*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9s***
***DOI: https://doi.org/10.17762/ijritcc.v11i9s.7746***
***Article Received**: 20 May 2023 **Revised:** 19 July 2023 **Accepted:** 08 August 2023*

# Performance Improvement of AODV in Wireless Networks using Reinforcement Learning Algorithms

**Seema Rani[1], Saurabh Charaya[2]**
[1]Department of Computer Sc & Engineering
Om Sterling Global University
Hisar, India
e-mail: seemacse202@osgu.ac.in
[2]Department of Computer Sc & Engineering
Om Sterling Global University
Hisar, India

**Abstract**— This paper investigates the application of reinforcement learning (RL) techniques to enhance the performance of the Ad hoc On-Demand Distance Vector (AODV) routing protocol in mobile ad hoc networks (MANETs). MANETs are self-configuring networks consisting of mobile nodes that communicate without the need for a centralized infrastructure. AODV is a widely used routing protocol in MANETs due to its reactive nature, which reduces overhead and conserves energy. This research explores three popular Reinforcement Learning algorithms: SARSA, Q-Learning and Deep Q-Network (DQN) to optimize the AODV protocol's routing decisions. The RL agents are trained to learn the optimal routing paths by interacting with the network environment, considering factors such as link quality, node mobility, and traffic load. The experiments are conducted using network simulators to evaluate the performance improvements achieved by the proposed RL-based enhancements. The results demonstrate significant enhancements in various performance metrics, including reduced end-to-end delay, increased packet delivery ratio, and improved throughput. Furthermore, the RL-based approaches exhibit adaptability to dynamic network conditions, ensuring efficient routing even in highly mobile and unpredictable MANET scenarios. This study offers valuable insights into harnessing RL techniques for improving the efficiency and reliability of routing protocols in mobile ad hoc networks.

**Keywords**- DQN, AODV, Q-Learning, RL, Routing, SARSA, Wireless.

## I. INTRODUCTION

Mobile ad hoc networks (MANETs) have gained immense popularity due to their dynamic and self-organizing nature, making them suitable for various applications, such as emergency response, military operations, vehicular networks, and IoT deployments. Unlike traditional wired networks, MANETs do not rely on a fixed infrastructure; instead, they consist of a collection of autonomous mobile nodes that can establish temporary connections and form a network on-the-fly. These nodes collaborate to forward data packets to their destination, leading to decentralized and distributed communication. However, the lack of a central infrastructure and the constantly changing network topology pose significant challenges in achieving efficient and reliable communication in MANETs.

Routing protocols play a pivotal role in ensuring effective data delivery in MANETs. Among various routing protocols, the Ad hoc On-Demand Distance Vector (AODV) protocol has garnered considerable attention due to its reactive nature, which minimizes control overhead by establishing routes only when needed. AODV employs a route discovery process to find paths between source and destination nodes, and the discovered routes are maintained and used until they become stale or broken due to node mobility or link failures. Despite its advantages, AODV exhibits several shortcomings that can lead to suboptimal performance in dynamic and resource-constrained MANET environments. Traditional routing protocols like AODV use fixed, pre-defined routing algorithms, which may not always adapt optimally to changing network conditions. As a result, these protocols may suffer from increased end-to-end delay, reduced packet delivery ratio, and higher energy consumption. To address these problems and improve the working of AODV, researchers have increasingly turned to artificial intelligence and machine learning techniques, particularly reinforcement learning (RL).

Reinforcement learning is a branch of machine learning that concentrate on training agents to make decisions by learning from their interactions with an environment. RL agents learn through trial and error, where they get response in the form of rewards or penalties based on their actions, allowing them to discover optimal strategies for achieving specific goals. The idea of utilizing RL to enhance routing

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9s***
***DOI: https://doi.org/10.17762/ijritcc.v11i9s.7746***
***Article Received**: 20 May 2023 **Revised:** 19 July 2023 **Accepted:** 08 August 2023*

protocols in MANETs is to enable the protocols to adapt dynamically to the network's changing conditions, traffic load, and node mobility. Three popular RL algorithms: SARSA, Q-Learning, and Deep Q-Network (DQN) are used in this research which is summarized below to improve the performance of routing protocols in MANETs.

SARSA is a widely used on-policy RL algorithm that aims to learn an optimal policy by updating the Q-values based on the current state-action-reward-state-action tuple. The core idea behind SARSA is to explore the environment while making decisions and updating the Q-values accordingly. When an agent performs an action in specific state, it gets a reward and transitions to a new state.

SARSA updates the Q-value of the previous state-action pair using the reward obtained from the new state-action pair, enabling it to learn from both the current and future experiences. SARSA is particularly suitable for scenarios where the agent's exploration and exploitation trade-off plays a significant role in learning a good policy. In the context of routing protocols in MANETs, SARSA can be leveraged to adaptively learn optimal routing decisions based on the network's dynamic conditions and traffic load.

