International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Reinforcement Machine Learning-based Improved Protocol for Energy Efficiency on Mobile Ad-Hoc Networks

Mrs. Biradar Ashwini Vishwanathrao *[1], Dr. Pradnya Ashish Vikhar [2]

Submitted: 09/10/2023 **Revised**: 27/11/2023 **Accepted**: 07/12/2023

Abstract: Mobile Ad-Hoc Networks (MANETs) are crucial in environments lacking permanent infrastructure, with energy efficiency being a primary concern due to the reliance on battery-powered devices. This study presents an innovative solution: the Reinforcement Machine Learning-enhanced Energy Efficient AODV (Ad-Hoc On-Demand Distance Vector) Protocol (RML-EEAODV). This novel approach integrates the adaptive capabilities of reinforcement machine learning with the AODV routing protocol to forge a smart, energy-conserving routing mechanism. The core challenge in MANETs is minimizing energy use and operational overhead while ensuring optimal packet delivery. RML-EEAODV addresses this by enhancing the AODV protocol's routing decisions. It employs machine learning to enable nodes to maintain and utilize a dynamic database of state information for intermediate nodes along potential routes. This database informs decision-making for forwarding packets, ensuring routes with guaranteed Quality of Service (QoS). The RML-EEAODV protocol significantly improves energy efficiency and reduces network overhead, while maintaining a satisfactory packet delivery ratio.

Keywords: *Mobile ad-hoc network, Reinforcement learning, K-Means Clustering, Machine Learning, Clustering, Ad-hoc on demand distance vector etc*

1. Introduction

The Mobile Ad-Hoc Network (MANET) represents a significant advancement in wireless communications, characterized by its independence from traditional network infrastructure. In MANETs, the absence of pre-installed infrastructure like fixed routers fundamentally changes the way data packets are transmitted. These networks, also known as mobile packet radio networks or mobile multihop wireless networks, offer a novel approach to providing network services where no established system exists. The term "ad hoc" in this context implies a system that operates without pre-established structures, not one that is makeshift or improvised.

As illustrated in Figure 1, the architecture of a MANET allows nodes to communicate directly within their wireless range. However, due to limitations such as signal attenuation, environmental noise, and restricted battery life, wireless networks like MANETs often have lower capacity and range compared to wired networks. Consequently, transferring data across the network can necessitate multiple hops from one node to another. This multi-hop nature of MANETs requires each node to function dually as a host and a router, taking on responsibilities for routing, packet forwarding, and executing various network functions autonomously. This unique operational mode sets MANETs apart from traditional wireless networks and underlines their versatility and adaptability in environments lacking conventional network infrastructure.

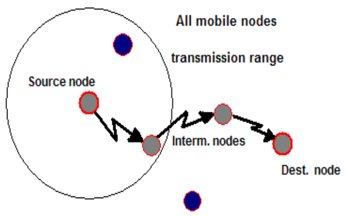

Fig 1. General structure of MANET

2. Routing protocols for Mobile Ad-Hoc Networks

The following part provides a concise explanation of the current AODV and DSDV protocols, as well as the Reinforcement Machine Learning-based Improved AODV (RMLB-AODV) protocols that have been developed.

2.1 Adhoc On-Demand Distance Vector (AODV) Routing Protocol

A routing protocol known as Ad-hoc On-Demand Distance Vector (AODV) was developed so that it may be used in Mobile Ad hoc Networks (MANETs). The AODV protocol makes it feasible to have a multi-hop routing that is both dynamic and self-starting between mobile nodes. Due to the fact that routes are only built when they are

[1,2]*Department of Computer Science and Engineering*
[1] *Research Scholar, Dr. A. P. J. Abdul Kalam University, Indore*
[2] *Research Supervisor, Dr. A. P. J. Abdul Kalam University, Indore*
E-mail Id: ashwinibiradar29@gmail.com
** Corresponding Author: Mrs. Biradar Ashwini Vishwanathrao*
Email: ashwinibiradar29@gmail.com

necessary, rather than on a regular basis, AODV is able to deliver fast dynamic routing.

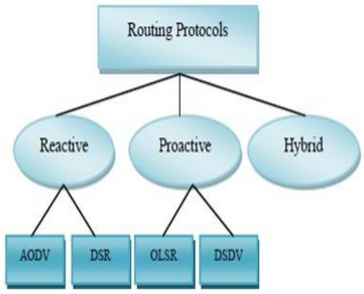

Fig 2. Routing Protocols including AODV

In order to determine the routes that may be taken between the nodes that make up a network, the Ad Hoc On-Demand Distance Vector (AODV) protocol makes use of a Route Discovery Process. The first step in this procedure involves a source node sending out a Route Request (RREQ) message to the nodes that are physically near to it. As soon as a node receives an RREQ, it examines whether or not it is in possession of a route that leads to the requested destination. In the event that it does not, the request is sent on the next step. In situations in which a node already has a route to the destination, it will send a Route Reply (RREP) back to the source node along the path that has been created. It is possible for the source node to create a route to the destination if it is able to successfully receive an RREP. It is possible for the source node to raise the Time-to-Live (TTL) value and reissue the RREQ in the event that an RREP is not received. This process will continue until either an RREP is acquired or the TTL hits zero.

2.2 Machine Learning-Enhanced Routing in MANETs

In Mobile Ad-Hoc Networks (MANETs), the absence of a centralized controller can compromise the security and efficacy of routing techniques compared to traditional networks. Machine learning algorithms can enhance MANET routing by learning from and adapting to environmental changes within the network. These algorithms can help in threat detection, predicting network traffic flow, and refining routing protocols for increased efficiency. Machine learning models can be trained to recognize shifts in traffic patterns and network structures, enabling MANET protocols to quickly adapt and fortify against such changes. By employing machine learning, mobile networks can achieve reliable, low-latency communication and provide a secure data transmission platform.

2.2.1 Reinforcement Machine Learning-based Improved Energy Efficient AODV (RML-EEAODV) Protocol

The RML-EEAODV protocol integrates reinforcement learning algorithms at each node in the network, based on the AODV framework. It utilizes a State Information Database (SIDB) containing data about all nodes along potential routes, focusing on two key metrics: maximum traffic load and energy consumption of the nodes. The traffic load at a node is quantified by the number of data packets in its queue. This database informs forwarding decisions to establish routes with guaranteed Quality of Service (QoS), considering the node's remaining energy and traffic load to minimize unnecessary route request rebroadcasting, which otherwise consumes bandwidth and energy.

When a source node requires a route to a destination with specific energy needs, it broadcasts an RREQ based on the current energy status and traffic load of neighboring nodes. Nodes that do not meet the energy and traffic load requirements refrain from rebroadcasting the RREQ. Conversely, nodes with sufficient energy and lower traffic load create a reverse route entry, aiding in forwarding the RREP to the source node. As packets are forwarded, each node along the path assesses its energy and traffic load before transmitting, ensuring efficient energy use and reducing network overhead. The RML-EEAODV protocol thus significantly improves energy consumption and network overhead while maintaining an acceptable packet delivery ratio.

3. Literature review

The objective of this study is to conduct a comprehensive evaluation of the existing literature on medium access control (MAC) algorithms for wireless ad hoc networks (WANETs) that are based on reinforcement learning (RL). The current traditional media access control (MAC) methods are inadequate in meeting the demands of the expanding scale of wireless ad hoc networks (WANETs). These systems are facing challenges like as changing topology, resource allocation, interference, limited bandwidth, and environmental constraints. This research highlights the need for more sophisticated MAC protocols in WANETs to adapt to the dynamic characteristics of these networks. After presenting the basic RL techniques, we will examine advanced MAC protocols for WANETs and analyse how these protocols specifically tackle challenges inside the MAC model. This portion of the study concludes by examining prospective research domains that may be explored to enhance the efficacy of the MAC procedure [1].

The Internet of Things (IoT) is undergoing fast expansion, giving rise to many crucial concerns. The issues included in this context are heterogeneity, reliability, and scalability. They emerge among a diverse array of distinct wireless devices that operate simultaneously. A novel framework has been developed to address these challenges by integrating Internet of Things (IoT) with Software-Defined Wireless Sensor Networks (SDWSN). The term used to describe this architecture is SDWSN-IoT. The goal of this research is to examine a distinct, intelligent, and energy-efficient multi-objective routing protocol for Internet of Things (IoT) networks. The system uses Reinforcement Learning with Dynamic Objective Selection (DNS-RL) as its foundation. The limited energy resources of wireless Internet of Things devices provide a considerable problem in optimising energy use. The primary objective of the DOS-RL is to provide adaptability to network modifications while also maximising energy efficiency. In order to expedite the process of learning, the technique utilises carefully crafted related objectives and incentives. Comparative simulations suggest that DOS-RL surpasses common protocols like OSPF and multi-objective Q-routing in SDN-Q in terms of energy economy and speedy flexibility [2].

This paper introduces a novel vertical routing method for flying ad hoc networks (FANETs), which are a subset of the 5G access network consisting of highly mobile unmanned aerial aircraft. The system is based on a new deep Q-network (DQN) approach. The objective of this research is to tackle the problem of frequent link disconnections and network partitions in FANETs, aiming to enhance the overall performance of the network. To facilitate the administration of both global and local information, our technology integrates a central controller (CC) with distributed controllers (DCs) deployed across several network planes (macro, pico, and femto). The strategy focuses on routing based on residual energy and mobility rates, an approach that has not been thoroughly investigated before. The idea enables the establishment of clusters on different network planes, which in turn promotes connections both inside and between these clusters. As a result, data traffic is efficiently distributed across the network. Our technique based on Deep Q-Network (DQN) has shown significant superiority over traditional Reinforcement Learning (RL) methods in enhancing network longevity, decreasing energy use, and minimising link failures [3].

