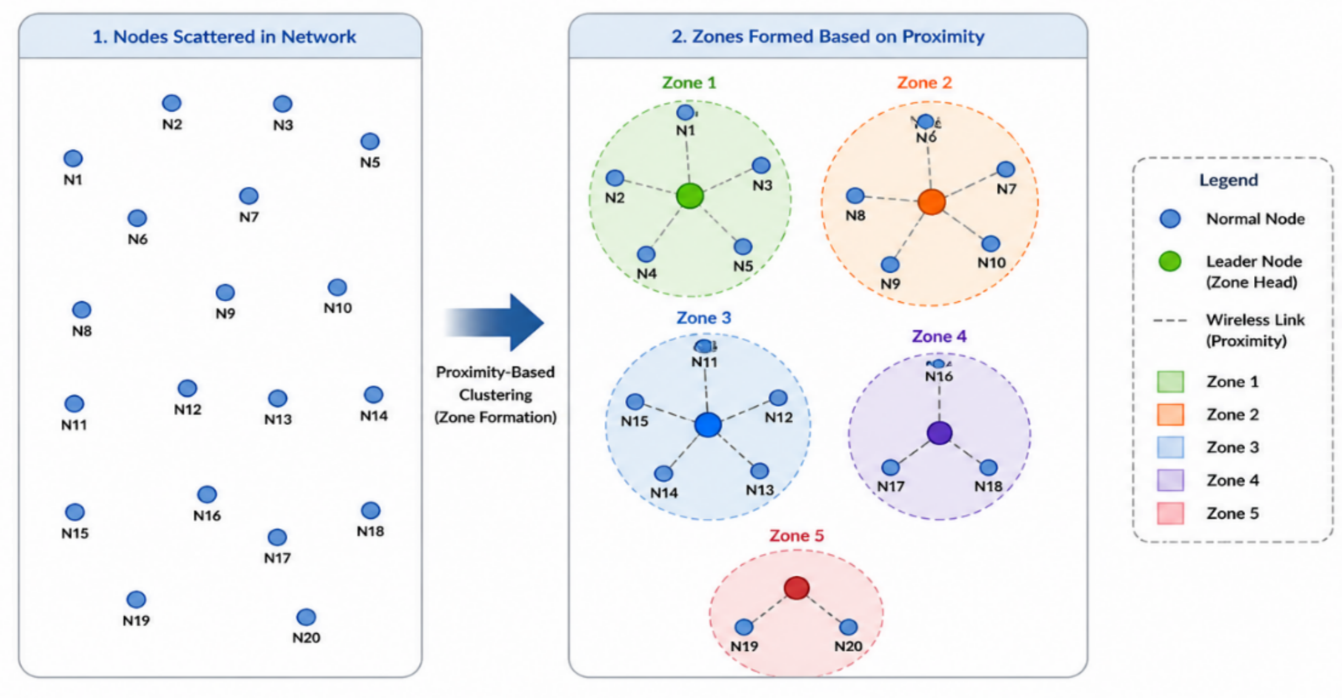

Figure 2: Proximity-Based Zone Formation Mechanism in MANETs

sinks under a certain threshold, a fresh leader election is triggered. This is done to keep the zone handling stable, and yes, to stay energy efficient. The protocol also incorporates an autonomous network expansion mechanism to accommodate newly joined nodes. When new nodes arrive, they are automatically assigned to the nearest available zone, based on proximity plus connectivity criteria, not only one factor. After any membership shifts, or a leader role change, or even a small network topology update, the zone information table gets refreshed, so the routing data stays correct and inter-zone communication remains smooth. Overall, this dynamic maintenance approach helps the network adapt properly to mobility, topology swings, and energy variations, which leads to better routing stability, improved scalability, and stronger total network performance.

3.4 Energy-Aware Leader Node Selection

Each zone selects a leader node that handles routing operations, learns new routes, and establishes connections with other zones. The leader node functions as a reinforcement learning agent that safeguards routing performance and network stability throughout the system.

An energy-aware scoring system performs leader node selection through evaluation of these specific node characteristics:

- The remaining energy resources of the node
- The current data transmission activities of the node
- The extent to which the node can connect with nearby nodes
- The ability of the link to maintain its connection

The selection process for leaders prioritizes nodes that possess greater remaining energy resources and reduced communication requirements, and better network connections. The leader selection process receives updates at regular intervals to maintain its effectiveness during changes in network topology while protecting the energy resources of all nodes in the system. The adaptive leader management strategy establishes a network resource allocation system that produces better routing results while extending the operational lifetime of the entire network. The Q-learning-based routing process becomes more effective through the selection of leaders who demonstrate both stability and energy-efficient performance. The energy-aware leader node selection strategy shown in Figure 3 demonstrates its multi-criteria evaluation method, which tests node performance based on their remaining energy and current load and their ability to connect with other nodes in order to find the best leader node for that specific zone.

Leader Score:

$$LS_i = \alpha E_i + \beta C_i + \gamma LQ_i - \delta TL_i$$

where

- E_i = Residual Energy
- C_i = Connectivity Degree
- LQ_i = Link Quality
- TL_i = Traffic Load

The node with the highest leader score is selected as the zone leader.

3.5 Reinforcement Learning Model

The proposed RML-ZEREM protocol uses Q-learning, which serves as a model-free Reinforcement Learning (RL) technique to develop intelligent adaptive routing systems that work in dynamic MANET networks. The framework designates every zone leader as an RL agent who studies network conditions to develop optimal routing methods through their

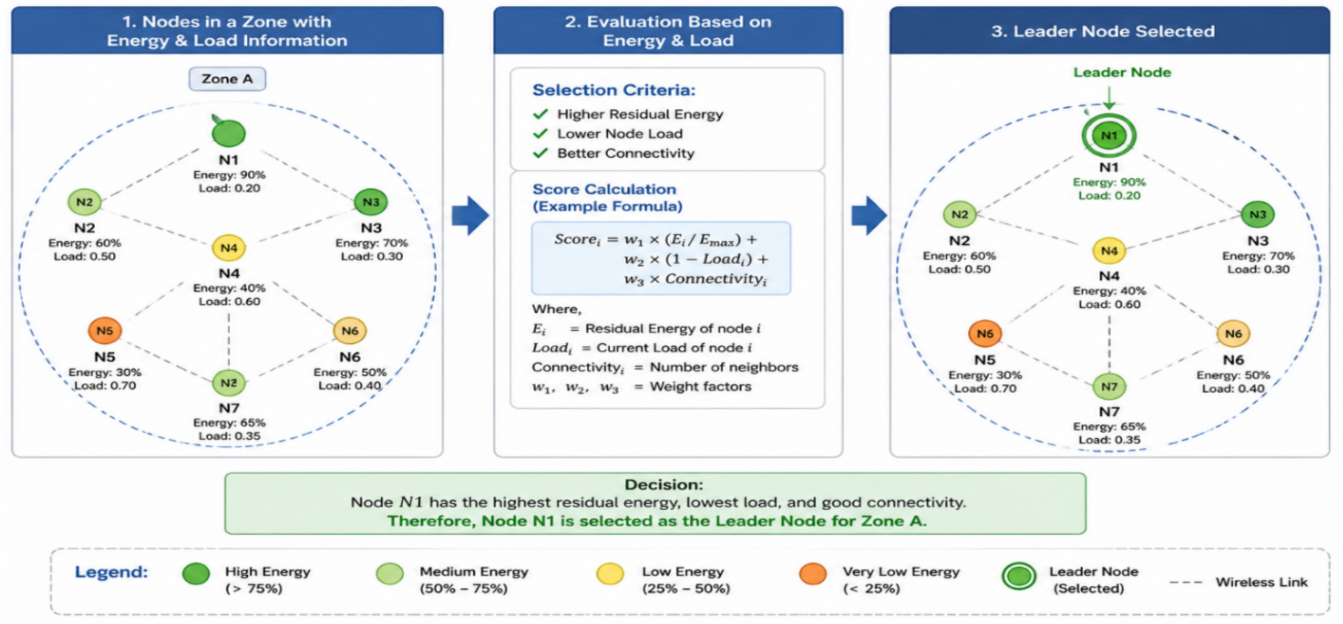

Figure 3: Energy-Aware Leader Node Selection Strategy

continuous network interactions. The Q-learning mechanism enables leader nodes to choose optimal routing paths by using maximum long-term rewards to decrease energy costs, routing overhead, and transmission delays. The routing system develops the capacity to handle topology variations and different network conditions through its ongoing learning process and feedback system.

3.5.1 State Space Definition

The state space shows the present network status, which the leader node observes. The state definition uses various network parameters that affect routing performance to establish each network state. The first part of the state representation exists in Equation (1) as its defined form.

$$s = \{E, L, M, D\} \tag{1}$$

where:

- E represents residual energy,
- L denotes traffic load,
- M indicates node mobility, and
- D represents inter-node distance or link stability.

The selected state variables jointly represent the energy condition, network congestion level, mobility behaviour, and communication stability of the network. The multi-parameter state information lets the RL agents make routing decisions from both energy efficacy and network stability perspectives.

3.5.2 Action Space

The action space defines the set of routing decisions available to the leader node at a given state. The system allows users to choose between two options, which involve selecting a routing path to follow or designating a next-hop node to handle packet transmission.

The system provides three routing options, which include the following:

- Users must choose a nearby node that will handle their packet transmission needs
- Users must choose an alternative routing path to follow
- Users must prevent their system from using both congested nodes and low-energy nodes
- Users must choose stable communication links that will enable their data transmission needs.

The action selection process is designed to optimize routing performance while minimizing energy consumption, transmission delay, and packet loss. The set of routing actions available to the Q-learning agent is presented in Table 2.

As shown in Table 2, the Q-learning agent can select neighbouring nodes, identify alternative routes, avoid congested and low-energy nodes, and select stable communication links. These actions enable the routing agent to adapt its decisions according to the current network conditions and improve overall routing efficiency.

Table 2: Action Set Used in the Q-Learning-Based Routing Process.

Action	Description
A1	Select Neighbour Node
A2	Select Alternative Route
A3	Avoid Congested Node
A4	Avoid Low-Energy Node
A5	Select Stable Link

3.5.3 Reward Function Design

The reinforcement learning process depends on the reward function because it directs the agent towards achieving optimal routing behaviour. The reward function in the proposed protocol functions to promote routing decisions that achieve energy efficiency while maintaining stability and reliability.

The reward is determined based on the following performance parameters:

- Energy efficiency
- Packet delivery performance
- Transmission delay and congestion level

The generalized reward formulation is expressed through Equation (2).

$$r = w_1 \cdot EE + w_2 \cdot PDR - w_3 \cdot Delay \tag{2}$$

where:

- EE represents energy efficiency,
- PDR denotes Packet Delivery Ratio,
- $Delay$ indicates transmission delay, and
- w_1, w_2, w_3 are weighting factors used to balance routing objectives.

In this study, the weighting factors were determined through preliminary simulation experiments to achieve a balanced trade-off among routing reliability, energy efficiency, and transmission performance. The values of the weighting factors were set as $w_1 = 0.4$, $w_2 = 0.4$, and $w_3 = 0.2$. More emphasis was given to energy efficiency as well as packet delivery performance because they directly shape the network lifetime and communication reliability in MANET environments.

The multi-objective reward function allows simultaneous optimization of three targets, which include energy consumption and routing reliability, and transmission performance. The RL agent develops effective routing methods for changing MANET environments through continuous updates of rewards, which depend on network feedback.

The proposed RML-ZEREM protocol uses reinforcement learning methods, which Figure 4 demonstrates. The leader nodes use current network conditions to determine which routing paths provide the best performance. The system generates reward values from routing decisions, which it uses to update the Q-values in the Q-table. The learning process starts with the observation of network states followed by the execution of forwarding actions, and ends with Q-table updates, which result from environmental rewards. The leader nodes use this ongoing learning system to enhance their routing decisions, which leads to better network performance.