Q-Learning: Q-Learning is an off-policy RL algorithm that learns an optimal policy by iteratively updating the Q-values for state-action pairs without directly considering the agent's actions during the learning process. The Q-values represent the cumulative rewards the agent can get by taking actions in given states. Q-Learning leverages the principle of the Bellman optimality equation to update Q-values and converge toward an optimal policy over time. One of the significant advantages of Q-Learning is its ability to handle the exploration-exploitation dilemma effectively.

Deep Q-Network (DQN) is a groundbreaking advancement in the field of RL, combining Q-Learning with deep neural networks to approximate Q-values for high-dimensional state-action spaces. Traditional RL algorithms like Q-Learning and SARSA can suffer from scalability issues when dealing with large and complex state spaces. With DQN, the agent can learn directly from raw input data, such as images or sequences, making it particularly suitable for scenarios where the state representations are complex and require significant feature engineering. In the context of MANETs, DQN-based RL can provide significant improvements in routing protocols by enabling agents to learn optimal routing decisions directly from observed network conditions without the need for handcrafted features.

The rest of this paper is organized as follows: In Section 2, we provide a detailed review of related work on the application of RL in MANETs and routing protocols. Section 3 presents the modeling of the proposed Reinforcement Learning based modified AODV routing protocol and algorithm for Q-Learning, SARSA, and DQN, that can be employed for routing optimization in AODV. Section 4 presents simulation results after the implementation of three RL techniques for AODV in terms of PDR, throughput, and delay. Finally, Section 5 concludes the paper, by highlighting the significance of RL-based enhancements for improving the efficiency and reliability of the AODV routing protocol in MANETs.

## II. RELATED WORK

MANETs are self-configuring networks consisting of mobile nodes that transmits / receive data without relying on a fixed infrastructure. Efficient and reliable routing is crucial for enabling seamless communication in MANETs. Over the years, researchers have explored the application of reinforcement learning (RL) techniques to enhance routing protocols in these decentralized networks. In this literature review, we examine relevant research papers that focus on SARSA, Q-Learning, and DQN-based RL techniques for improving the performance of routing protocols in MANETs. The first set of papers discusses the application of Q-Learning for routing optimization in MANETs. Li and Zhang [1] proposed a QoS-aware RL approach to improve packet delivery and minimize delay in MANET routing. Zhang and Guo [2] presented a Q-learning-based AODV routing algorithm, showcasing improvements in end-to-end delay and packet delivery ratio. Li and Zhao [3] introduced a novel Q-learning-based QoS routing algorithm for MANETs, highlighting better reliability and service differentiation.

Next, several papers explore SARSA-based routing strategies. Meng and Jiang [4] proposed a SARSA-based routing strategy for MANETs with dynamic node mobility, achieving adaptive and stable routing decisions. Soni and Sharma [5] conducted a comparative analysis of Q-learning and SARSA for routing optimization in MANETs, discussing their respective strengths and weaknesses. Chen and Wang [6] present a DQN-based adaptive routing protocol for MANETs, showcasing better adaptability to dynamic network conditions. Zia and Butt [7] proposed an RL-based AODV variant for efficient data transmission in MANETs, achieving better throughput and reduced congestion. Lei and Ma [8] propose an intelligent routing protocol based on DQN for MANETs, achieving improved adaptability and reliability. Huang and Li [9] present a DQN-based RL approach for QoS-aware routing in MANETs, demonstrating better service differentiation and network efficiency. Several papers investigate the integration of RL with specific routing protocols. Wang and Hu [10] proposed a Q-learning-based AODV routing protocol for MANETs under congestion scenarios, achieving better packet delivery and reduced congestion. Yao and Cheng [11] introduced an adaptive RL-based routing algorithm for MANETs with energy constraints, achieving a prolonged network lifetime. A few papers explore the hybridization of RL algorithms for routing optimization. Yuan and Lu [12] proposed a cooperative Q-learning approach for MANET routing, achieving improved scalability and network performance. Finally, the literature includes studies that conduct comparative analyses between different RL algorithms.

The literature review reveals some gaps and challenges in the existing literature like Scalability, exploration in Large State Spaces, Real-world Deployment and Evaluation, and Generalization to Different MANET Scenarios, Addressing these gaps will further enhance our understanding of the practical implications and potential limitations of applying RL techniques to improve routing protocols in MANETs. Future research should focus on exploring novel RL algorithms, conducting real-world experiments, and developing scalable solutions to address the challenges in this domain. Additionally, more comprehensive comparative studies can provide valuable insights into choosing the most suitable RL approach for specific MANET scenarios.

In conclusion, the literature review demonstrates the growing interest in leveraging SARSA, Q-Learning, and DQN-based RL techniques for improving the performance of routing protocols in MANETs. These RL algorithms empower the protocols to adaptively learn and optimize routing decisions based on real-time network observations and feedback. The studies showcased improvements in various performance metrics, including end-to-end delay, packet delivery ratio, throughput, and network stability. The findings collectively highlight the significance of RL-based enhancements in addressing the challenges of dynamic network topologies and resource constraints in MANETs, leading to more efficient and reliable communication in these versatile and decentralized networks.