Ad hoc car networks are becoming recognised as a viable option for smart city communication, given the ongoing progress of intelligent transportation systems. However, the increasing popularity of wireless technology in highly mobile contexts presents some intriguing concerns. This paper proposes reinforcement learning (RL) as a viable solution to the challenges, particularly in the field of routing. Our purpose is to create a complex objective space that considers the geographic placement of cars, the signal strength, and the environmental path loss, which includes city maps and barriers. The main focus of this technique is to optimise both the stability of the route and the number of hops. The results suggest a significant improvement in the robustness of the route when compared to traditional protocols and other reinforcement learning approaches that rely on a single decision-making parameter [6].

In the wake of the current outbreak, there has been a surge in demand for wireless communication to support vital operations in remote and inaccessible areas. This study explores the potential of aerial ad hoc networks (AANETs) created by flying vehicles to accomplish various tasks. While AANETs are recognised for their high level of mobility, they are vulnerable to issues such as connection disruptions, energy depletion, and packet loss. Given that AANETs rely mostly on batteries for power, it is crucial to consider factors like as flight length and speed. By combining deep learning with NS3 simulation, we provide an energy-efficient and resilient deep scheduling technique for hello packets. This approach aims to prolong the duration of unmanned aerial vehicle (UAV) missions by conserving energy and ensuring crucial network performance criteria are maintained. Furthermore, it will provide valuable information on the geographical and temporal patterns of AANETs in selected scenarios that include three-dimensional spaces [5].

This research delves into the realm of Flying Ad-Hoc Networks (FANETs), which are a crucial subset of wireless ad-hoc networks that include unmanned aerial vehicles (UAVs) for various tasks and communications. FANETs are increasingly used in commercial and civilian domains for diverse applications such as traffic management, remote data gathering, sensing, network relaying, and product transportation. Outlined below are many critical challenges encountered by FANETs. The issues include adaptive routing protocols, flight trajectory optimisation, energy constraints, charging mechanisms, and autonomous deployment strategies. Reinforcement learning (RL) has been a prominent method in recent years because of the increased mobility and dynamic structure of FANETs. The objective of this study is to provide a comprehensive examination and comparison of the use of Reinforcement Learning (RL) in several scenarios within the Flying Ad-hoc Network (FANET) domain. These scenarios include routing, flight routes, relaying, and charging. Furthermore, it offers potential research areas that might guide future investigations in this field [6].

The advancements in vehicle communication that have taken place recently necessitate the use of efficient security measures. Security is the primary problem in Vehicle Ad Hoc Networks (VANETs). This paper aims to tackle the crucial issue of identifying rogue nodes in VANETs, with a particular focus on detecting distributed denial of service attacks. Despite the existence of other proposed solutions, this study introduces a real-time detection system that utilises machine learning. We provide a distributed multi-layer classifier that has been verified by OMNET++ and

SUMO simulations, utilising machine learning classifiers like GBT, LR, MLPC, RF, and SVM. The simulation shows significant progress in assault categorization, with a classification accuracy of up to 99%. This is achieved by considering both ordinary and aggressive vehicle datasets. Although the number of network nodes has increased, the use of Amazon Web Services has led to a notable improvement in network performance, especially during the training and testing phases [7].

Unmanned aerial vehicles, sometimes known as UAVs, are increasingly being used to provide wireless connections in situations when terrestrial networks are either overloaded or nonfunctional. However, the challenges that emerge include limitations on energy and interference from neighbouring cells of unmanned aerial vehicles. This study introduces a system called collaborative multi-agent decentralised double deep Q-network (CMAD–DDQN), which focuses on facilitating direct collaborative communication among unmanned aerial vehicles (UAVs). The CMAD–DDQN enables unmanned aerial vehicles (UAVs) to perform data interchange, optimise flight paths in three dimensions, and mitigate interference, user mobility, and energy constraints. The simulation results demonstrate that this technique significantly outperforms the existing baselines in terms of both system energy efficiency and network coverage. This emphasises the benefits of UAV collaboration in dynamic environments [8].

The device-to-device (D2D) link is a crucial feature of 5G technology. It enables high-speed, low-latency, energy-efficient, and spectrum-efficient peer-to-peer networking. This work introduces a multi-hop routing protocol that utilises a double deep Q learning technique, which is rooted on deep reinforcement learning. Furthermore, the system also considers energy usage. The Gannet Chimp optimisation (GCO) technique is used by the protocol to identify the most optimal approach. This technique is evaluated based on several factors, such as packet delivery ratio, latency, residual energy, throughput, and network lifetime. The results indicate that the proposed method is successful in multi-hop D2D communication scenarios, as it attains excellent performance in all evaluated metrics. The findings suggest significant improvements [9].

Cellular networks are crucial for the evolution of the Internet of Things, particularly for machine-to-machine (M2M) communication. This study investigates the difficulty of resource management in cellular networks that handle both Human-to-Human (H2H) and Machine-to-Machine (M2M) traffic concurrently. Our approach entails a network architecture that integrates simultaneous wireless information and power transfer (SWIPT) technology. This architecture is specifically developed to tackle the high energy consumption associated with machine-to-machine (M2M) communication. Under this methodology, machine type communication devices (MTCDs) are classified into two distinct categories: critical and acceptable. In addition, a resource management plan focused on optimising energy efficiency (EE) is established. A multi-agent deep reinforcement learning (DRL) approach has been developed to achieve optimal resource allocation. This technique considers the links between spectrum, power, and power splitting. In comparison to other intelligent solutions, this approach excels in terms of its rapid convergence speed and its ability to meet the criteria for Enterprise Environment (EE) and Quality of Service (QoS) [10].

In Wireless Multimedia Sensor Networks (WMSN), there is a need for a solution that is both energy-efficient and capable of providing a Quality of Service (QoS) guarantee. This is due to the complex job processing and frequent data exchanges that take place in these networks. This is especially crucial inside the sensing layer of the Internet-of-Vehicles. The presence of heterogeneity and uneven energy distribution in WMSNs poses challenges, since existing routing approaches generally neglect energy considerations while maintaining quality of service. This paper introduces an Energy-Efficient Distributed Adaptive Cooperative Routing (EDACR) model for wireless sensor networks (WMSN). Our solution focuses on finding a balance between quality of service and energy utilisation by using a reinforcement learning mechanism driven by dependability and delay metrics. Simulation results demonstrate that EDACR, in comparison to traditional and distributed adaptive cooperative routing protocols, achieves a significant decrease in energy consumption without compromising quality of service [11].

Mobile Ad-Hoc Networks (MANETs) are crucial in environments lacking a fixed infrastructure. Energy efficiency is a crucial concern in such environments due to the reliance on gadgets powered by batteries. This paper introduces a protocol called Reinforcement Machine Learning-based Improved Energy Efficient AODV (RML-EEAODV) Protocol. The goal of this protocol is to optimise the energy use of MANETs, hence optimising their efficiency. RML-EEAODV utilises reinforcement machine learning with the AODV routing protocol to provide an intelligent and flexible energy-efficient routing solution. The major goal is to minimise energy consumption and network overhead while ensuring a suitable packet delivery ratio. The RML-EEAODV architecture uses state information databases for intermediate nodes to facilitate decision-making and provide assured quality of service pathways [12].

Wireless networks, including Mobile Adhoc Networks (MANETs), are increasingly focusing on congestion management to optimise resource sharing in terms of efficiency and fairness. This study departs from traditional rule-based techniques and adopts a machine learning approach to fulfil future network needs. This is due to the demonstrated efficacy of machine learning in resolving complex problems. The disclosed technique offers a cross-layer protocol for wireless Mobile Ad hoc Networks (MANETs). This protocol encompasses the administration of traffic, the maintenance of connections, and the scheduling of concurrent transmission in a distributed manner. Our integrated congestion management and scheduling solution enhances active radio communication networks via the use of deep reinforcement learning. This is achieved by combining scheduling schemas with adaptation modelling and an optimised congestion dominance and flow management model [13].

MANETs that rely on 5G face challenges such as heavy traffic loads and strict quality of service requirements. This paper proposes a tailored solution for 5G-based MANETs, which enhances the AODV protocol. This system employs reinforcement learning to optimise routes. To ensure that the routing algorithm can identify routes that provide quality of service, nodes are tasked with maintaining a database containing state information about intermediate nodes. Based on the simulation results, the enhanced protocol demonstrates a significant degree of effectiveness in terms of throughput, end-to-end latency, and signal-to-noise ratio (SNR) [14].

Due to their dynamic structure and constrained resources, Mobile Ad Hoc Networks (MANETs) pose difficulties in terms of multicast routing and quality of service provisioning. This paper presents an agent-based solution for routing quality of service (QoS). This approach uses fuzzy logic to ascertain the optimal path by considering independent quality of service characteristics such as buffer occupancy, battery capacity, and hop count. The study examines the resilience to various attacks, including both efficacy and safety. This study presents a distributed technique for achieving optimal resource allocation and a sleep scheduling algorithm for selecting network flows in an energy-efficient manner. The resource allocation and flow selection approaches provide near-optimal outcomes with little computing effort, leading to significant performance improvements across various network topologies [15].