3.6 Q-Learning-Based Routing Mechanism

The Q-learning algorithm enables dynamic MANET networks to make intelligent routing decisions through the RML-ZEREM protocol. The Q-table, which each leader node operates, contains Q-value information for all possible state-action combinations. Networking feedback together with routing results drives the ongoing development of Q-value information within that system.

The Q-learning update equation used in the proposed protocol is defined in Equation (3)

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \tag{3}$$

Where:

- α represents the learning rate,
- γ denotes the discount factor,
- r is the immediate reward received after action execution,
- s' indicates the next network state,
- $\max_{a'} Q(s',a')$ represents the maximum future reward for the next state.

The proposed protocol routing process executes the following sequence of steps. The current network state needs to be observed. The ε-greedy policy enables the selection of

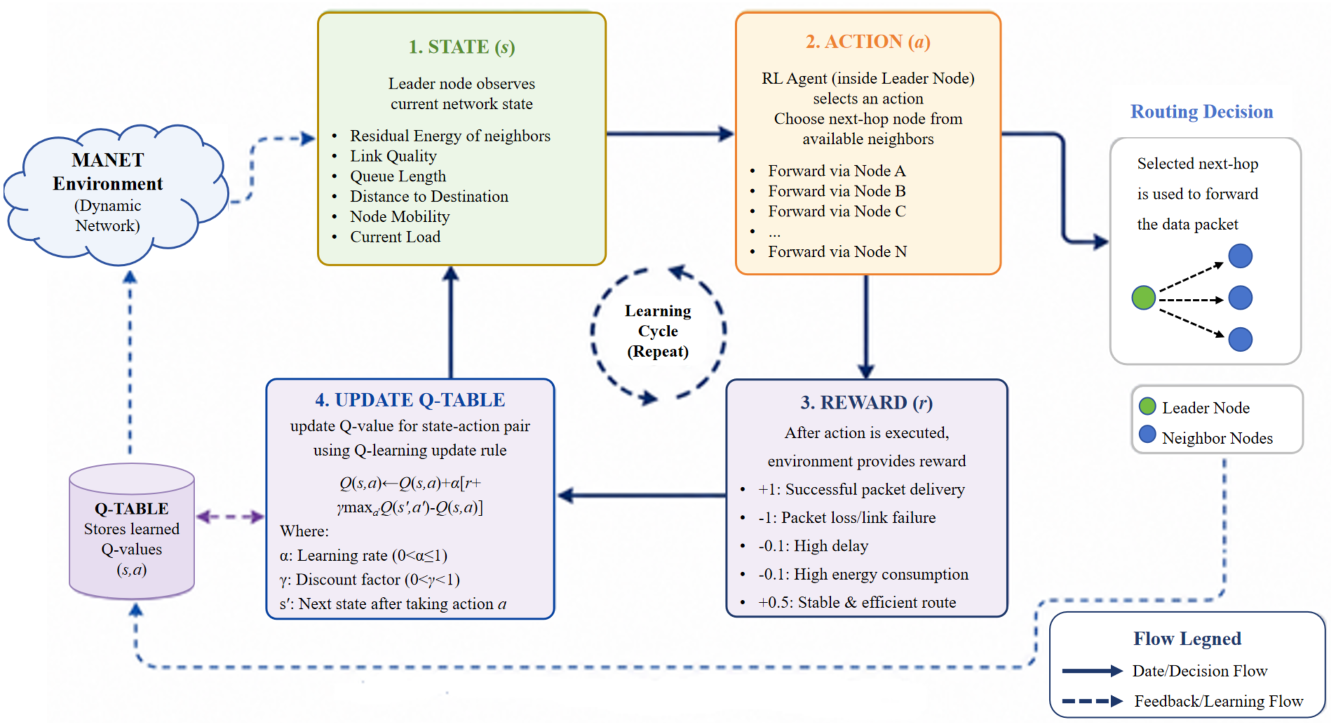

Figure 4: Reinforcement Learning Framework for Adaptive Routing Decisions

an action. The system performs its routing decision, which determines how packets will be sent. The system delivers a reward that assesses how routing functions. The system computes Q-value updates through the Q-learning equation implementation. The RL agent continuously learns from network interactions to discover the best routing paths that achieve maximum rewards while ending up consuming the least energy and time and routing resources.

3.6.1 Parameter Configuration

The key Q-learning parameters used in the proposed RML-ZEREM protocol are presented in Table 3. The learning rate (α) was set to 0.7 to control the extent to which newly acquired knowledge influences the Q-value update. The discount factor (γ) was set to 0.8 to balance immediate and future rewards, while the exploration rate (ε) was set to 0.1 to maintain a balance between exploring alternative routing actions and exploiting the currently learned optimal actions. These parameter settings were selected to support stable learning and effective adaptation to changing network conditions during the routing process.

Table 3: Q-Learning Parameter Settings.

Parameter	Value
Learning Rate (α)	0.7
Discount Factor (γ)	0.8
Exploration Rate (ε)	0.1

3.7 Route Discovery and Maintenance

The RML-ZEREM protocol uses reinforcement learning adaptive routing mechanisms to improve its route discovery and maintenance functions. The proposed system uses zone leader-based techniques to manage routes because conventional routing systems need to send out network-wide broadcasts at frequent intervals.

Route Discovery

Leader nodes use their Q-value knowledge during route discovery to determine the best routing paths that connect different zones. The system limits its route discovery process to specific zones, which leads to reduced network routing costs and decreased control packet delivery. The RL agent evaluates multiple candidate paths based on parameters such as residual energy, delay, traffic load, and link stability before selecting the optimal route.

Route Maintenance

The RL agent uses Q-value updates to find new routes after a link failure or topology change disrupts the network. The adaptive system reduces the need to discover new routes while enabling fast recovery from communication breakdowns. The protocol enhances routing reliability, decreases packet loss, and sustains consistent network performance in changing environmental conditions.

3.7.1 Inter-Zone Communication Mechanism

During inter-zone communication, the source node forwards packets to its corresponding zone leader. The source leader evaluates neighbouring zone leaders using Q-values and forwards the route request toward the destination zone. Intermediate leaders exchange routing information and collaboratively establish an energy-efficient path between zones. Once the destination zone is reached, the destination leader forwards packets to the target node. This hierarchical routing process significantly reduces network-wide route discovery overhead and improves routing scalability.

Figure 5 shows how the proposed RML-ZEREM protocol establishes and maintains routes through its design. The source node sends route requests to its zone leader, who contacts adjacent zone leaders for the purpose of building effective inter-zone routing connections. The system uses reinforcement learning to select the best paths by evaluating energy efficiency, delay times, and link stability. The system uses localized route maintenance to quickly adjust to changes in network topology.

3.8 Algorithm Design of RML-ZEREM

The complete operational procedure of the proposed RML-ZEREM protocol is summarized in Algorithm 1. The algorithm uses a combination of zone-based routing and energy-aware leader node selection and Q-learning-based

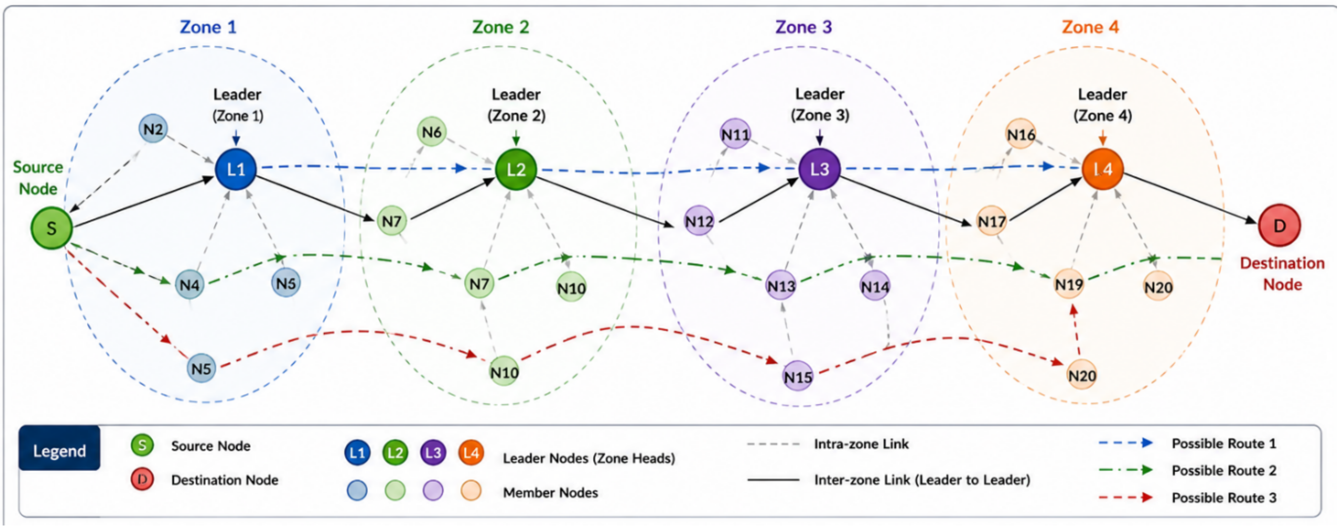

Figure 5: Route Establishment and Maintenance Process in RML-ZEREM

adaptive routing to establish effective communication within MANET networks.

Algorithm 1 RML-ZEREM Routing Protocol

Require: Network $G(N,L)$
Ensure: Optimal Energy-Efficient Routing Paths

Initialize network parameters and node configurations.
Divide the network into multiple logical zones.
for each zone **do**
  Evaluate node characteristics, including residual energy, traffic load, connectivity degree, and link stability.
  Compute the leader score for each node.
  Select the node with the highest leader score as the zone leader.
end for
Initialize the Q-table for each zone leader.
Set Q-learning parameters:
  Learning Rate (α)
  Discount Factor (γ)
  Exploration Rate (ε)
while network is active **do**
  Observe the current network state s.
  Select action a using the ε-greedy policy.
  Execute route selection and packet forwarding.
  Calculate the reward r based on energy efficiency, packet delivery ratio, and transmission delay.
  Update the Q-value using Equation (3)
  Update the network state to s'.
  if a link failure or topology change occurs **then**
    Update zone information.
    Re-elect the leader node if required.
    Discover an alternative routing path.
  end if
end while
Evaluate routing performance using:
  Energy Consumption
  Throughput
  Packet Delivery Ratio (PDR)
  End-to-End Delay
Compare the performance of RML-ZEREM with existing routing protocols.
Return the optimal energy-efficient routing path.

The RML-ZEREM protocol uses an energy-aware data forwarding system, which Figure 6. Data packets are sent through leader nodes and intermediate nodes, which are chosen based on their remaining energy, current network traffic, and their link performance. The reinforcement learning system uses Q-values to select the best routing paths, which helps the system to stay away from low-energy nodes and unstable nodes while it extends network lifespan and improves routing performance.

4 Mathematical Modelling

The mathematical formulation of the proposed RML-ZEREM protocol. The proposed model uses graph theory and energy consumption analysis with reinforcement learning and policy optimization methods to create adaptive energy-efficient routing systems for dynamic MANET environments.

4.1 Network Representation (Graph Model)

The MANET describes its structure through an undirected graph, which Equation (4) shows.

$$G = (N, L) \tag{4}$$

where:

- $N = \{n_1, n_2, \ldots, n_k\}$ represents the set of mobile nodes
- $L \subseteq N \times N$ represents the set of communication links

A link $l_{ij} \in L$ exists between nodes n_i and n_j if they are within transmission range.