This paper introduces a novel model that enhances the DYMO protocol for Mobile Ad-Hoc Networks (MANETs). The model incorporates substantial improvements in the domains of route discovery and maintenance. The route discovery technique incorporates an authentication step between nodes. This process uses the MD5 hashing algorithm. In order to improve the process of route maintenance, we use reinforcement learning, which is a machine learning method. The Diffie-Hellman key management system is used to securely encrypt and decode data sent between the source and the destination. Due to the implementation of enhanced authentication and encryption measures, our evaluation of the modified protocol indicates that the performance of the MANET has been enhanced, but with a little increase in latency from start to finish. The user's text is enclosed in tags.

This study introduces a new routing protocol called Reputation Opportunistic Routing based on Q-learning (RORQ), which utilises reinforcement learning. The objective of this protocol is to tackle the intricacy of routing in mobile ad hoc networks (MANETs), especially in the presence of malicious nodes. RORQ utilises game theory to construct a reputation system that identifies and removes unauthorised nodes, ensuring efficient routing. Our simulations demonstrate that RORQ outperforms alternative protocols. It significantly enhances the reduction of packet loss, end-to-end latency, and energy consumption in the presence of blackhole and grayhole attacks [17].

Vehicular ad hoc networks (VANETs) face challenges such as control overhead and routing complexities. This paper proposes the use of Improved Deep Reinforcement Learning (IDRL) as a feasible option to address these issues. The IDRL system efficiently adjusts routing channels and reduces the time required for convergence in dynamic vehicle densities. Vehicle-to-Infrastructure (V2I) connection use the vehicle's data and transmission capacity to enable efficient packet transfer. Compared to other methods, the simulation results show that IDRL is better in reducing latency, increasing packet delivery ratios, and enhancing data dependability [18].

Robust connectivity between drones are crucial in the context of using several Unmanned Aerial Vehicles (UAVs). The current study introduces a novel routing mechanism called the Geolocation Ad Hoc Network (GLAN). This system uses geolocation data. An Adaptive GLAN (AGLAN) system is being developed to facilitate adaptation to environmental changes. This system integrates reinforcement learning. Using a pseudo-attention function may accelerate the learning process. The system demonstrates enhanced efficiency by minimising memory and processor resource use, as shown by the test conducted against conventional routing approaches [19].

The ability of wireless networks to broadcast information enables the possibility of routing in dynamic environments like MANETs and VANETs. This paper introduces DeepMPR, a multicast routing system based on multi-agent deep reinforcement learning. This technique outperforms the usual OLSR MPR selection process and does not need MPR announcement messages from neighbouring nodes. DeepMPR's analysis reveals superior efficacy in multicast forwarding as compared to other commonly used methods, leading to enhanced network reliability and reduced costs associated with broadcasting [20] .

During a crisis, prompt mobilisation and reliable communication are crucial. UAVs, or unmanned aerial aircraft, provide a means of rapid deployment that is unaffected by physical constraints. This paper presents a novel methodology for managing a WiFi ad-hoc network that operates in the air. The model employs deep Q-learning to improve Quality-of-Service (QoS), coverage, and power efficiency. The idea underwent evaluation at Istanbul Technical University's campus, where it demonstrated a packet delivery ratio of 90 percent, user coverage of 97 percent, and efficient power utilisation [21].

The mobility of nodes and resource constraints give rise to issues in Mobile Ad-Hoc Networks (MANETs). These issues might affect the stability of connections and lead to network congestion and loss of data packets. To tackle the issues in MANETs, this study proposes an Adaptive Congestion and Energy Aware Multipath Routing Scheme, abbreviated as ACEAMR. The main goal of ACEAMR is to identify routes that are stable, energy-efficient, and devoid of congestion. This is achieved by using stable link prediction to ensure reliable data transfer and an adaptive feedback mechanism to accurately detect congestion. The simulation findings clearly indicate that ACEAMR outperforms other competing approaches in terms of throughput, packet delivery ratio, latency reduction, and energy economy [22].

4. Proposed methdology

4.1 Proposed algorithm

Algorithm: RML for Energy Efficiency in MANETs

Step 1: Define Network Parameters

- **Input:** A set of mobile nodes $N=\{n1,n2,...,nk\}$ in a MANET.
- **Network Properties:** Node positions, transmission range, initial energy levels.

Step 2: Initialize Environment

- **State Space (S):** Network metrics such as node energy levels, neighbor count, and traffic load.
- **Action Space (A):** Possible actions like route changes, transmission power adjustments.
- **Reward Function (R):** A function that quantifies energy efficiency, e.g., remaining energy or successful data transmission with minimal energy use.

Step 3: Model the Reinforcement Learning Agent

- **Agent:** Each node in the MANET acts as an independent learning agent.
- **Learning Algorithm:** Choose an algorithm like Q-learning or Deep Q-Learning.
- **Initialization:** Initialize the Q-table or neural network for Deep Q-Learning with random weights.

Step 4: Define Q-Learning Parameters

- **Learning Rate (α):** Typically between 0 and 1, determining how much new information overrides old information.
- **Discount Factor (γ):** Also between 0 and 1, indicating the importance of future rewards.

Step 5: Q-Learning Algorithm

- For each episode (or time step):
- **Select Action (a):** Based on the current state (s), choose an action from the action space using a policy like ϵ-greedy.
  - **Perform Action:** Apply the chosen action in the network, leading to a new state (s').
  - **Observe Reward:** Calculate the reward based on the energy efficiency after taking action.
  - **Update Q-Table:** $Q(s, a) \leftarrow Q(s, a) + \alpha[R(s,a) + \gamma \max a' Q(s',a') - Q(s,a)]$
  - **Update State:** Set $s \leftarrow s'$.

Step 6: Energy-Efficient Routing Decision

- **Routing Algorithm:** Incorporate the learned Q-values to make routing decisions that prioritize energy efficiency.

Step 7: Repeat Learning Process

- **Iterations:** Run the algorithm for a sufficient number of episodes to ensure adequate learning.

Step 8: Network Adaptation

- **Dynamic Adjustment:** Allow the algorithm to adapt to changes in network topology and node energy levels.

Step 9: Performance Evaluation

- **Metrics:** Evaluate the algorithm based on energy efficiency, throughput, latency, and packet delivery ratio compared to traditional routing protocols.

4.2 Compare differences between Traditional Reinforcement Learning (RL) methods and the Proposed RL method

Table 1. Compare differences between Traditional Reinforcement Learning (RL) methods and the Proposed RL method

Feature	Traditional Reinforcement Learning	Proposed RL for Energy Efficiency in MANETs
Objective	Generally focuses on maximizing cumulative reward or achieving a specific goal.	Specifically aims to optimize energy efficiency while maintaining network performance in MANETs.
Environment	Could be any environment (games, robotics, simulations) where an agent learns from interactions.	Specifically a MANET environment with dynamic topology and energy constraints.
State Space (S)	Defined by the problem domain (e.g., positions in a game, sensor readings).	Includes network-specific metrics like node energy levels, neighbor count, traffic load.
Action Space (A)	Varies with the application (e.g., moving directions in games, control actions in robotics).	Involves actions like route selection, transmission power adjustments, data packet forwarding decisions.
Reward Function (R)	Designed to reinforce desired behaviors or achievements in the general application.	Tailored to promote energy-saving actions, such as choosing less energy-intensive routes or minimizing re-transmissions.
Learning Algorithm	Standard algorithms like Q-learning, SARSA, or Deep Q-Networks.	Similar algorithms but fine-tuned for energy-aware decision-making in the network context.
Performance Metrics	Depends on the specific application, like score in games or accuracy in tasks.	Energy efficiency, throughput, latency, packet delivery ratio, and network lifetime in MANETs.
Adaptation	Generally adapts to the learning task or environment.	Specifically adapts to changing network conditions, node mobility, and varying energy levels.
Implementation Complexity	Varies with the application but generally agnostic to specific types of environments.	Higher, due to the need to consider network dynamics, energy models, and node mobility in MANETs.
Application Scope	Broad and varied across many fields.	Specifically focused on improving energy efficiency in mobile ad-hoc network environments.
Result Assessment	Based on how well the agent learns to perform the task or maximize rewards.	Based on improvements in energy efficiency and sustaining network functionality under constraints.

4.3. Proposed flowchart

This flowchart provides a structured approach to implementing a specialized RL algorithm for improving energy efficiency in MANETs. Each step is crucial for adapting traditional RL methods to the specific challenges of mobile ad-hoc networks.

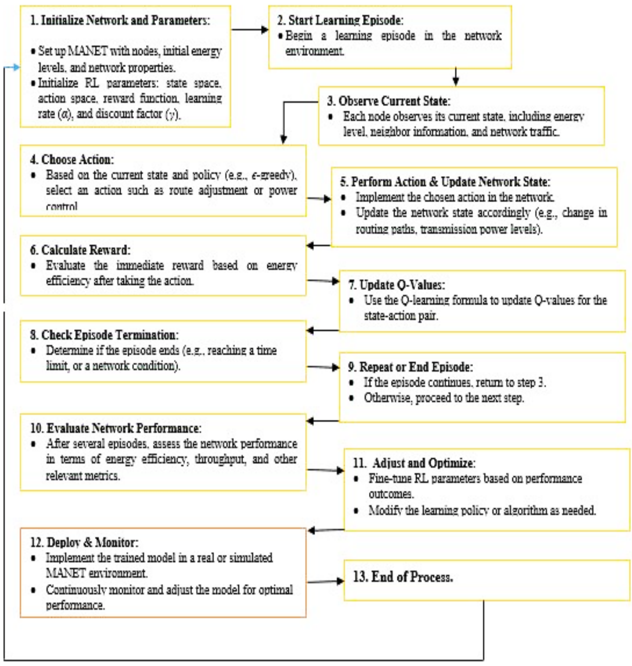

Fig 3. Proposed flowchart

4.4 Mathematical equations for proposed Reinforcement Learning (RL)

1. Q-Learning Update Equation

The core of the Q-learning algorithm is the update equation, which iteratively improves the Q-values (quality of actions) for each state-action pair. The equation is:

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

Where:

- $Q(s,a)$: Current Q-value for state s and action a.
- α: Learning rate, determining the weight given to new information.
- $R(s,a)$: Reward received after executing action a in state s.
- γ: Discount factor, representing the importance of future rewards.
- $max_{a'} \, Q(s',a')$: Maximum predicted reward for the next state s', considering all possible actions a'.

2. Reward Function for Energy Efficiency

The reward function in the context of energy efficiency in MANETs can be designed to encourage actions that save energy. It could be formulated as:

$$R(s,a) = E_{saved} - E_{spent}$$

Where:

- E_{saved}: Energy saved due to efficient actions, like choosing shorter routes or lower transmission powers.
- E_{spent}: Energy spent in executing the action, including factors like transmission energy and processing energy.

. Energy Consumption Model

To calculate E_{saved} and E_{spent}, you may need an energy consumption model, which could be represented as:

$$E_{\text{transmit}}(d) = E_{elec} + \epsilon_{amp} \times d^n$$

$$E_{\text{receive}} = E_{elec}$$

- $E_{\text{transmit}}(d)$: Energy used to transmit a packet over distance d.
- E_{receive}: Energy used to receive a packet.
- E_{elec}: Energy dissipated per bit to run the transmitter or receiver circuit.
- ϵ_{amp} : Amplification energy required per bit for a specific transmission distance.
- n: Path-loss exponent, depending on the environment.

4. Policy Equation

A policy equation, like ε-greedy, can be used to balance exploration and exploitation:

With probability ϵ choose a random action a, otherwise choose

$$A = arg \ max_{a'} Q(s,a')$$

Where:

- ϵ: Exploration rate.

These equations form the backbone of the RL methodology for enhancing energy efficiency in MANETs, incorporating aspects unique to network environments and energy constraints.

4.5 Proposed Integrate RML with AODV

This algorithm provides a detailed approach to incorporating RML into the AODV protocol for MANETs, with a focus on mathematical formulations for key components such as the state space, action space, reward function, and Q-learning update.

Step 1: Network Initialization

- **Initialize MANET:** Set up nodes, each with its own energy level and communication range.
- **Define AODV Protocol:** Implement standard AODV routing protocol.

Step 2: Define RL Framework

- **State Space (S):** The state of a node is defined by parameters such as energy level E, number of neighbors N, and traffic load T. State can be represented as $S=\{E,N,T\}$.
- **Action Space (A):** Possible actions include route selection R_s and power adjustment P_a. Actions can be represented as $A=\{R_s,P_a\}$.
- **Reward Function (R):** Define a reward function that could be a function of successful transmission T_s and energy used E_u. An example reward function: $R(S,A)=\alpha\times T_s-\beta\times E_u$ where α and β are weighting factors.

Step 3: Integrate RML with AODV

- **Route Discovery with RL:** Modify AODV's route discovery to use RL. When a node needs a route, it uses its current state S and selects an action A based on a policy derived from its Q-table or RL model.
- **Route Maintenance with RL:** Enhance the route maintenance process using RL. When a route error occurs, RL is used to decide whether to repair the route or find a new one.

Step 4: Implement Q-Learning

- **Q-Learning Algorithm:** Use the Q-learning algorithm for the RL model.
- **Q-Value Update Rule:**

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

Where α is the learning rate and γ is the discount factor.

Step 5: Policy for Action Selection

- **Policy (e.g., ε-greedy):** Define a policy for selecting actions. For ε-greedy:
- With probability ϵ, select a random action.
- Otherwise, select $A=arg\ max_{a'} Q(s,a')$

Step 6: Continuous Learning and Update

- **Iterative Learning:** Each routing decision is a learning episode. Update Q-values after each decision.
- **Adaptation:** Continuously adapt the Q-table or RL model to changing network conditions.

Step 7: Performance Evaluation

- **Network Performance Metrics:** Evaluate using metrics like . Consumed Energy ,packet delivery ratio *PDR*, and Network Overhead .
    - Consumed energy = Total energy-Remaining energy
- PDR= Total Packets Received / Total Packets Sent
- Network Overhead =Total number of network control packets transmitted during the simulation time.

Step 8: Deployment and Monitoring

- Deploy in MANET: Implement the enhanced AODV with RML in the network.
- Monitoring: Continuously monitor performance and adjust parameters as necessary.

5. Implementation

We have compared three routing protocols across six distinct scenarios, as outlined in Table 2, while shared simulation parameters are presented in Table 3.

Table 2 : Six different simulation scenarios

Sr. No	Scenario	Various values
1.	Pause Time (Sec)	30,50,60,and 70 sec.
2.	Node movements Speed(m/s)	0,1,10,15,20,and 25 m/s.
3.	Number of Nodes	40, 50, 60, 70, 80, and 90.
4.	Number of Sources	20, 30, 40, 50, and 60.
5.	Simulation Area Sizes (m.)	500m x 500m, 750m x 750m, 1000m x 1000m, and 1250m x 1250m.
6.	Sending Rate(Kbps)	48, 64, 80, 96, 112, and 128 kbps.

Table 3: Common simulation parameters

Simulator	NS 2.34
Operating System	Linux (Fedora 13)
Simulation Time	1500 sec
Moment Model	Random way point Mobility model
MAC Layer	IEEE 802.11
Traffic Type	CBR
Data Payload	1024 bytes
Energy Model	3000 Jules

- Enhanced AODV through Reinforcement Machine Learning: In this scenario, we conducted a comparative analysis between the Reinforcement Machine Learning-

Enhanced AODV protocol and the conventional AODV and DSDV protocols. This comparison aimed to evaluate the improvements in the AODV protocol when augmented with reinforcement machine learning techniques, as opposed to its standard form and the DSDV protocol.

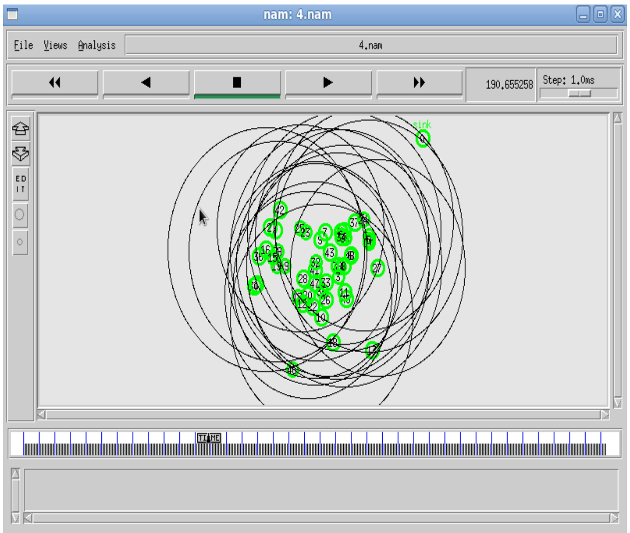

Fig 4: Route Establishment using Reinforcement Machine Learning-Based Improved AODV

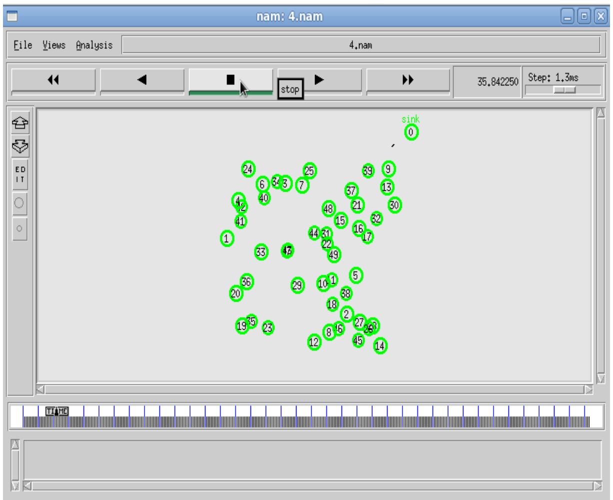

Fig 5: Data forwarding using Reinforcement Machine Learning-Based Improved AODV

Figures 4 and 5 illustrate the processes of Route Establishment and Data Forwarding, respectively, using the Reinforcement Machine Learning-Based Improved AODV (RML-AODV) protocol.

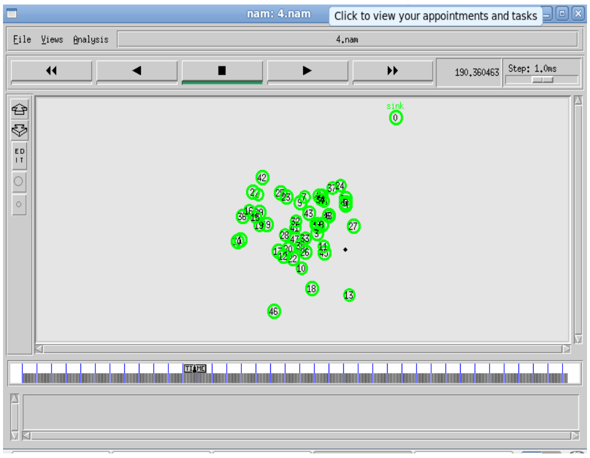

Fig 6: Q- Learning Agent takes Action of Packet dropping according State value.