Each node n_i is associated with a residual energy value, as defined in Equation (5):

$$E_i(t) \tag{5}$$

where $E_i(t)$ denotes the remaining energy of the node i at time t.

The network is partitioned into zones using a mapping function, as defined in Equation (6).

$$Z : N \to \{Z_1, Z_2, \ldots, Z_m\} \tag{6}$$

where each node belongs to exactly one zone based on proximity constraints and communication range.

4.2 Energy Consumption Model

The energy usage of mobile ad hoc networks occurs mainly through the transmission and reception of data packets. Equation (7) shows the total energy usage of a node through its different operational activities.

$$E_{\text{total}} = E_{\text{tx}} + E_{\text{rx}} \tag{7}$$

The transmission energy is modelled as defined in Equation (8).

$$E_{\text{tx}}(d) = E_{\text{elec}} \cdot k + \varepsilon_{\text{amp}} \cdot k \cdot d^{\eta} \tag{8}$$

The reception energy is defined in Equation (9).

$$E_{\text{rx}} = E_{\text{elec}} \cdot k \tag{9}$$

where:

- E_{elec} = energy consumed per bit for transmission/reception
- ε_{amp} = amplifier energy constant
- k = packet size (bits)
- d = distance between nodes
- η = path loss exponent (typically 2–4)

The energy model system lets the routing protocol select routes by using two factors, which are the distance to communicate and the remaining power of the nodes.

4.3 Q-Learning Formulation

The Q-learning algorithm serves as the main learning method, which allows leader nodes to discover the best route patterns

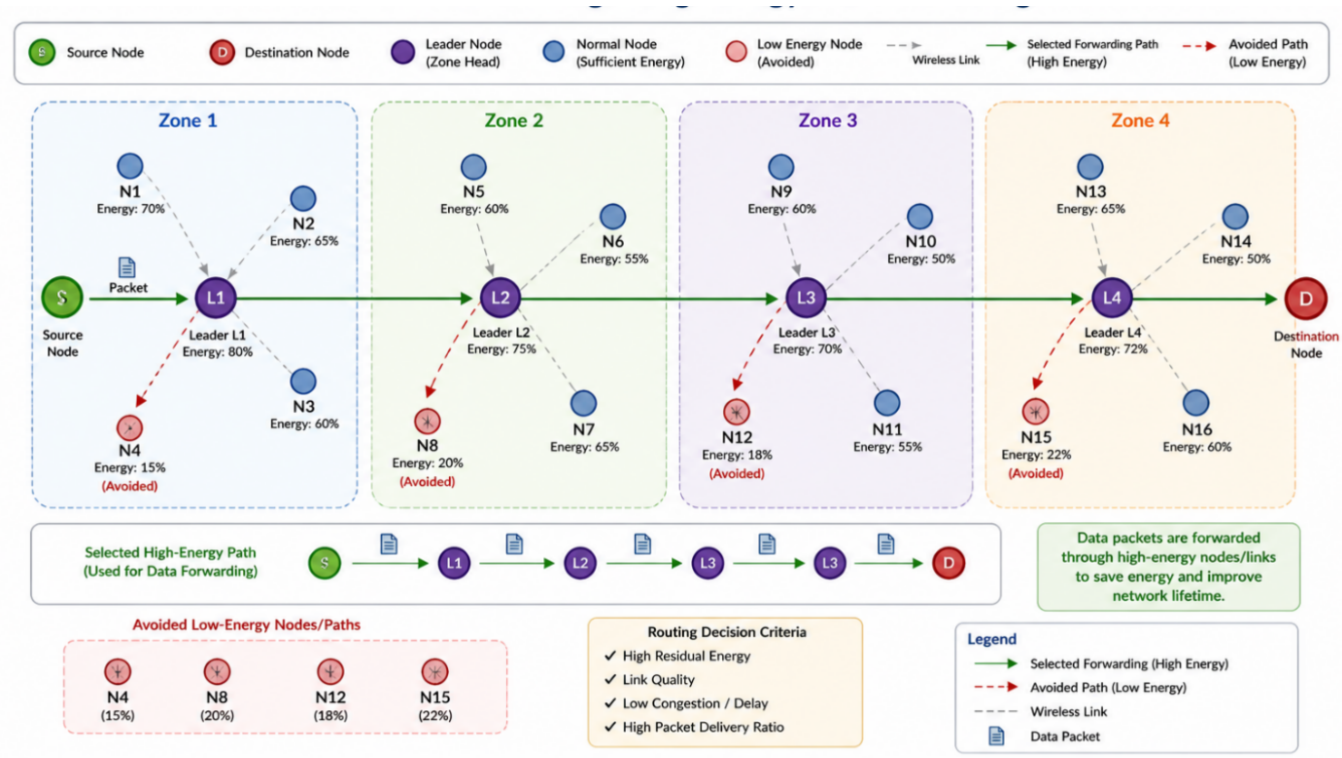

Figure 6: Energy-Aware Data Forwarding Mechanism through Leader Nodes

through their ongoing experience with the network environment. The Q-learning update rule introduced in Equation (3) in Section 3.6.

$$Q(s,a) \leftarrow Q(s,a) + \alpha\,[r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$$

Where:

- $Q(s,a)$ = current Q-value for state-action pair
- $\alpha \in (0,1)$ = learning rate
- $\gamma \in (0,1)$ = discount factor
- r = immediate reward
- s' = next state
- $\max_{a'} Q(s',a')$ = maximum future reward

The routing agent improves its routing decisions through an ongoing process that uses network experience together with system feedback. The convergence behaviour of the Q-learning agent is further evaluated in the simulation results section through reward convergence analysis.

The Q-learning-based routing decision process, which the proposed protocol uses, is shown in Figure 7. The figure demonstrates how network states are mapped to routing actions through Q-values, which enable adaptive route selection and optimized route selection throughout the time period. As illustrated in Figure 7, the routing agent continuously interacts with the network environment by observing network states, selecting routing actions, receiving rewards, and updating Q-values. This learning cycle enables the routing agent to gradually identify optimal routing paths that maximize energy efficiency and packet delivery performance while minimizing transmission delay.

4.4 Reward Function Formulation

The reward function exists to optimize three different routing objectives, which include energy efficiency, packet delivery performance, and delay reduction. Based on Equation 2 in Section 3.5.3, the reward formulation is defined as:

$$r = w_1 \cdot \frac{E_{\text{res}}}{E_{\max}} + w_2 \cdot PDR - w_3 \cdot D_{\text{avg}} \tag{10}$$

where:

- E_{res} = residual energy of the node
- $E_{\max}$ = maximum initial energy
- PDR = packet delivery ratio
- D_{avg} = average delay
- w_1, w_2, w_3 = weighting factors such that $w_1 + w_2 + w_3 = 1$

In the proposed protocol, the weighting factors are selected as $w_1 = 0.4$, $w_2 = 0.4$, and $w_3 = 0.2$, as discussed in Section 3.5.3. Higher weights are assigned to residual energy and packet delivery ratio because these parameters directly influence network lifetime and communication reliability. A comparatively lower weight is assigned to delay to maintain a balance between routing efficiency and transmission performance. The reward system drives people to choose nodes that possess greater remaining power. The system needs to achieve both dependable packet transmission and transmission time efficiency. The proposed reward function enables adaptive routing to work efficiently in different network environments by balancing multiple performance metrics.

4.5 Policy Optimization Strategy

The proposed protocol uses an ε-greedy policy mechanism to achieve a balance between exploration and exploitation throughout route learning. The routing policy is defined in

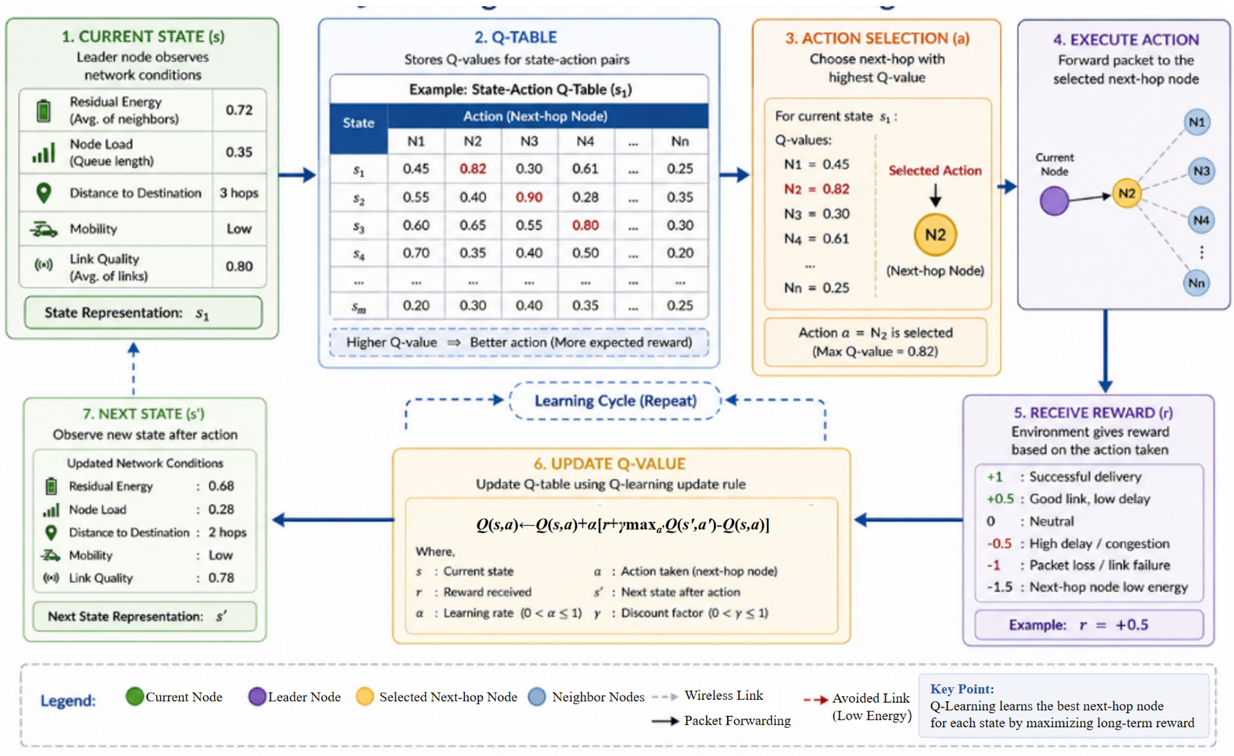

Figure 7: Q-Learning-Based Decision Process for Optimal Path Selection

Equation (11).

$$\pi(s) = \begin{cases} random\ action, & \text{with probability } \varepsilon \\ arg\ max_a Q(s,a), & \text{with probability } 1-\varepsilon \end{cases} \tag{11}$$

where ε represents the exploration rate.

In this study, the exploration rate ε is set to 0.1. This value provides an effective balance between route exploration and exploitation of previously learned routing knowledge.

The ε-greedy strategy enables the RL agent to search for new routing paths with probability ε while using its existing knowledge of optimal routes at $1-\varepsilon$. The system achieves better convergence results because it maintains better performance results while preventing the routing agent from choosing wrong routing paths. The computational complexity of the proposed RML-ZEREM protocol mainly depends on how zone formation works, leader selection, and the Q-learning based route optimization. For the zone formation part, the complexity is $O(N)$, here N is the total number of nodes. Then, for leader selection inside each zone, there is the need to inspect or evaluate node parameters, which ends up costing about $O(ZN)$, where Z is the number of zones. After that, the Q-learning update step relies on the number of states S and the number of actions A, so the cost becomes $O(SA)$. So overall, the routing protocol complexity can be roughly written as:

$$O(ZN+SA)$$

Even if the reinforcement learning mechanism brings extra computational overhead, the fact that zone leaders take over many coordination tasks greatly curbs network wide routing operations. This is why the approach keeps scaling well and stays efficient for large-scale MANET settings.