Figure 6 depicts a Q-learning agent executing a packet dropping action based on the current state value. This state is characterized by network conditions, including the queue's packet count and the node's available energy. The agent has two possible actions: to drop a packet or to retain it. The objective or reward for the agent is to enhance the network's throughput while simultaneously reducing the rate of packet loss. The RML-AODV algorithm represents an advancement over the conventional Ad-hoc On-demand Distance Vector (AODV) routing protocol, employing reinforcement machine learning to comprehend the node's state for more effective routing decisions..

6. Results and analyis

6.1. Pause Time

In the previously mentioned section, we conducted a comprehensive evaluation of the three routing protocols, focusing on key performance indicators such as Energy Consumption, Packet Delivery Ratio, and Routing Overhead. The specific parameters used for this simulation are outlined below, and further details can be found in below:

- Pause Time intervals: 30, 50, 60, and 70 seconds.
- Dimensions of the simulation area: 500m x 500m.
- Total number of nodes involved: 50.
- Data transmission rate per node: 90 Kbps.

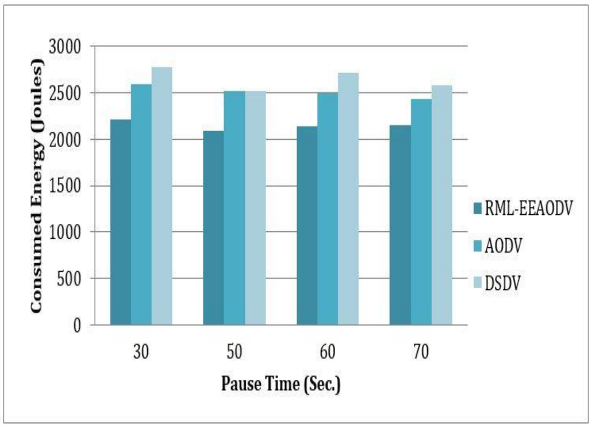

Fig 7. Consumed Energy of RML-EEAODV, AODV and DSDV with various Pause Times (Sec.).

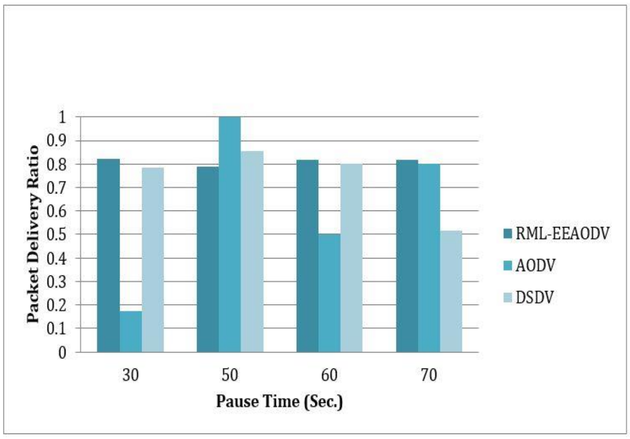

Fig 8. Packet Delivery Ratio of AODV, RML-EEAODV and DSDV with various Pause Times (Sec.).

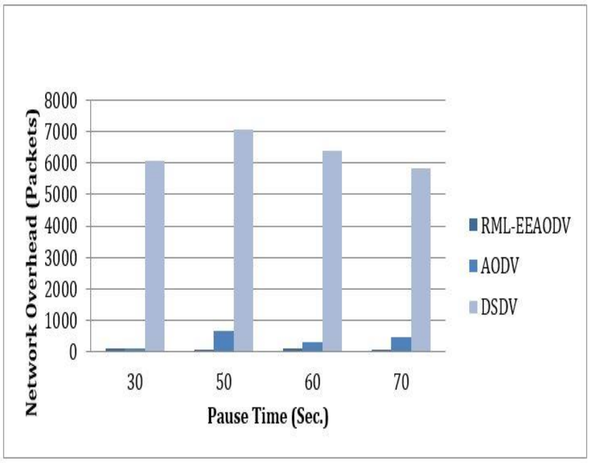

Fig 9. Network Overhead of AODV, RML-EEAODV and DSDV with different Pause Time (Sec.).

Figures 7, 8, and 9 present a comparative analysis of the AODV, DSDV, and the newly developed RML-EEAODV protocols across various network parameters, with a particular emphasis on the Pause Time in seconds.

In this evaluation, we explore how different pause times of 30, 50, 60, and 70 seconds, combined with a consistent mobility speed of 5 m/s, affect the network's performance. This combination typically results in more stable routing paths during the given time, consequently lowering the energy needed for both establishing and maintaining routes.

The interplay between node mobility and pause time is crucial, as they tend to be inversely related. From Figure 7, it's evident that the RML-EEAODV protocol is more energy-efficient than both the AODV and DSDV protocols. Moving to Figure 8, we note that the RML-EEAODV exhibits a comparatively stable Packet Delivery Ratio (PDR) across varying pause times. In contrast, the AODV shows optimal PDR at a pause time of 50 seconds, but its performance dips at pause times of 30 and 60 seconds when compared to RML-EEAODV and DSDV. Finally, Figure 9 highlights RML-EEAODV's efficient routing strategy, which minimizes unnecessary control packet transmission, thereby generating less network overhead in comparison to AODV and DSDV.

6.2. Node Mobility Speed

In this setup, we evaluate the three routing protocols across three crucial performance indicators: Energy Consumption, Packet Delivery Fraction, and Routing Overhead. The simulation for this scenario was conducted according to the parameters listed below, in addition to other specified values:

- Node Mobility Speed: Set at various speeds of 0, 1, 5, 10, 15, 20, and 25 m/s.
- Area Size: The simulation area was a square of 500m x 500m.
- Total Nodes in the Network: 50 nodes were included in the simulation.
- Data Transmission Rate per Node: Each node transmitted data at a rate of 90 Kbps.

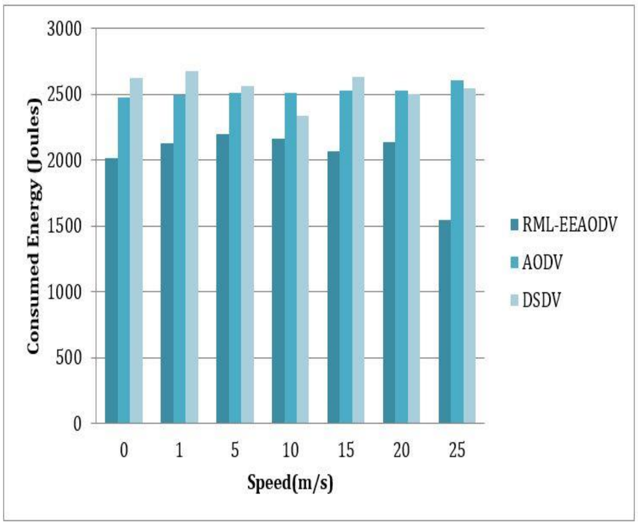

Fig 10. Consumed Energy of AODV, RML-EEAODV and DSDV with various node mobility Speeds (m/s).

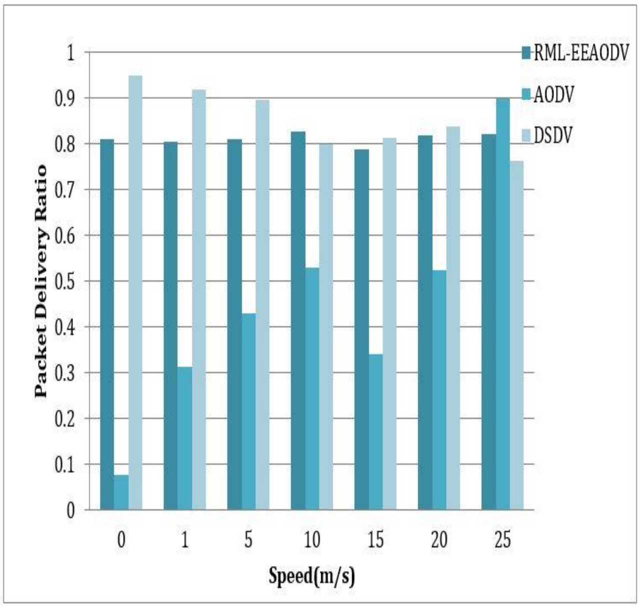

Fig 11. Packet Delivery Ratio of AODV, RML-EEAODV and DSDV with various node Mobility Speed (m/s).

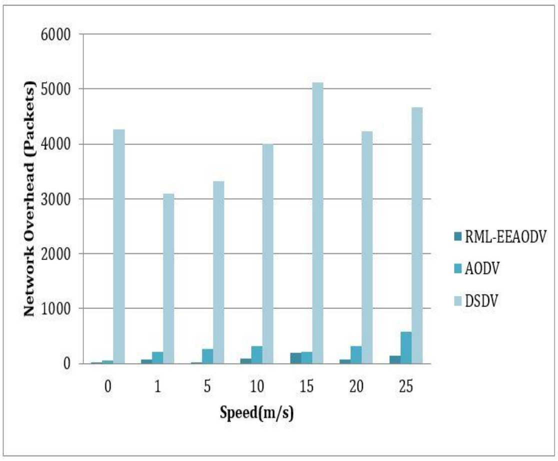

Fig 12. Network Overhead of AODV, RML-EEAODV and DSDV with various node Mobility Speed (m/s).