4.6 Performance Metrics

The proposed RML-ZEREM protocol performance assessment uses standard MANET performance metrics, which include energy efficiency, throughput and packet delivery ratio. The routing mechanism assessment uses these metrics to evaluate its performance across various network situations.

4.6.1 Energy Efficiency

Energy efficiency measures the ratio of successfully transmitted useful data to the total energy spent during network operations. The energy efficiency measure has its definition established through Equation (12).

$$EE = \frac{\text{Total Data Delivered}}{\text{Total Energy Consumed}} \tag{12}$$

Higher energy efficiency values indicate better utilization of network energy resources and improved network lifetime.

4.6.2 Throughput

Throughput measures the successful data transmission rate achieved during the entire testing period. The measurement is determined through the application of Equation (13).

$$T = \frac{\sum_{i=1}^{N} \text{Packets Received}_i \times \text{Packet Size}}{\text{Total Simulation Time}} \tag{13}$$

The equation includes N as the total number of received packets and Packet Size as the size of each transmitted packet. Throughput measurement uses two standard units, which are kilobits per second (kbps) and megabits per second (Mbps). The higher throughput values bring about better data transmission efficiency and reliability.

4.6.3 Packet Delivery Ratio (PDR)

The Packet Delivery Ratio (PDR) measures the ratio between successfully received packets and all transmitted packets. It is defined in Equation (14).

$$PDR = \frac{\text{Total Packets Received}}{\text{Total Packets Sent}} \tag{14}$$

PDR functions as a critical measurement that assesses both routing reliability and communication efficiency. The dynamic MANET environments show better packet transmission performance because higher PDR values lead to improved routing stability.

4.6.4 Q-Learning Parameter Selection

The Q-learning parameters were selected based on commonly adopted values reported in reinforcement learning-based MANET routing studies and were further refined through preliminary simulation experiments.

5 Simulation Setup and Parameters

This section presents the simulation environment, network configuration, and evaluation parameters employed to assess the performance of the proposed RML-ZEREM protocol. The suggested protocol is set side by side with ordinary MANET routing protocols across several simulation scenarios, to check how well it performs. In this way they validate efficiency, scalability, and adaptability, even when the network is changing fast and the topology is moving around all the time.

5.1 Simulation Environment

The performance evaluation of the proposed routing protocol is done using the NS-2.35 (Network Simulator 2) platform. NS-2.35 was selected for this study due to its widespread use in MANET research and its ability to provide a flexible simulation platform supporting wireless communication models, routing protocols, and node mobility patterns. In general, the simulation framework includes something like:

- NS-2.35 simulator for the network implementation
- TCL scripting for building the simulation scenarios
- Trace file analysis for the performance evaluation

The simulated environment shows quite real MANET conditions, since it mixes dynamic topology shifts, wireless communication constraints and node travel, plus multi hop routing. To make the results more dependable, every simulation scenario was run several times with different random seeds. The performance numbers we report are the mean outcomes from all those runs. This method reduces the influence of randomness, that comes from node mobility and also from traffic generation patterns.

5.2 Network Configuration

In the simulation, we look at a mobile ad hoc network, where there are mobile nodes that keep moving around a fixed geographical area. The whole setup relies on multi-hop wireless communication for how nodes connect, but every single node has a limited battery power supply, like it can only go so far before it fails.

During the simulation, several assumptions and configurations are put in place, such as:

- Nodes are dropped randomly across the whole simulation region, so the space is fully covered.
- Each node must begin with a restricted power capacity, from the start.
- Communication happens through multi-hop wireless transmission, rather than a direct link all the time.
- The Random Waypoint Mobility Model is used to describe how the nodes travel and change direction.
- Network behaviour is examined by varying both the node speed and the pause duration between moves.
- Traffic is created using Constant Bit Rate sources, and these sources run over UDP links.

To properly evaluate the proposed protocol, we run many simulation scenarios, including trying different simulation durations, different degrees of node mobility, multiple network size settings, and also different dimensions for the simulation area. Overall, the RML-ZEREM protocol shows what it can do across these tests, and the scenarios are meant to highlight things like its ability to scale, its routing efficiency, its flexibility under change, and its energy-related performance.

5.3 Simulation Parameters

The key simulation parameters used for performance evaluation are summarized in Table 4.

Table 4: Simulation Parameters for Performance Evaluation

Parameter	Value
Simulator	NS-2.35
Simulation Time	100–600 seconds
Network Area	500×500 m^2 to 1000×1000 m^2
Number of Nodes	50, 75, 100
Mobility Model	Random Waypoint
Node Speed	3–9 m/s
Pause Time	100 seconds
Traffic Type	CBR (UDP)
Packet Size	512 bytes
Transmission Range	250 m
Initial Energy	1000 Joules
Routing Protocols Compared	AODV, AOMDV, RML-ZEREM
MAC Protocol	IEEE 802.11
Queue Type	DropTail/PriQueue
Learning Rate (α)	0.7
Discount Factor (γ)	0.8
Exploration Rate (ε)	0.1
Reward Weight (w_1)	0.4
Reward Weight (w_2)	0.4
Reward Weight (w_3)	0.2

In the proposed RML-ZEREM protocol, the reinforcement learning parameters were selected based on preliminary simulation studies and commonly recommended values reported in previous Q-learning-based routing research. For the learning rate, we used $\alpha = 0.7$, so the routing agent can take on new situations without completely forgetting what it has already learned. Then, the discount factor $\gamma = 0.8$ is there to keep the focus on future rewards, and to make the long-run routing behave in a more stable way. The exploration rate $\varepsilon = 0.1$ is set to somehow balance exploration of routes versus exploitation of the best option found so far. Likewise, the reward weighting factors were set as $w_1 = 0.4$, $w_2 = 0.4$, and $w_3 = 0.2$. They were selected so that energy friendliness and packet delivery outcome get prioritized, while the transmission delay is pushed down a bit, not totally ignored, but reduced.

6 Results and Performance Evaluation

So, the performance of the proposed RML-ZEREM protocol was checked in a few simulation scenarios and also set side-by-side with common MANET routing protocols, like AODV and AOMDV. In general, we look at key performance indicators such as Energy Consumption, Throughput, and Packet Delivery Ratio (PDR), just to see how efficient, dependable, and adaptable the proposed routing framework is. The evaluation is done using those same key performance metrics, Energy Consumption, Throughput, and Packet Delivery Ratio (PDR) again, so the comparison stays consistent.

6.1 Scenario 1: Impact of Simulation Time

6.1.1 Energy Consumption Analysis

Table 5 shows the energy consumption performance of AODV, AOMDV, and the proposed RML-ZEREM protocol, for simulation durations that go from 100 s up to 600 s. With longer simulation time, energy usage also climbs for every protocol, because the network keeps working longer, packets keep being sent continuously, and routing activities happen over and over again. Nevertheless, the proposed RML-ZEREM protocol achieves consistently lower energy consumption than AODV and AOMDV across the entire simulation period. More specifically, RML-ZEREM's energy consumption moves from 152.00 J at 100 s to 185.00 J at 600 s, while AODV goes from 157.84 J to 196.13 J, and AOMDV changes from 157.84 J to 196.08 J over the same interval. So overall, the gap remains noticeable even as the load grows. Overall, the outcomes suggest that RML-ZEREM delivers energy savings of around 3.70% to 5.67% compared with AODV, and about 3.70% to 5.65% compared with AOMDV. Moreover, the performance improvement becomes more pronounced with increasing simulation time, highlighting the effectiveness of the proposed routing strategy in long-duration MANET scenarios. The better behaviour of RML-ZEREM can be linked to what it combines: energy-conscious leader selection, zone-based routing that adapts, and route refinement using Q-learning. The reinforcement learning part keeps learning smarter route choices under different network conditions, while the zone structure reduces unneeded route discoveries and also cuts down routing control overhead. As a result, this protocol helps lower energy spending and extends the overall network lifetime, even when things become more demanding.

Figure 8 shows how energy use changes over simulation time. As the network keeps running, energy consumption still goes up for all routing protocols, but RML-ZEREM stays at the bottom, consistently, with the lowest energy usage the whole time. This also hints that the proposed Q-learning-based routing setup really works for reducing extra routing actions, and in the end, it helps energy utilization to be more efficient. You can also see that RML-ZEREM keeps the least energy consumption for every simulation duration, so it proves it can do energy-efficient and adaptive routing, even when MANET conditions are dynamic and shift around.

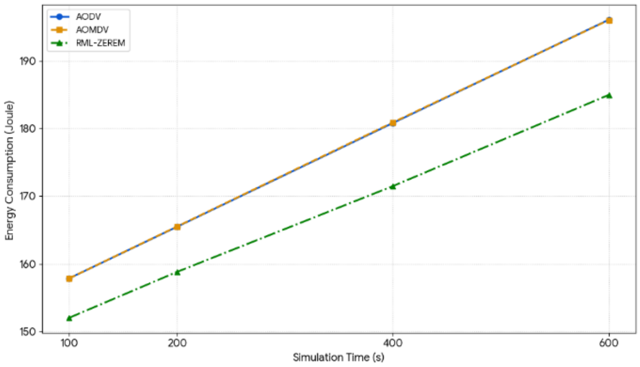

Figure 8: Energy Consumption vs. Simulation Time Comparison

6.1.2 Throughput Analysis

Table 6 shows how AODV, AOMDV, and the proposed RML-ZEREM protocol perform in terms of throughput for simulation times that go from 100 s up to 600 s. When the simulation time keeps getting longer, throughput usually goes up too, mainly because the nodes can keep transmitting data, and the overall communication in the network lasts longer. Still, the pace of that improvement is not the same for every routing option, since their route upkeep strategies and traffic handling methods are a bit different. In this set of results, RML-ZEREM keeps landing the top throughput across all scenarios. In particular, its throughput climbs from 1965.74 kbps at 100 s, then it reaches 11402.93 kbps by 600 s, which suggests that RML-ZEREM can maintain a solid and steady packet delivery during longer simulation durations too.

The results show that RML-ZEREM gets throughput improvements of 25.00% over AODV at both 100 s and 200 s, then the gain keeps moving up, to 35.99% at 400 s and it reaches 50% at 600 s. Compared with AOMDV, the proposed protocol delivers throughput gains of 42.50%, 50.15%, 11.52% and 4.83% for simulation durations of 100 s, 200 s, 400 s and 600 s, respectively. The better throughput performance of RML-ZEREM is mostly because of its Q-learning-based adaptive routing behaviour and energy-aware zone management. In other words, the reinforcement learning agent keeps refining its routing choices using network responses, remaining energy, and link state. At the same time, the zone-style design keeps routing operations more localized and that, reduces control chatter, more or less. Taken together, these methods boost route steadiness and packet delivery effectiveness, so more data ends up getting through during the limited simulation window. Figure 9 shows the

Table 5: Performance Comparison: Energy Consumption vs. Simulation Time

Simulation Time (s)	AODV (J)	AOMDV (J)	RML-ZEREM (J)	Improvement over AODV (%)	Improvement over AOMDV (%)
100	157.84	157.84	152.00	3.70	3.70
200	165.48	165.52	158.80	4.03	4.05
400	180.84	180.93	171.50	5.16	5.21
600	196.13	196.08	185.00	5.67	5.65

throughput patterns of the routing protocols that were tested. As shown in the results, the throughput achieved by RML-ZEREM increases progressively with simulation time and consistently remains higher than that of both AODV and AOMDV. Overall, the findings suggest the proposed approach can still support dependable and efficient data transfer when MANET conditions keep changing around it.

The observed throughput improvements, seem to show that the proposed Q-learning framework effectively, adapts to shifting network circumstances and keeps fairly stable routing tracks in place, so it ends up boosting the overall network throughput and communication efficiency, like it works pretty smoothly even when conditions change.