Figures 10, 11, and 12 showcase a comparative analysis of the AODV, DSDV, and the newly formulated RML-EEAODV protocols, examining their performance across different network parameters in relation to node mobility speeds in meters per second (m/s).

For this analysis, the performance metrics were evaluated under a range of mobility speeds, specifically 0, 1, 5, 10, 15, 20, and 25 m/s, coupled with a consistent pause time of 100 seconds. Figure 10 indicates that the RML-EEAODV protocol is more energy-efficient than both the AODV and DSDV protocols. When observing packet delivery ratios in Figure 11, there are minor yet noticeable fluctuations when comparing RML-EEAODV with the AODV and DSDV protocols. It's important to note that DSDV, as a table-driven protocol, tends to experience reduced effectiveness with higher node mobility. This is attributed to the demanding nature of frequent routing table updates necessary in fast-moving MANET environments. In contrast, RML-EEAODV excels by effectively managing the routing of control packets, thereby achieving lower network overhead than AODV and DSDV, as illustrated in Figure 12.

6.3. Numbers of nodes

In this setup, we undertake a detailed evaluation of the three routing protocols, focusing on three critical performance measures: Energy Consumption, Packet Delivery Fraction, and Routing Overhead. The simulation for this analysis was conducted in accordance with the following parameters:

- Number of Nodes: The network consisted of varying node counts, specifically 40, 50, 60, 70, 80, and 90 nodes.
- Area Size: The simulation was carried out in a 500m x 500m area.
- Data Transmission Rate per Node: Each node in the network transmitted data at a rate of 90 Kbps.

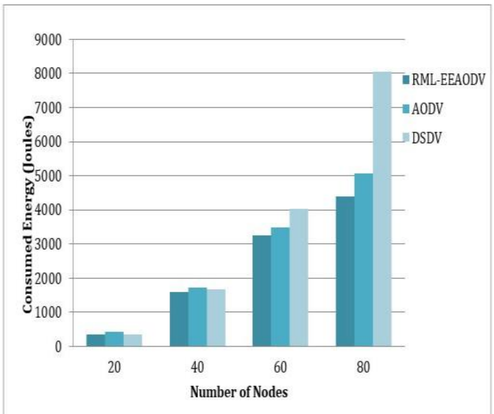

Fig 13. Consumed Energy of AODV, RML-EEAODV and DSDV with various numbers of Nodes.

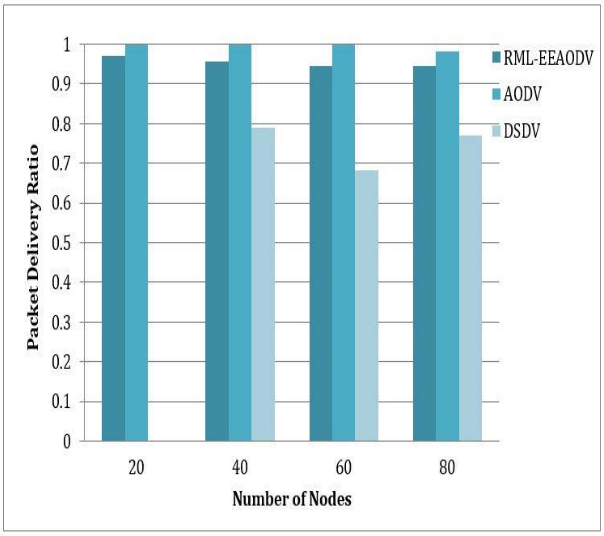

Fig 14. Packet Delivery Ratio of AODV, RML-EEAODV and DSDV with various numbers of Nodes.

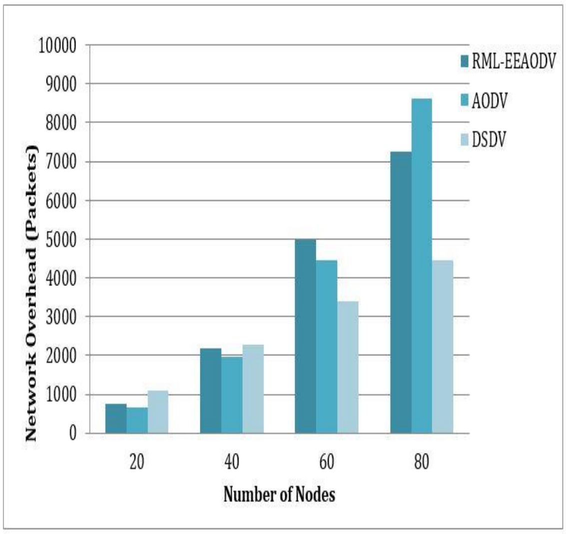

Fig 15. Network Overhead of AODV, RML-EEAODV and DSDV with various numbers of Nodes.

Figures 13, 14, and 15 present a comparison of the AODV, DSDV, and the proposed RML-EEAODV protocols across various network parameters, focusing particularly on how they perform with different numbers of nodes in the network.

In this scenario, our evaluation is centered on the network's functionality with varying node counts. Figure 13 highlights the energy consumption in the network as the number of nodes increases, accounting for both data and control packet transmissions. Here, the RML-EEAODV protocol demonstrates superior energy efficiency when compared to the AODV and DSDV protocols.

Moving to Figure 14, we observe that although the packet delivery count for RML-EEAODV is marginally lower than AODV, the difference is minimal and within acceptable limits. More importantly, RML-EEAODV surpasses DSDV in terms of packet delivery efficiency.

In terms of Routing Overhead, depicted in Figure 15, RML-EEAODV and AODV appear to perform similarly. However, upon closer examination, RML-EEAODV actually performs better due to its effective management of unnecessary control packets, thereby reducing the overall burden on the network. In contrast, DSDV tends to generate lesser overhead with an increasing number of nodes, proving to be more efficient in this aspect than both AODV and RML-EEAODV.

6.4. Numbers of sources

In this particular scenario, the three routing protocols are scrutinized based on three key performance metrics: Energy Consumption, Packet Delivery Fraction, and Routing Overhead. The simulation environment for this assessment is characterized by the following parameters, along with other specified values:

- Number of Sources: The simulation includes various source counts, specifically 20, 30, 40, 50, and 60 sources.
- Duration of Simulation: The simulation is conducted over a period of 1500 seconds.
- Size of CBR (Constant Bit Rate) Packets: Each CBR packet is sized at 1024 bytes.
- Dimensions of the Simulation Area: The area for the simulation is a square measuring 500m x 500m.

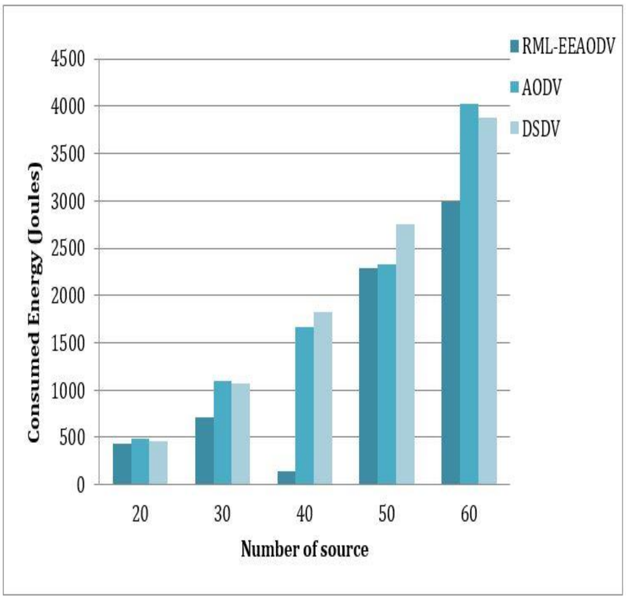

Fig 16. Consumed Energy of AODV, RML-EEAODV and DSDV with various number of Sources.

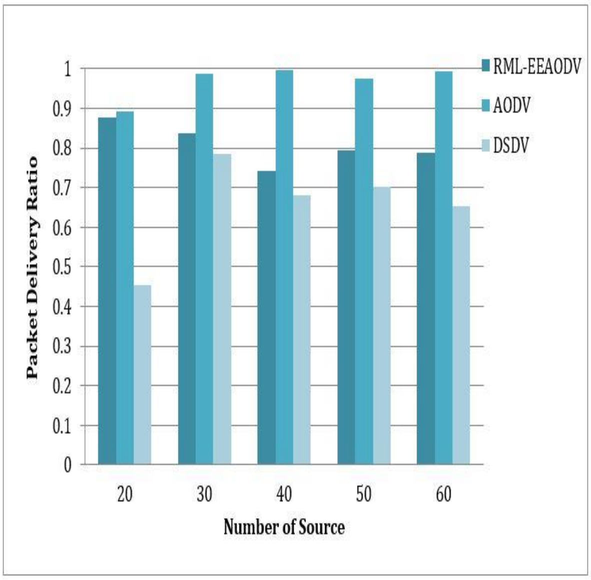

Fig 17. Packet Delivery Ratio of AODV, RML-EEAODV and DSDV with various number of Sources.