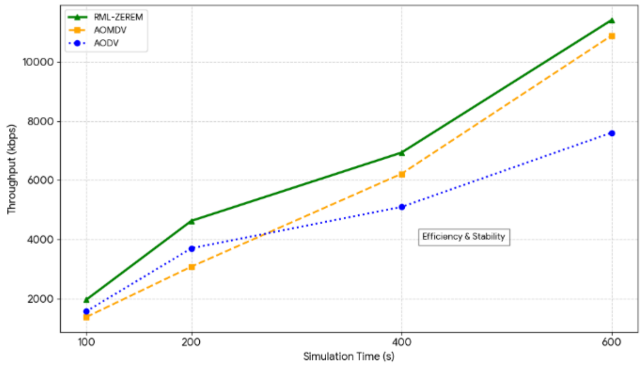

Figure 9: Throughput Trends over Extended Simulation Periods

Table 6: Throughput Comparison across Different Simulation Durations

Simulation Time (s)	AODV (kbps)	AOMDV (kbps)	RML-ZEREM (kbps)	Improvement over AODV (%)	Improvement over AOMDV (%)
100	1572.59	1379.51	1965.74	25.00	42.50
200	3700.73	3080.63	4625.92	25.00	50.15
400	5095.57	6213.76	6929.80	35.99	11.52
600	7602.21	10878.48	11402.93	50.00	4.83

6.1.3 Packet Delivery Ratio (PDR) Analysis

Table 7 shows the Packet Delivery Ratio (PDR) performance of AODV, AOMDV, and the proposed RML-ZEREM protocol, for simulation durations about 100 s up to 600 s. Overall, RML-ZEREM consistently achieves the highest packet delivery performance across all considered scenarios, demonstrating its robustness and adaptability under varying network conditions. More detailed, the PDR for RML-ZEREM climbs from 64.20% at 100 s, then up to 88.48% by 600 s, which suggests better route steadiness and more efficient packet delivery as the simulation gets longer. Still, the PDR is not fixed, it fluctuates with node mobility, because topology keeps shifting, and route failures show up more often. Even with that, the proposed protocol continues to hold higher PDR values than both AODV and AOMDV during the whole simulation window, no matter the time step. In addition, the results highlight that RML-ZEREM provides PDR gains in the neighbourhood of 8.30% to 15.15% relative to AODV, and around 4.14% to 9.00% compared with AOMDV. This better packet delivery behaviour is mainly linked to combining Q-learning based route optimization, energy aware leader selection, and zone-oriented routing management. The reinforcement learning agent keeps judging the network state, including residual energy, link stability, node mobility, and traffic load, so it can pick dependable forwarding paths and steer clear of unstable routes. Finally, the zone-based routing design helps confine route handling, so the routing overhead drops, and fewer packets get lost during repeated route rediscovery, or when links fail. Taken together, these parts strengthen route stability and boost the likelihood of successful delivery, which leads to improved communication reliability.

Figure 10 illustrates the Packet Delivery Ratio (PDR) trends of the evaluated routing protocols. As the simulation time increases, the Packet Delivery Ratio (PDR) of all protocols improves; however, RML-ZEREM consistently achieves the highest packet delivery performance across the evaluated scenarios. These PDR improvements seem to suggest that the proposed Q-learning framework actually learns and adapts well when the network conditions change, it also keeps routing paths steady and helps reduce packet drops too. As a result, the proposed approach enhances communication reliability in dynamic MANET environments.

The observed improvements in packet delivery ratio indicate that the proposed Q-learning framework effectively adapts to dynamic network conditions, maintains route stability, reduces packet loss, and thereby enhances overall communication reliability.

Table 7: Packet Delivery Ratio (PDR) vs. Simulation Time

Simulation Time (s)	AODV (%)	AOMDV (%)	RML-ZEREM (%)	Improvement over AODV (%)	Improvement over AOMDV (%)
100	55.75	58.90	64.20	15.15	9.00
200	65.41	68.45	74.10	13.28	8.25
400	74.96	78.83	84.20	12.33	6.81
600	81.70	84.96	88.48	8.30	4.14

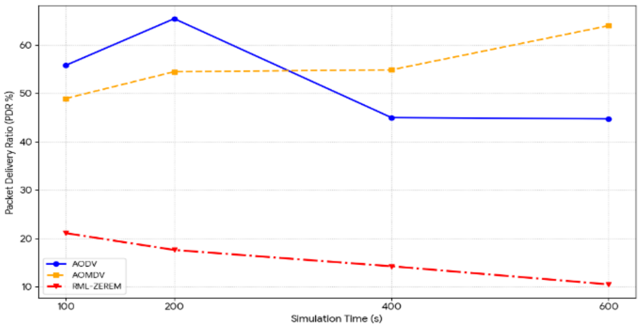

Figure 10: Packet Delivery Ratio (PDR) Analysis vs. Simulation Duration

6.2 Scenario 2: Impact of Node Mobility

Node mobility represents one of the major challenges in MANETs, since frequent node movements result in rapid topology variations, causing link failures and route instability. In order to judge how well the proposed routing protocol adapts when the network is constantly changing, the node speed was moved around from 3 m/s up to 9 m/s. Then the influence of mobility on energy usage, overall throughput, and packet delivery was observed and matched against AODV and AOMDV, to see how it behaves under those shaky conditions.

6.2.1 Energy Consumption Analysis

Table 8 presents the energy consumption performance of AODV, AOMDV, and the proposed RML-ZEREM protocol under different node mobility levels. When the node speed goes up, the energy usage slightly climbs for all routing schemes, mostly because route breaks happen more often, then there is the extra need for route rediscovery and additional control packet sending. Nevertheless, RML-ZEREM consistently maintains lower and more stable energy consumption than both AODV and AOMDV across all mobility scenarios, demonstrating its superior energy efficiency. More specifically, the energy consumption of RML-ZEREM remains relatively stable, decreasing slightly from 173.47 J at 3 m/s to 173.37 J at 9 m/s. In comparison, AODV exhibits fluctuating energy consumption, ranging from 180.50 J to 195.86 J, while AOMDV ranges from 180.78 J to 196.69 J over the same speed interval. Overall, RML-ZEREM achieves energy-consumption reductions of 3.96–11.46% compared with AODV and 4.11–11.85% compared with AOMDV. The superior energy efficiency of RML-ZEREM can be attributed to the integration of Q-learning-based adaptive routing, energy-aware leader selection, and zone-based route management within a unified framework. The reinforcement learning agent keeps adjusting its routing choices using residual energy, link stability, and network topology cues, so it tends to pick routes that are both steadier and less power hungry. On top of that, the zone-based structure helps confine routing tasks to smaller areas, which cuts down the route discovery work and also reduces those unnecessary control packet exchanges. Taken together, these parts help in keeping the routes more stable while pushing more balanced energy usage, even when mobility is constantly changing. Figure 11 shows how energy consumption shifts as node speed goes up. Even though higher mobility usually means more frequent topology disturbances and route breaks, RML-ZEREM still manages to keep the lowest energy consumption compared with the other routing protocols. The observed energy savings demonstrate the effectiveness of the proposed routing framework in achieving energy-efficient communication and enhancing overall network lifetime, even in highly dynamic environments.

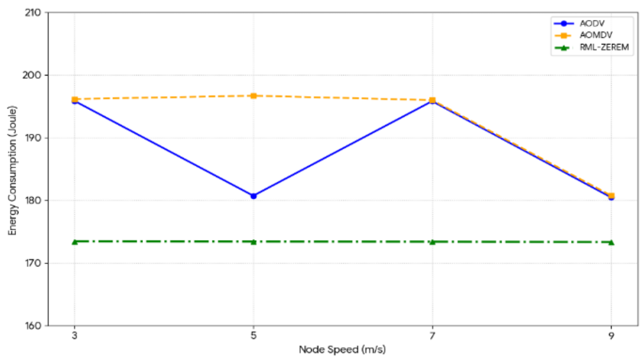

Figure 11: Energy Consumption Performance under Variable Node Mobility

Table 8: Impact of Node Speed on Energy Consumption

Node Speed (m/s)	AODV (J)	AOMDV (J)	RML-ZEREM (J)	Reduction vs AODV (%)	Reduction vs AOMDV (%)
3	195.86	196.17	173.47	11.46	11.61
5	180.74	196.69	173.44	4.04	11.85
7	195.82	195.99	173.42	11.43	11.56
9	180.50	180.78	173.37	3.96	4.11

6.2.2 Throughput Analysis

Table 9 presents the throughput performance of AODV, AOMDV, and the proposed RML-ZEREM protocol under varying node mobility conditions ranging from 3–9 m/s. Throughput is an important performance metric, as it reflects the efficiency of data transmission in dynamic MANET environments. From the results, it looks clear that RML-ZEREM consistently achieves higher throughput than the two reference routing methods in every mobility case. More specifically, the throughput of RML-ZEREM goes from 14831.45 kbps at 3 m/s down to 9597.85 kbps at 9 m/s, so it still holds fairly solid communication efficiency even as mobility keeps rising. When the node speed is 3 m/s, the proposed protocol brings throughput uplift of 69.48% and 46.67% relative to AODV and AOMDV, respectively. At 5 m/s, we can observe gains of 41.29% compared with AODV, and 9.18% compared with AOMDV. Similarly, at node speeds of 7 m/s and 9 m/s, RML-ZEREM consistently maintains its performance advantage, achieving throughput improvements of 18.19% and 33.30% over AODV, and 28.04% and 47.66% over AOMDV, respectively.

The superior throughput performance mainly comes from Q-learning-based adaptive route choice, energy-aware leader picking, and zone-style route management. The reinforcement learning agent keeps on updating its routing choices by looking at network feedback, leftover energy signals, and link steadiness readings, thereby favoring more reliable communication paths. Furthermore, the zone-based architecture confines routing operations to localized regions, which significantly reduces routing overhead and enhances packet forwarding efficiency with minimal disruption, even under dynamic node mobility conditions

Figure 12 shows throughput changes as node speed increases. Even though throughput drops a bit when mobility rises, because the topology keeps getting remixed and links break more often, RML-ZEREM still stays at the top, i.e., it has the highest throughput across all tested protocols. The throughput gains that were seen suggest the routing framework really does adapt well to changing network conditions, and it supports steady data delivery that is also efficient in high mobility MANETs.

The consistently higher throughput values indicate that the proposed reinforcement learning framework effectively adapts to varying mobility conditions and maintains stable routing performance. Consequently, the overall network efficiency is enhanced, and the elevated throughput serves as solid evidence of the framework's effective learning capability.

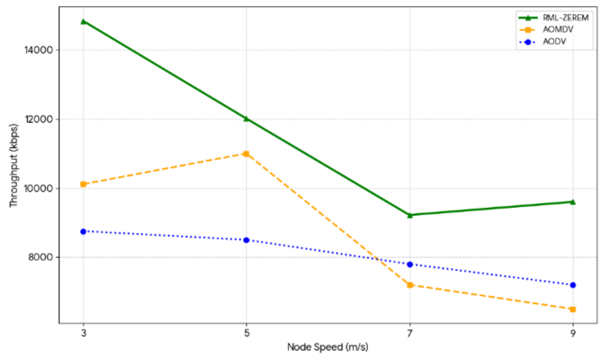

Figure 12: Throughput Variation with Increasing Node Speed

Table 9: Throughput Performance under Variable Node Mobility

Node Speed (m/s)	AODV (kbps)	AOMDV (kbps)	RML-ZEREM (kbps)	Improvement over AODV (%)	Improvement over AOMDV (%)
3	8752.32	10110.70	14831.45	69.48	46.67
5	8500.00	11000.00	12009.58	41.29	9.18
7	7800.00	7200.00	9218.64	18.19	28.04
9	7200.00	6500.00	9597.85	33.30	47.66

6.2.3 Packet Delivery Ratio (PDR) Analysis

Table 10 shows the Packet Delivery Ratio (PDR) performance for AODV, AOMDV, and the proposed RML-ZEREM protocol when node mobility changes, from 3 m/s up to 9 m/s. The results clearly demonstrate that RML-ZEREM keeps reaching better packet delivery outcomes than the two benchmark routing protocols in every scenario, even as node speed increases. Specifically, the PDR of RML-ZEREM increases from 65.12% at 3 m/s to 78.45% at 9 m/s. That means the protocol seems capable of holding steady, reliable packet delivery, despite the fact that the node mobility keeps increasing. Also, compared with AODV, RML-ZEREM gives PDR gains of 26.57%, 37.60%, 10.64%, and 30.75% for the 3 m/s, 5 m/s, 7 m/s, and 9 m/s cases. In a similar way, when compared with AOMDV, the proposed approach brings improvements of 9.53%, 22.06%, 15.25%, and 42.64% at the same mobility levels.

The superior packet delivery performance of RML-ZEREM seems to come from how it mashes together Q-learning driven route optimization, energy mindful leader selection, and zone style routing management. In other words, the reinforcement learning agent keeps checking the state of the network; residual energy, link stability, node mobility, and traffic load are among the things it watches, so it can pick forward paths that are more dependable and, at the same time, dodge the unstable ones. So, even when mobility changes on the fly, route failures and packet losses are noticeably lowered, not just a little.

Furthermore, the zone-based routing architecture reduces the routing workload by confining routing operations to localized regions. This lowers the routing overhead that usually appears due to repeated route discoveries. As a result, route recovery feels quicker, and packet forwarding stays steadier, which then boosts communication reliability and pushes the packet delivery performance up.

Table 10: Packet Delivery Ratio (PDR) across Different Node Speeds

Node Speed (m/s)	AODV (%)	AOMDV (%)	RML-ZEREM (%)	Improvement over AODV (%)	Improvement over AOMDV (%)
3	51.46	59.45	65.12	26.57	9.53
5	55.00	62.00	75.68	37.60	22.06
7	50.00	48.00	55.32	10.64	15.25
9	60.00	55.00	78.45	30.75	42.64

Figure 13 shows the variation of PDR with respect to node speed. Although, greater mobility in general raises the chance of route breaks and link failures, but RML-ZEREM still holds the top PDR compared with every other routing protocol that was tested. Those PDR gains indicate that the routing framework actually adapts well to topology changes and consistently achieves packets in a reliable manner in those highly dynamic MANET environments.

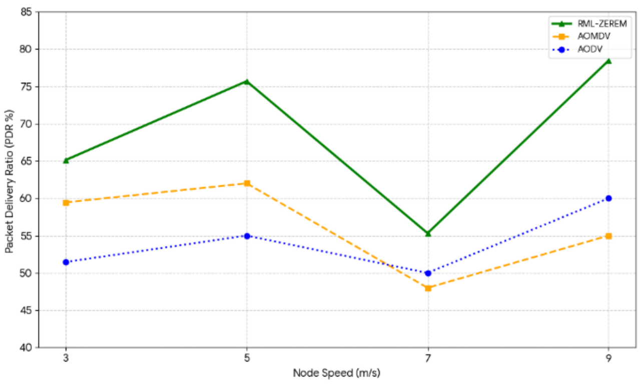

Figure 13: Packet Delivery Ratio (PDR) Performance at Different Node Speeds

6.3 Scenario 3: Impact of Network Size

The network size really affects MANET performance a lot, because when there are more nodes around, the routing gets more complicated, there's also bigger channel contention, and of course, extra communication overhead. To check how scalable the proposed routing framework is, the number of nodes was tuned from 50 to 100. After that, the effect of the overall network size on energy consumption, throughput, and packet delivery performance was looked at, then contrasted with AODV and AOMDV.

6.3.1 Energy Consumption Analysis

Table 11 shows the energy consumption behaviors of AODV, AOMDV, and the proposed RML-ZEREM protocol when the network size changes from 50 to 100 nodes. When the count of nodes gets higher, the energy usage usually goes up too, because more packet relay work is happening, plus routing overhead, and these control exchanges that keep happening in the background. In bigger networks, the competition for the channel is also higher, and the communication gets more tangled, so energy keeps getting spent. More precisely, RML-ZEREM energy consumption grows from 65.16 J at 50 nodes to 165.53 J at 100 nodes, while AODV goes from 97.51 J to 180.59 J and AOMDV from 80.88 J to 180.95 J over the same span. For the proposed method, the reduction in energy use is 33.17% versus AODV and 19.44% versus AOMDV at 50 nodes. Then at 75 nodes, the gains are still present, 15.79% over AODV and 8.64% over AOMDV. At 100 nodes, it doesn't drop to zero; the protocol still reports reductions of 8.34% and 8.52% compared with AODV and AOMDV, respectively.

The superior energy efficiency of RML-ZEREM is mainly caused by the combination of Q-learning-based route optimization, energy-aware leader selection, and zone-based routing management together. The superior performance of RML-ZEREM can be attributed to the reinforcement learning agent, which selects routes based on residual energy and link stability. In addition, the zone-based architecture confines routing operations to localized regions, thereby reducing route discovery overhead. When the network gets denser, these same mechanisms help spread traffic in a more balanced way and also reduce those unneeded control packet transmissions, so scalability improves, and energy usage goes down.

Figure 14 shows the effect of network size on energy consumption. Even if energy consumption goes up when the number of nodes increases for all routing protocols, RML-ZEREM still keeps the lowest energy usage across every network size. The energy savings that we see confirm the proposed routing framework really works for scalable and energy-efficient communication, especially in large MANET setups.

The results seem to confirm that the proposed routing framework works well as the network size goes up, like it scales effectively, but still keeps lower energy consumption and gives better resource use than conventional MANET routing protocols do. In other words, it's more efficient while staying flexible when the nodes grow in number.

Table 11: Energy Consumption Analysis based on Network Size (Number of Nodes)

Number of Nodes	AODV (J)	AOMDV (J)	RML-ZEREM (J)	Reduction over AODV (%)	Reduction over AOMDV (%)
50	97.51	80.88	65.16	33.17	19.44
75	196.09	180.72	165.13	15.79	8.64
100	180.59	180.95	165.53	8.34	8.52

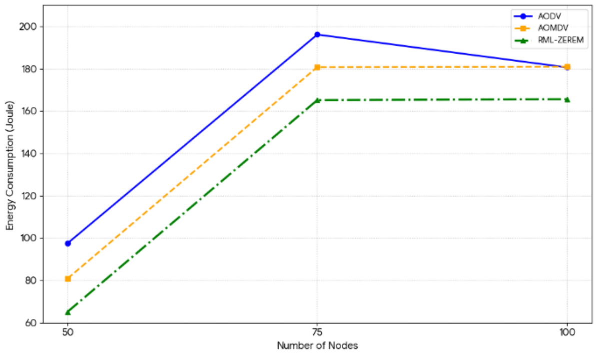

Figure 14: Impact of Network Size on Total Energy Consumption

6.3.2 Throughput Analysis

Table 12 presents the throughput performance of AODV, AOMDV, and the proposed RML-ZEREM protocol under varying network sizes ranging from 50 to 100 nodes. Throughput itself is a key performance number; it reflects the capacity of a routing protocol to successfully transmit data packets within a specified time interval. The results indicate that RML-ZEREM outperforms AODV and AOMDV under all tested network densities regardless of variations in network scale. More details, the throughput for RML-ZEREM starts around 95.34 kbps when there are 50 nodes, then it falls to roughly 78.12 kbps as you reach 100 nodes. Even though throughput does drop a bit with larger networks, because there is more routing complexity, plus extra channel contention and communication overhead, the RML-ZEREM protocol still holds that edge over the compared routing protocols. When the network size is set to 50 nodes, RML-ZEREM gives throughput gains of 2.63% and 16.63% against AODV and AOMDV. At the same time, the overall absolute throughput tends to drop at 75 nodes, mostly because of extra routing overhead and the usual channel contention. RML-ZEREM keeps a noticeably bigger relative lead over AODV, and even over AOMDV. In other words, this outcome hints that the proposed routing approach is good at handling moderate densities while still supporting smoother packet forwarding and less routing burden. Then at 100 nodes, where channel contention, routing complexity, and network congestion show up more clearly, RML-ZEREM still beats the traditional protocols. It reports throughput improvements of 23.08% relative to AODV, and 16.81% compared with AOMDV. Overall, these findings suggest the routing framework maintains reliable data delivery even when the network becomes quite dense. RML-ZEREM's superior throughput performance mainly comes from how Q-learning-based adaptive route choice is tied in with energy-aware leader picking, and also with that zone style routing control. The reinforcement learning agent keeps revising its routing moves from what the network tells it, including residual energy details, plus link stability indicators, so it can pick more dependable communication paths. On top of that, the zone-based layout limits where routing actions happen, which lowers routing overhead, so packet forwarding gets smoother, and congestion-induced drops are kept small.

Figure 15 shows how throughput changes as the network size gets larger. Even though bigger networks bring more communication complexity and higher channel contention, RML-ZEREM still keeps a higher throughput than both AODV and AOMDV. Overall, the throughput gains suggest the routing framework works well for scalable, adaptive, and trustworthy data delivery in changing MANET scenarios.

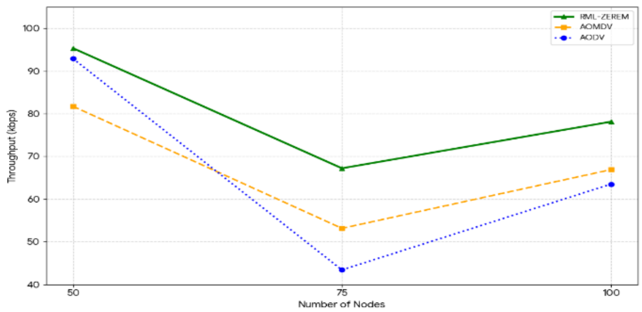

Figure 15: Throughput Analysis for Scalable Network Sizes

The variation in throughput with increasing network size is influenced by factors such as dynamic node distribution, channel contention, and route availability. Consequently, throughput does not necessarily increase linearly with node density.

Table 12: Throughput Comparison for Different Network Sizes

Number of Nodes	AODV (kbps)	AOMDV (kbps)	RML-ZEREM (kbps)	Improvement over AODV (%)	Improvement over AOMDV (%)
50	92.90	81.73	95.34	2.63	16.63
75	43.40	53.14	67.21	54.89	26.45
100	63.50	66.89	78.12	23.08	16.81

6.3.3 Packet Delivery Ratio (PDR) Analysis

Table 13 shows the Packet Delivery Ratio (PDR) results for AODV, AOMDV, and the suggested RML-ZEREM protocol for different network sizes. From what we can see, RML-ZEREM keeps reaching better packet delivery performance than the comparison routing protocols in each mobility scenario, without exception. As mobility goes higher, every routing protocol shows some wavering PDR, mostly because the topology keeps changing, routes get broken more often, and the links become unstable. Still, the proposed protocol keeps its edge, adapting to those shifting network conditions in a steadier manner. RML-ZEREM ends up with PDR gains that fall between 8.30% and 15.15% when matched against AODV, and also between 4.14% and 9.00% compared with AOMDV at the mobility levels tested. The enhanced packet delivery performance is mostly linked to how Q-learning-based route optimization gets merged with energy-aware zone management. In practice, the reinforcement learning agent keeps checking what the network looks like, for example, residual energy, link stability, node motion, and also traffic load, then it selects dependable forwarding paths and tries to sidestep routes that become shaky. In addition, this zone-oriented routing design keeps routing work more local, so the overhead of route handling stays lower, which then makes route maintenance smoother and also cuts down packet loss.