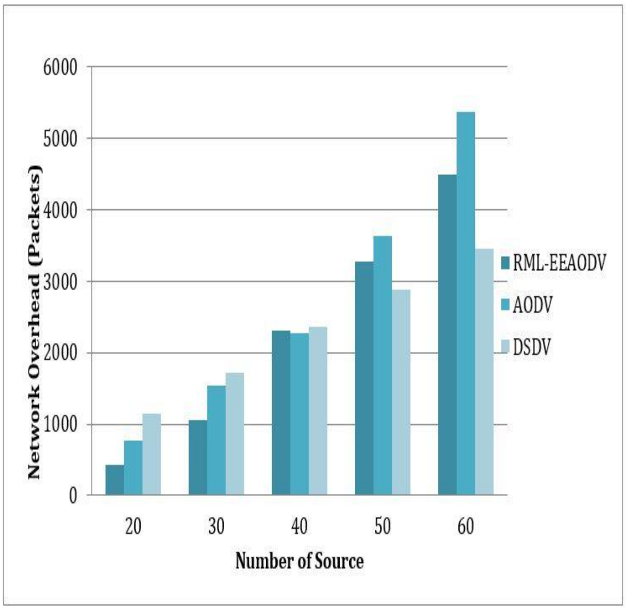

Fig 18. Network Overhead of AODV, RML-EEAODV, and DSDV with number of sources.

Figures 16, 17, and 18 provide a comparative analysis of the AODV, DSDV, and the proposed RML-EEAODV protocols, examining their performance across various network parameters with a focus on the number of sources.

Figure 16 demonstrates that the RML-EEAODV protocol is more efficient in scenarios with fewer sources, consuming less energy compared to both AODV and DSDV protocols in Mobile Ad-Hoc Networks. This efficiency is a key advantage of RML-EEAODV in energy conservation.

In Figure 17, it's observed that while RML-EEAODV exhibits a slightly lower packet delivery ratio than AODV, the performance is still within acceptable limits. Additionally, it maintains a superior packet delivery ratio when compared to DSDV. This balance between energy efficiency and packet delivery makes RML-EEAODV a viable option in diverse networking scenarios.

Figure 18 highlights the impact of increasing the number of sources on network overhead. With more sources, RML-EEAODV still manages to generate less routing overhead than AODV, indicating its effectiveness in handling increased network traffic. On the other hand, DSDV tends to incur more overhead as the number of sources rises due to the enlarged routing table and the need for more frequent update messages. This results in a higher overall network load in the DSDV protocol as compared to RML-EEAODV and AODV.

6.5 Area Sizes

In this particular scenario, we conduct an evaluation of three routing protocols, focusing on three critical performance metrics: Energy Consumption, Packet Delivery Fraction, and Routing Overhead. The simulation environment for this assessment includes various configurations, with additional details as follows:

- Flat Area Sizes: The simulations are performed over different area dimensions, specifically 500m x 500m, 750m x 750m, 1000m x 1000m, and 1250m x 1250m.
- Total Number of Nodes: Each simulation setup includes 50 nodes.
- Data Transmission Rate per Node: The nodes are configured to send data at a rate of 90 Kbps.

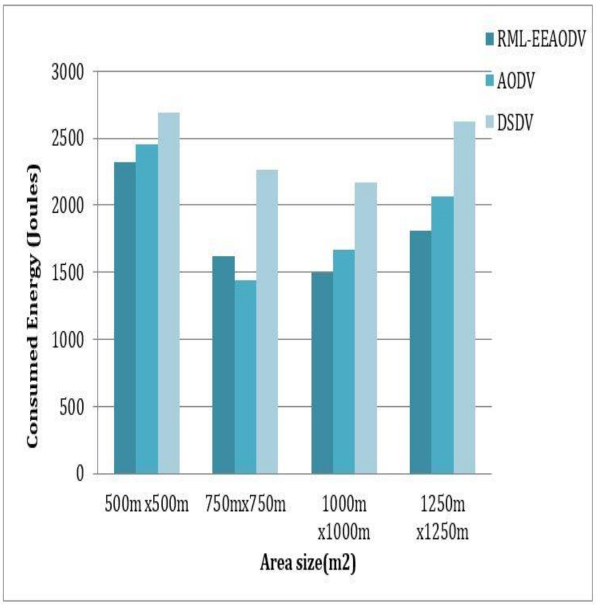

Fig 19. Consumed Energy of AODV, RML-EEAODV and DSDV with various Area sizes.

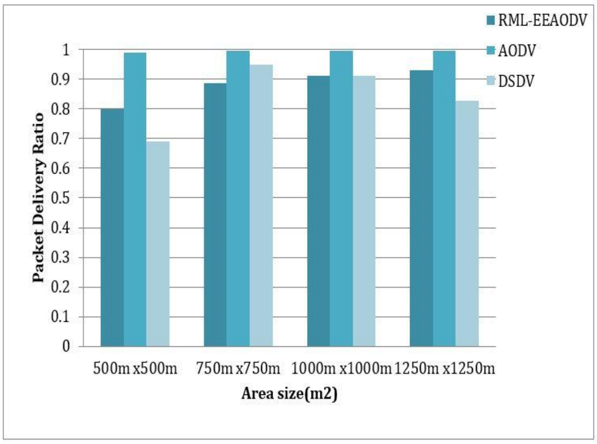

Fig 20. Packet Delivery Ratio of AODV, RML-EEAODV and DSDV with various Area size.

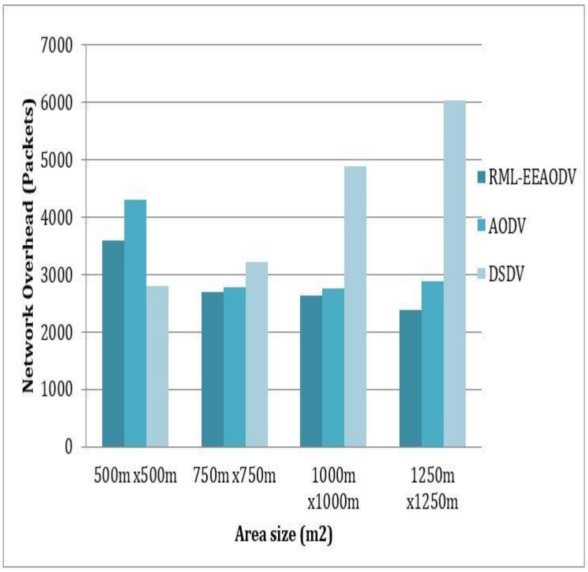

Fig 21. Network Overhead of AODV, RML-EEAODV and DSDV with various Area size.

Figures 19, 20, and 21 offer a comparative study of the AODV, DSDV, and the proposed RML-EEAODV protocols, focusing on how they perform across different network parameters in relation to varying area sizes.

Throughout the simulation, we expanded the network area and observed the performance of all three protocols. Figure 19 reveals that RML-EEAODV is more energy-efficient in larger simulation areas. This efficiency is attributed to the decreased energy consumption as the simulation area expands while maintaining a fixed number of nodes. In denser areas, energy utilization tends to increase due to the proximity of the nodes. Overall, RML-EEAODV shows a lower energy consumption compared to both AODV and DSDV protocols in Mobile Ad-Hoc Networks.

According to Figure 20, RML-EEAODV outperforms DSDV but falls short in packet delivery ratio when compared to AODV. This limitation in RML-EEAODV arises from its method of transferring packets, which is governed by the remaining energy of the nodes as a control condition. This approach restricts packet transmission, resulting in a lower packet delivery ratio.

Figure 21 demonstrates that RML-EEAODV generates less routing overhead than AODV and DSDV. In denser areas, the presence of more nodes along the path from source to destination contributes to higher overhead compared to larger areas, where nodes are more dispersed. Consequently, paths in larger areas consist of fewer nodes, necessitating fewer control packets and thereby introducing less overhead in the network.

6.6 Sending Rate

In this analysis, we evaluate three routing protocols – AODV, DSDV, and RML-EEAODV – based on three key performance metrics: Energy Consumption, Packet Delivery Fraction, and Routing Overhead. The simulation for this scenario was conducted under the following conditions, with additional parameters specified below:

- Simulation Area: The network was set up within a 500m x 500m area.
- Number of Nodes: A total of 50 nodes were included in the simulation.
- Node Sending Rates: The nodes transmitted data at various rates, specifically 48, 64, 80, 96, 112, and 128 kbps.

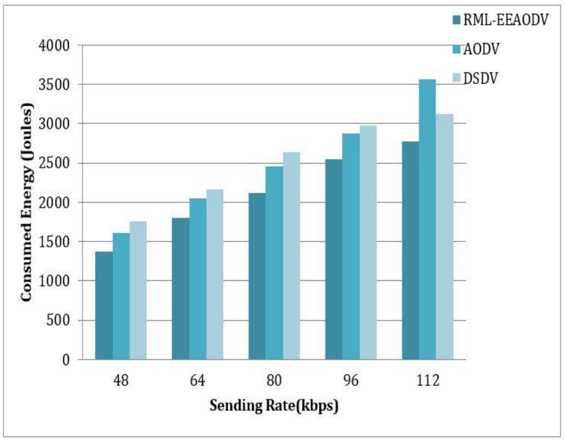

Fig 22. Consumed Energy of AODV, RML-EEAODV and DSDV with various Sending Rates (kbps).

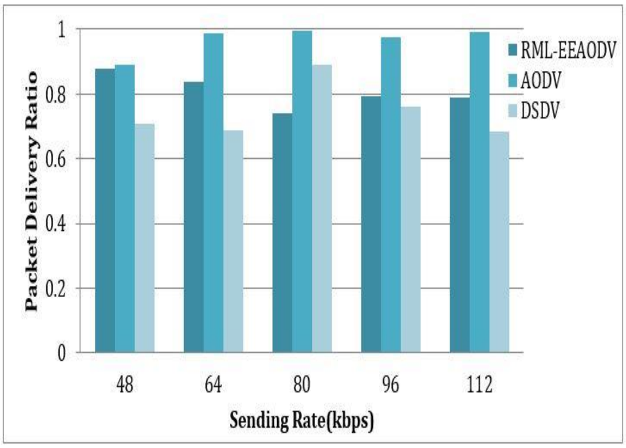

Fig 23. Packet Delivery Ratio of AODV, RML-EEAODV and DSDV with various Sending Rates (kbps).

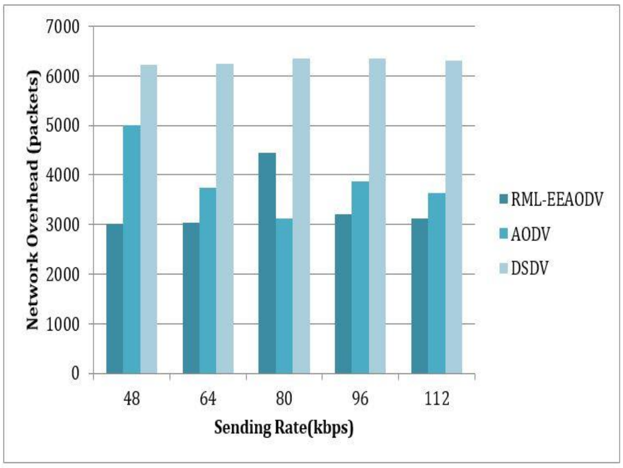

Fig 24. Network Overhead of AODV, RML-EEAODV and DSDV with various Sending Rates (kbps.).

Figures 22, 23, and 24 offer a comparative analysis of the AODV, DSDV, and the proposed RML-EEAODV protocols, assessing their performance across various network parameters against different data sending rates measured in Kbps.

Figure 22 reveals that at lower data transmission rates, RML-EEAODV is notably more energy-efficient. However, as the sending rate increases, its energy consumption also rises. This is attributed to the increased battery power usage when nodes transmit data at higher speeds. Compared to AODV and DSDV protocols, RML-EEAODV still shows lesser energy consumption in Mobile Ad-Hoc Networks, especially noticeable as the sending rate escalates with a fixed number of nodes.

In Figure 23, it is evident that RML-EEAODV outperforms DSDV but not AODV in packet delivery efficiency. This is because RML-EEAODV prioritizes the remaining energy of nodes as a deciding factor for packet forwarding, while AODV does not consider node energy, leading to continuous packet transmission until the node's energy is depleted. With limited queuing capacities in nodes, increased sending rates lead to buffer overflows and subsequent packet drops, thus reducing the packet delivery ratio.

Figure 24 shows that RML-EEAODV has a superior performance compared to AODV and DSDV in terms of routing overhead, particularly at higher sending rates. RML-EEAODV's efficiency in managing routing overhead becomes more pronounced with increased data transmission rates, distinguishing it from AODV and DSDV which tend to generate more overhead under similar conditions.

7. Conclusion and Summary

Machine learning algorithms have the potential to significantly improve the functionality of Mobile Ad-Hoc Network (MANET) routing protocols by adeptly recognizing and adapting to changes in the network environment. This section summarizes the results obtained from implementing an energy-saving mechanism in MANET routing protocols through a machine learning approach. Specifically, reinforcement learning is utilized to refine routing decisions, prompting nodes to select more efficient pathways for packet transmission. This not only enhances routing efficiency but also leads to lower energy consumption and reduced network overhead.

The Reinforcement Machine Learning-based Improved Energy Efficient AODV (RML-EEAODV) protocol underwent extensive testing and comparative analysis against traditional AODV and DSDV protocols using the NS-2.34 network simulator. The findings consistently show that the RML-EEAODV algorithm outperforms in terms of energy efficiency and minimizes network overhead in various test scenarios, despite a slight compromise in packet delivery ratio under certain conditions. As a result, RML-EEAODV stands out as a preferable choice in this study, offering substantial benefits in energy savings and enhanced efficiency in route discovery and maintenance.

The core strategy of RML-EEAODV involves minimizing unnecessary broadcasts by assessing the remaining energy and traffic load of nodes forwarding Route Request (RREQ) packets, a method informed by reinforcement learning techniques. This approach not only ensures reliable communication but also provides a secure platform for data transmission in mobile networks, thereby making it a valuable addition to MANET routing protocols.

Author contributions

Mrs. Biradar Ashwini Vishwanathrao: Conceptualization, Methodology, Software, Field study, Data curation, Writing-Original draft preparation, Software, Validation., Field study **Dr. Pradnya Ashish Vikhar:** Visualization, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Zheng Z, Jiang S, Feng R, Ge L, Gu C. Survey of Reinforcement-Learning-Based MAC Protocols for Wireless Ad Hoc Networks with a MAC Reference Model. Entropy. 2023 Jan 3;25(1):101.

[2] Godfrey D, Suh B, Lim BH, Lee KC, Kim KI. An Energy-Efficient Routing Protocol with Reinforcement Learning in Software-Defined Wireless Sensor Networks. Sensors. 2023 Oct 13;23(20):8435.

[3] Khan MF, Yau KL, Ling MH, Imran MA, Chong YW. An Intelligent Cluster-Based Routing Scheme in 5G Flying Ad Hoc Networks. Applied Sciences. 2022 Apr 6;12(7):3665.

[4] Teixeira LH, Huszák Á. Reinforcement learning environment for advanced Vehicular Ad Hoc Networks communication systems. Sensors. 2022 Jun 23;22(13):4732.

[5] Saini HK, Jain K. A Novel Deep Energy Efficient Hello Packet Scheduling for Ad Hoc Networks in Unmanned Aerial Vehicles. Engineered Science. 2023 Mar 19;22:843.

[6] Rezwan S, Choi W. A survey on applications of reinforcement learning in flying ad-hoc networks. Electronics. 2021 Feb 11;10(4):449.

[7] Rashid K, Saeed Y, Ali A, Jamil F, Alkanhel R, Muthanna A. An Adaptive Real-Time Malicious Node Detection Framework Using Machine Learning in Vehicular Ad-Hoc Networks (VANETs). Sensors. 2023 Feb 26;23(5):2594.

[8] Omoniwa B, Galkin B, Dusparic I. Communication-enabled deep reinforcement learning to optimise energy-efficiency in UAV-assisted networks. Vehicular Communications. 2023 Oct 1;43:100640.

[9] Khan MT, Adholiya A. Energy Efficient Multi Hop D2D Communication Using Deep Reinforcement Learning in 5G Networks.

[10] Li X, Wei X, Chen S, Sun L. Multi-agent deep reinforcement learning based resource management in SWIPT enabled cellular networks with H2H/M2M co-existence. Ad Hoc Networks. 2023 Oct 1;149:103256.

[11] Wang D, Liu J, Yao D, Member IE. An energy-efficient distributed adaptive cooperative routing based on reinforcement learning in wireless multimedia sensor networks. Computer Networks. 2020 Sep 4;178:107313.

[12] Vishwanathrao BA, Vikhar PA. Reinforcement Machine Learning-based Improved protocol for Mobile Ad-Hoc Networks. Scandinavian Journal of Information Systems. 2023 Oct 19;35(3):469-83.

[13] Jothi Lakshmi S, Karishma M. A Modified DSR Protocol Using Deep Reinforced Learning for MANETS. IETE Journal of Research. 2023 Jun 16:1-2.

[14] Duong TV. An improved method of AODV routing protocol using reinforcement learning for ensuring QoS in 5G-based mobile ad-hoc networks. ICT Express. 2023 Jul 13.

[15] Rengarajan G, Ramalingam N, Suriyan K. Performance enhancement of mobile ad hoc network life time using energy efficient techniques. Bulletin of Electrical Engineering and Informatics. 2023 Jun 15;12(5):2870-7.

[16] Hamamreh RA, Ayyad MR, Abutaha M. RAD: reinforcement authentication model based on DYMO protocol for MANET. International Journal of Internet Protocol Technology. 2023;16(1):46-57.

[17] Ryu J, Kim S. Reputation-Based Opportunistic Routing Protocol Using Q-Learning for MANET Attacked by Malicious Nodes. IEEE Access. 2023 Feb 6.

[18] Upadhyay P, Marriboina V, Goyal SJ, Kumar S, El-Kenawy ES, Ibrahim A, Alhussan AA, Khafaga DS. An improved deep reinforcement learning routing technique for collision-free VANET. Scientific Reports. 2023 Dec 8;13(1):21796.

[19] Park C, Lee S, Joo H, Kim H. Empowering Adaptive Geolocation-Based Routing for UAV Networks with Reinforcement Learning. Drones. 2023 Jun 9;7(6):387.

[20] Kaviani S, Ryu B, Ahmed E, Kim D, Kim J, Spiker C, Harnden B. DeepMPR: Enhancing Opportunistic Routing in Wireless Networks through Multi-Agent Deep Reinforcement Learning. arXiv preprint arXiv:2306.09637. 2023 Jun 16.

[21] Ariman M, Akkoç M, Sari TT, Erol MR, Seçinti G, Canberk B. Energy-efficient RL-based aerial network deployment testbed for disaster areas. Journal of Communications and Networks. 2023 Jan 9;25(1):25-34.

[22] Arun M, Jayanthi R. An adaptive congestion and energy aware multipath routing scheme for mobile ad-hoc networks through stable link prediction. Measurement: Sensors. 2023 Dec 1;30:100926.