Figure 16 shows how PDR changes as node speed goes up. Even if higher mobility brings extra routing problems, RML-ZEREM stays more effective in terms of packet delivery than AODV and AOMDV throughout the whole range. Overall, the findings suggest the proposed routing approach supports better throughput and stronger packet delivery under different operating conditions. In the end, it supports reliable communication and more stable routing performance inside very dynamic MANET settings.

The consistently higher PDR values show that the proposed reinforcement-learning setup actually adapts well to mobility-caused topology changes, and it keeps packet delivery reliable even when the network is changing all the time.

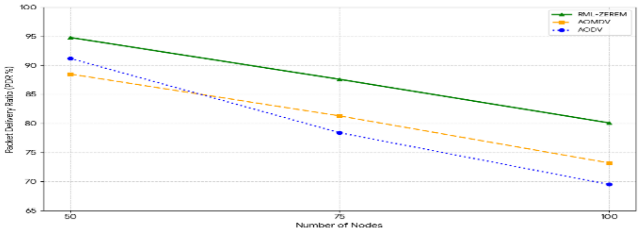

Figure 16: Packet Delivery Ratio (PDR) Trend across Different Node Counts.

Table 13: Impact of Network Density on Packet Delivery Ratio (PDR)

Number of Nodes	AODV (%)	AOMDV (%)	RML-ZEREM (%)	Improvement over AODV (%)	Improvement over AOMDV (%)
50	91.20	88.50	94.80	3.95	7.12
75	78.40	81.30	87.60	11.73	7.75
100	69.50	73.20	80.10	15.25	9.43

6.4 Scenario 4: Impact of Area Size

This scenario checks how several network area sizes mess with the energy consumption performance of AODV, AOMDV, and the proposed RML-ZEREM protocol. During the simulation, the coverage plane is moved around, from 500 × 500 m^2 up to 1000 × 1000 m^2, while the count of nodes stays the same. When the network area gets larger, the node density drops, so the distance between nodes becomes greater, connectivity turns weaker, and routing becomes more complicated. In other words, the node connectivity within the network decreases, and routing protocols require additional overhead to discover valid transmission paths.

6.4.1 Impact of Area Sizes on Energy Consumption

Table 14 shows the energy consumption results for various simulation area sizes, and yes, the numbers look pretty clear. In general, RML-ZEREM tends to use less energy than both AODV and AOMDV in all network arrangements, even the layout changes a bit. For example, when the area size is 500 × 500 m^2, the proposed protocol delivers energy consumption reductions of 13.45% and 8.16% relative to AODV and AOMDV, respectively. Similar improvements are observed for larger network areas, indicating that the proposed routing framework maintains robust performance under varying node distribution conditions. This better energy efficiency of RML-ZEREM is mostly coming from the way Q-learning-based route optimization gets combined with energy-aware zone management. The reinforcement learning agent keeps looking for dependable communication paths that are also energy-efficient. Meanwhile, the zone-driven design reduces unnecessary route discovery, and it also trims down routing control overhead. Because of that, the protocol ends up with fewer retransmissions, and it sidesteps the extra energy drain that usually happens when routing stays unstable.

Figure 17 presents how energy consumption changes as the simulation area gets larger. Overall, energy consumption seems to drop little by little for every protocol as the network area expands. This is probably because node contention becomes lower, and collision probability decreases too, especially in less crowded environments. Nevertheless, RML-ZEREM maintains the lowest energy consumption across all test cases, demonstrating its capability to efficiently manage routing tasks and energy resources under both dense and sparse MANET deployment scenarios.

The results confirm that the proposed RML-ZEREM protocol works well even when node distribution shifts and when network density changes, so it keeps a higher kind of energy efficiency along with better routing performance across different deployment areas, in general.

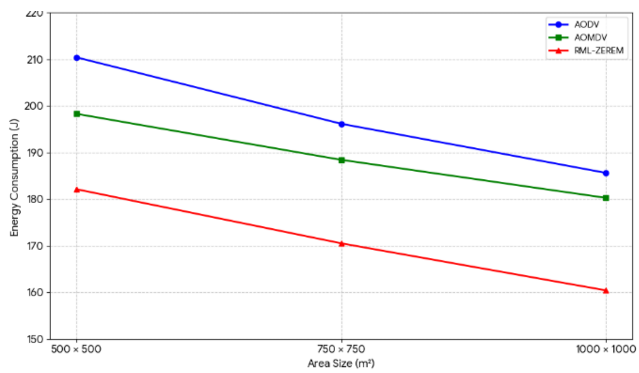

Figure 17: Energy Consumption Trends for Various Simulation Area Sizes.

Table 14: Energy Consumption vs. Simulation Area Size

Area Size	AODV (J)	AOMDV (J)	RML-ZEREM (J)	Reduction over AODV (%)	Reduction over AOMDV (%)
500×500 m^2	210.45	198.32	182.14	13.45	8.16
750×750 m^2	196.18	188.44	170.52	13.08	9.51
1000×1000 m^2	185.63	180.27	160.41	13.58	11.02

6.4.2 Throughput Analysis

Table 15 shows the throughput performance of AODV, AOMDV, and the proposed RML-ZEREM protocol for different network area sizes, from 500×500 m^2 up to 1000×1000 m^2. Throughput is a key performance metric, as it reflects the efficiency of data delivery across the network. Overall, RML-ZEREM keeps landing as the top performer, consistently higher throughput than the other evaluated routing protocols across all the simulation area sizes. When the simulation area gets bigger, the throughput of the proposed protocol also climbs clearly. In particular, RML-ZEREM posts throughput gains of 26.30%, 62.05%, and 89.84% relative to AODV for network areas of 500×500 m^2, 750×750 m^2, and 1000×1000 m^2, respectively. In the same settings, it also beats AOMDV by 115.32%, 98.63%, and 80.19%, which is pretty strong.

This kind of throughput boost seems mostly driven by reduced network congestion and less channel rivalry when the deployment area grows. As node density goes down, there is a smaller chance of packet collisions and communication disturbance, so packet forwarding feels more trustworthy and steadier. Plus, the Q-learning-based routing mechanism keeps refining route choice based on what the network is doing, and it manages to spot dependable communication routes that support better delivery results. Also, the zone-based routing structure kind of fences routing work into nearby regions only, so the routing burden drops and the end-to-end transmission efficiency goes up. With adaptive route learning and zone handling working together, RML-ZEREM can keep a superior throughput level even when the node placement patterns change around a bit.

Figure 18 illustrates the throughput variation as the network area size increases. As the deployment area increases, throughput improves for all protocols; however, RML-ZEREM consistently maintains a significant performance advantage over both AODV and AOMDV. Overall, these findings highlight that the proposed routing framework actually works well for scalable data delivery, and it stays reliable across various MANET deployment setups.

The increasing throughput trend indicates that the proposed reinforcement learning framework effectively adapts to variations in node distribution while maintaining efficient packet forwarding, thereby enhancing network utilization and overall communication performance.

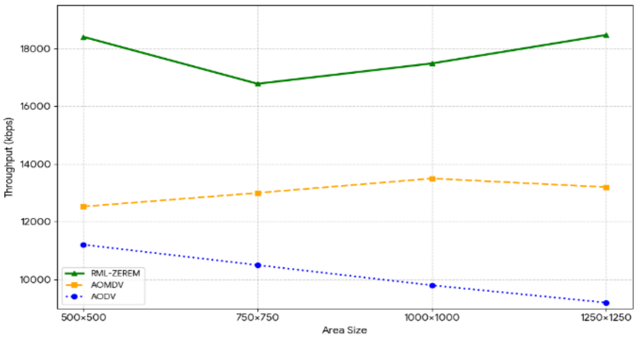

Figure 18: Throughput Performance across Different Network Dimensions

Table 15: Throughput Performance across Various Area Dimensions

Area Size	AODV (kbps)	AOMDV (kbps)	RML-ZEREM (kbps)	Improvement over AODV (%)	Improvement over AOMDV (%)
500×500 m^2	7902.54	4633.37	9979.89	26.30	115.32
750×750 m^2	7382.03	6021.81	11962.20	62.05	98.63
1000×1000 m^2	7196.04	7580.50	13660.06	89.84	80.19

6.4.3 Packet Delivery Ratio (PDR) Analysis

Table 16 shows Packet Delivery Ratio (PDR) results for AODV, AOMDV, and the proposed RML-ZEREM protocol when the network area varies from 500×500 m^2 up to 1000×1000 m^2. When the simulation area gets larger, but the node count stays the same, the node density drops and the typical communication distance between nodes gets longer. Due to this, the links are less steady, route breaks happen more often, and the packet delivery results slowly go down for every routing approach. Even with that, RML-ZEREM keeps producing higher PDR values than AODV and AOMDV for all the scenarios we test. In the case of 500×500 m^2, RML-ZEREM boosts PDR by 3.57% versus AODV, and by 6.20% versus AOMDV. After that, as the deployment area keeps expanding, the gap gets even clearer. At 1000×1000 m^2, RML-ZEREM delivers PDR gains of 11.77% and 7.24% over AODV and AOMDV, respectively. The superior packet delivery performance of RML-ZEREM can be explained by the way it blends Q-learning-based adaptive routing with energy-aware zone management, yes, kind of both, at the same time. A reinforcement learning agent keeps checking the network state, then it chooses dependable communication routes by looking at route stability, remaining energy, and how well the network is connected. Also, the zone-oriented routing structure actually trims down routing complexity and makes route maintenance a bit more efficient. So overall, the protocol becomes more capable in sparse and fast-changing MANET settings.

Figure 19 shows how PDR shifts when the network area size keeps getting larger. Even though every routing approach shows a slow decrease in packet delivery performance as the nodes are deployed over a wider space, RML-ZEREM stays on top and keeps the highest PDR values. This indicates that the suggested routing framework responds well to changes in how nodes are distributed and how the network topology evolves, while still keeping packet delivery dependable and communication performance stable.

The rather consistently higher PDR values show that the suggested reinforcement learning setup manages to learn rather stable routing behaviors and keep adapting even when the network is sparse. So, you end up with fewer packet losses, and the whole routing reliability gets better across the various deployment areas.

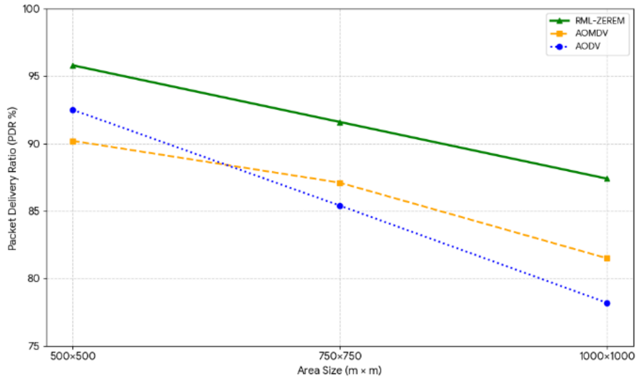

Figure 19: Packet Delivery Ratio (PDR) Analysis based on Area Size.

Table 16: Packet Delivery Ratio (PDR) vs. Network Area Size

Area Size	AODV (%)	AOMDV (%)	RML-ZEREM (%)	Improvement over AODV (%)	Improvement over AOMDV (%)
500×500 m^2	92.50	90.20	95.80	3.57	6.20
750×750 m^2	85.40	87.10	91.60	7.26	5.17
1000×1000 m^2	78.20	81.50	87.40	11.77	7.24

6.5 Q-Learning Convergence Analysis

The convergence behavior of the proposed Q-learning mechanism is analyzed by monitoring the average cumulative reward throughout training episodes. In the initial training phase, the routing agent explores diverse routing alternatives, leading to frequent fluctuations in reward values. As the training proceeds, Q-values gradually converge to stable states, and the average reward eventually stabilizes. This suggests that the agent actually learned energy-efficient routing paths. Overall, the convergence pattern confirms that this reinforcement learning framework can adapt well to changing MANET environments, while still keeping stable routing performance.

7 Conclusion

The results from the performance evaluation show that the suggested Reinforcement Learning driven Multi-Leader Zone Energy-Efficient Routing Protocol (RML-ZEREM) really enhances MANET routing performance when we compare it to the standard AODV and AOMDV protocols. By mixing in Q-learning, energy cognizant leader node selection, and zone driven routing, the method supports more adaptive route optimization even when the network keeps changing, or in other words, dynamic network conditions. Overall, the proposed protocol tends to lower energy usage, while at the same time it boosts throughput and Packet Delivery Ratio (PDR) across several simulation situations, with different simulation times, node mobility settings, network sizes, and deployment zones. What makes RML-ZEREM stand out most, is that it can keep learning routing choices on the fly, using residual energy, link steadiness, traffic conditions, and the way the topology changes. Also, the zone-based structure cuts down routing overhead, because it confines routing operations to smaller areas, so the protocol becomes more scalable and resource efficient in bigger MANET deployments. The simulation findings therefore indicate that the proposed protocol can keep communications stable, and still reliable, even under very high dynamics in the network. Even if the proposed protocol brings in extra computational overhead, mainly because of Q-learning operations plus ongoing route evaluation, the net gains still seem worth it, in terms of energy efficiency, higher throughput, and better routing reliability. So, overall RML-ZEREM looks like a quite promising routing option for energy-constrained and dynamic MANET scenarios, even with that extra processing cost hanging around. One caveat in the current study is that the protocol was only tested against more conventional MANET routing approaches like AODV and AOMDV. For future work, we plan to compare it with more recent reinforcement learning-driven, energy-aware, and zone-based routing protocols, so we can validate the proposed framework's real-world effectiveness even more, and in a broader way.

8 Future Work

Future work will focus on extending the proposed framework by incorporating Deep Reinforcement Learning (DRL) and Multi-Agent Reinforcement Learning (MARL) techniques to enhance learning efficiency and accelerate convergence in large-scale MANET environments. At the same time, adaptive zone formation mechanisms, and secure routing strategies will be explored too, so the network stays more resilient when faced with malicious attacks and also those constant, dynamic topology changes. The routing framework that we propose, will be tested further in up-and-coming wireless networking environments including Internet of Things (IoT), Vehicular Ad Hoc Networks (VANETs), Unmanned Aerial Vehicle (UAV) networks, and 5G/6G-enabled communication systems.

Funding

This research received no external funding.

Author Contributions

Rani Sahu: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Visualization, Writing – original draft, Writing – review & editing. Dr. Babita Rathore: Supervision, Validation, Resources, Writing – review & editing. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

All the authors declare that they have no conflict of interest.

References

[1] Abdullah, A.M.: An efficient approach for minimizing control packets and enhancing route stability in mobile ad-hoc networks. The Journal of Supercomputing **82**(5), 305 (2026). https://doi.org/10.1007/s11227-026-08480-y

[2] Saxena, P., Phade, G.: Analysis of routing protocols adopted in FANET and RANET communication network. Wireless Networks **32**, 1131–1157 (2026). https://doi.org/10.1007/s11276-026-04099-2

[3] Zhu, M., Zhou, Y., Chu, J., Chen, Z., Ai, H., Li, J., Han, J.: I-AOMDV: QoS-aware multipath routing protocol for enhanced in dynamic multi-node vehicle area network. China Communications **23**(3), 286–297 (2026). https://doi.org/10.23919/JCC.fa.2025-0186.202603

[4] Harshitha, V.S., Surekha, V.R., Ananthi, V.: A Comprehensive Review on Q-Learning Based Optimal Routing Protocol for MANET. In 2026 International Conference on Machine Learning and Autonomous Systems (ICMLAS), pp.1754–1758 (2026). https://doi.org/10.1109/ICMLAS67792.2026.11483818

[5] Tonin, A., Algabri, M.N.A., Ali, M.N., Ghurab, M., Al-Khulaidi, A.A.G., Al-Baltah, I.A.: A novel energy model for optimizing energy efficiency and prolonging network lifetime in MANET. IEEE Access **13**, 117771–117787 (2025). https://doi.org/10.1109/ACCESS.2025.3586795

[6] Shah, I.A., Ahmed, M.: Survey on protocols in MANETs, WSNs and DTNs. International Journal of System Assurance Engineering and Management **16**(6), 2089–2120 (2025). https://doi.org/10.1007/s13198-025-02760-1

[7] Yakubu, S., Mindila, A., Kihato, P.: Design of Self-Healing Mesh Architecture: A Proof-of-Stake AODV Routing Protocol With Autonomous Adaptability. Engineering Reports **8**(2), e70656 (2026). https://doi.org/10.1002/eng2.70656

[8] Wijonarko, D., Arifin, S., Faisal, M., Pratama, M.N., Priambodo, O.N., Nugraha, E.S.: Mobile Ad-Hoc Network (MANET) Method: Some Trends and Open Issues. Recent in Engineering Science and Technology **3**(02), 49–74 (2025). https://doi.org/10.59511/riestech.v3i2.108

[9] Maray, M.: Hybrid deep learning techniques for adaptive routing and congestion control in urban VANET for wireless mobile networking. Scientific Reports **16**, 16046 (2026). https://doi.org/10.1038/s41598-025-33193-2

[10] Mazhar, T., Guizani, S., Hamam, H.: Deep Reinforcement Learning and IoT for Renewable Energy Optimization in Smart Buildings: A Comprehensive Review. IET Generation, Transmission & Distribution **20**(1), e70255 (2026). https://doi.org/10.1049/gtd2.70255

[11] Vivekanandan, S.D., Sharmila, P., Dhakshnamoorthy, M., R., S.: MANETs and WSNs in Hybrid Genghis Khan Shark and Lotus Effect Optimizer for Secure Routing. International Journal of Communication Systems **39**(6), e70461 (2026). https://doi.org/10.1002/dac.70461

[12] Uma, B., Sumathi, S.: Secure routing and attack detection in mobile ad hoc networks via bidirectional cascade residual convolutional neural networks. OPSEARCH, 1–31 (2026). https://doi.org/10.1007/s12597-026-01157-3

[13] Singh, S.B., Rizvi, M.A., Saxena, K., Gupta, R., Tripathi, A.N., Dewangan, N.K.: An adaptive, energy-efficient and secure routing protocol for zone-related mobile Ad-hoc networks using reinforcement learning, Scientific Reports **16**, 3002 (2026). https://doi.org/10.1038/s41598-025-32918-7

[14] Selim, I.M., Abdelrehem, N.S., Alayed, W.M., Elbadawy, H.M., Sadek, R.A.: MANET Routing Protocols' Performance Assessment Under Dynamic Network Conditions. Applied Sciences **15**(6), 2891 (2025). https://doi.org/10.3390/app15062891

[15] Sarkar, N.I., Ali, M.J.: A study of MANET routing protocols in heterogeneous networks: a review and performance comparison. Electronics **14**(5), 872 (2025). https://doi.org/10.3390/electronics14050872

[16] Shakya, A., Ali, S., Nand, P.: Simulation-Based Performance Analysis of MANET Routing Protocols: Proactive vs. Reactive. In 2025 Optical Communication, Photonics, Telecommunications, and Intelligent Machine Applications (OPTIMA), pp. 348–353 (2025). https://doi.org/10.1109/OPTIMA67660.2025.11380316

[17] Karthikeyan, M., Manimegalai, D., Rajagopal, K.: Optimizing energy efficiency in wireless sensor networks: dynamic routing with capuchin search algorithm. Multimedia Tools and Applications **84**(15), 15453–15477 (2025). https://doi.org/10.1007/s11042-024-19625-7

[18] Tegulapalle, V.R., Gurram, R., Thangam, S., Kumari, J.J.J.: Optimizing Rural Healthcare Monitoring Using MANET: A Comparative Analysis of AODV, DSDV, DYMO, and OSLR. In International Conference on Power Engineering and Intelligent Systems (PEIS), pp. 218–294 (2025). https://doi.org/10.1007/978-981-96-9716-8_22

[19] Ghanem, M., Sabaliauskaite, G., Correia-Hopkins, S., Jones, J.-L., Micallef, N.: A dynamic simulation framework for mobile ad hoc networks in search and rescue operations. Simulation Modelling Practice and Theory **149**, 103281 (2026). https://doi.org/10.1016/j.simpat.2026.103281

[20] Tirumalasetti, R., Singh, S.K., Roy, P.K., Mishra, S.: A Systematic Review of VANET Routing Protocols for Intelligent Transport Systems (ITSs). Journal of Advanced Transportation **2026**(1), 7999623 (2026). https://doi.org/10.1155/atr/7999623

[21] Sahu, R., Yogi, S., Khan, N., Nigam, R.K., Vyas, S., Danger, B.: "Optimizing MANET Routing with SINTM: A Smarter Load Balancing Approach", in Emerging Perspectives and Applications of Computational Intelligence and Smart Systems, CRC Press (2026).

[22] More, A.P., Kale, R.S., Rizvi, M.: Enhancement in Performance Metrics for Mobile Ad Hoc Networks (MANETs) Using Hybrid Ant Lion Optimization (HALO) Algorithm. Journal of Electrical and Computer Engineering **2026**(1), 9947624 (2026). https://doi.org/10.1155/jece/9947624

[23] Abdullah, A.M.: Hybrid energy-efficient routing protocol for extended network lifetime in wireless body area networks. The Journal of Supercomputing **82**(3), 130 (2026). https://doi.org/10.1007/s11227-026-08277-z

[24] Yang, H., Bai, W., Shi, Y., Zhou, D., Sheng, M., Li, J.: Optimal Zone Routing Scheme for LEO Mega-Constellation Networks. IEEE Transactions on Mobile Computing **25**(8), 13116–13129 (2026). https://doi.org/10.1109/TMC.2026.3674959

[25] Sahu, R., Sahu, V., Joshi, S., Kumar, S., Malviya, S., Kakkad, A.: "Networks Design and Evaluation of a Zone-Based Energy-Efficient Routing Protocol for Hybrid Wireless Networks", in Emerging Perspectives and Applications of Computational Intelligence and Smart Systems, CRC Press, (2025).

[26] Sahu, R., Sharma, S., Rizvi, M.: ZBLE: zone based leader election energy constrained AOMDV routing protocol. International Journal of Computer Networks and Applications **6**(3), 39–46 (2019). https://doi.org/10.22247/ijcna/2019/49643

[27] Sahu, R., Sharma, S., Rizvi, M.A.: ZBLE: Zone based efficient energy multipath protocol for routing in mobile Ad Hoc networks. Wireless Personal Communications **113**(4), 2641–2659 (2020). https://doi.org/10.1007/s11277-020-07345-8

[28] Zelani, S.K., Ramakrishna, K.V.S.S., Asiri, F., Alyahya, A., Basheer, S.: A Reinforcement Learning–Driven Payoff-Adaptive Game-Theoretic Framework for Secure and Reliable Operation of Mobile Ad Hoc Networks. Transactions on Emerging Telecommunications Technologies **37**(4), e70394 (2026). https://doi.org/10.1002/ett.70394

[29] Kurkina, N., Papaj, J., Badar, J.: Enhancing routing efficiency in Cloud MANET using KNN and Fitness Function for dynamic network environments. Peer-to-Peer Networking and Applications **19**(2), 40 (2026). https://doi.org/10.1007/s12083-026-02198-7