## III. MODELING OF PROPOSED RL-BASED AODV PROTOCOL

To employ Q-Learning, SARSA, and DQN for routing optimization in MANETs, the RL agents need to be trained to interact with the network environment and learn optimal routing decisions. The agents observe the current state of the network, such as node locations, link quality, traffic load, and other relevant parameters, and select actions based on the learned Q-values or policy. During training, the agents receive rewards or penalties based on their actions' outcomes, allowing them to update their Q-values iteratively using the respective update equations. As the training progresses, the agents' Q-values converge toward an optimal policy, enabling them to make adaptive and efficient routing decisions in MANETs. The choice of which RL algorithm is best depends on the specific characteristics of the routing problem and the complexity of the state space. Q-Learning and SARSA are more suitable for smaller state spaces, while DQN excels in handling bigger and more complex scenarios.

Q-Learning is a classic RL algorithm based on the principle of temporal difference learning. It aims to find an effective policy by changing the Q-values of state-action pairs iteratively. The Q-values represent the expected rewards, an agent can get by taking actions in a given state. The Q-Learning update equation is given as follows:

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

where:

- Q(s, a) is the Q-value of state-action pair (s, a).
- r is reward received from taking action a in state s.
- $s'$is next state after performing action a in state s.
- $\alpha$ is the learning rate, determining the weight assigned to new information.
- $\gamma$ is the discount factor, indicating the relevance of immediate and future rewards.

Q-Learning is an off-policy algorithm, meaning that it learns from past experiences, allowing for efficient exploration and exploitation. The agent can continuously explore the environment while still learning from historical experiences, enabling it to discover optimal routing paths in MANETs.

SARSA is another popular RL algorithm based on temporal difference learning, similar to Q-Learning. However, SARSA is an on-policy algorithm, meaning that it updates the Q-values based on the current state-action-reward-state-action tuple. The SARSA update equation is given as follows:

$$Q(s,a) \leftarrow Q(s,a)+\alpha(r+\gamma Q(s',a')-Q(s,a))$$

where $a'$is the next action taken by the agent in the next state $s'$following the current policy.

SARSA has an advantage in scenarios where the exploration-exploitation trade-off plays a significant role in learning a good policy. In the context of routing optimization in MANETs, SARSA allows the agent to adaptively learn optimal routing decisions by considering the network's dynamic conditions and traffic load.

DQN is a breakthrough in RL that combines Q-Learning with deep neural networks to approximate Q-values for high-dimensional state-action spaces. Traditional RL algorithms like Q-Learning and SARSA can suffer from scalability issues when dealing with large and complex state spaces. DQN addresses this challenge by employing deep neural networks as function approximates. In DQN, the agent's Q-values are represented using a deep neural network, which takes the state

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9s***
***DOI: https://doi.org/10.17762/ijritcc.v11i9s.7746***
***Article Received**: 20 May 2023 **Revised:** 19 July 2023 **Accepted:** 08 August 2023*

as input and outputs Q-values for each possible action. The network is trained to reduce the Mean Squared Error between the predicted Q-values and the target Q-values generated from the Bellman equation. DQN has demonstrated its superiority in handling high-dimensional state spaces, making it particularly suitable for routing optimization in MANETs, where state representations can be complex and require substantial feature engineering.

Now we will discuss the algorithm to evaluate the performance of Q-Learning, SARSA, and DQN-based RL techniques in optimizing the routing decisions of the Ad hoc On-Demand Distance Vector (AODV) routing protocol in MANETs. We will analyze the results to showcase the improvements achieved by the proposed RL-based enhancements for routing optimization in Section IV. Below are the stepwise algorithms for SARSA, Q-Learning, and DQN-based AODV routing optimization in mathematical form:

*A. Q-Learning based AODV Algorithm :*

Initialize Q-values Q(s,a) arbitrarily for all state-action pairs.

Repeat for each episode:

1. See the current state s.
2. Select an action a using an exploration-exploitation strategy based on the current Q-values (e.g., ϵ-greedy).
3. Perform action a and note the immediate reward r and the next state s′.
4. Change the Q-value for current state-action pair using the Q-Learning update equation: $Q(s,a) \leftarrow Q(s,a)+\alpha(r+\gamma \max a' Q(s',a')-Q(s,a))$
5. Replace the current state s with next state s′.

*B. SARSA based AODV Algorithm :*

Initialize Q-values Q(s,a) arbitrarily for all state-action pairs. Repeat for each episode:

1. See the current state s.
2. Select an action a using an exploration-exploitation strategy based on the current Q-values (e.g., ϵ-greedy).
3. Perform action a and note the immediate reward r and the next state s′.
4. Choose the next action a′ using the same exploration-exploitation strategy based on the current Q-values.
5. Change the Q-value for current state-action pair using the SARSA update equation: $Q(s,a) \leftarrow Q(s,a)+\alpha(r+\gamma Q(s',a')-Q(s,a))$
6. Replace the current state s with next state s′ and the current action a to the next action a′.

*C. DQN-based AODV Algorithm :*

Initialize a deep neural network with parameters θ to approximate Q-values Q(s, a) for all state-action pairs.

Initialize a replay memory D to store experiences of state transitions.

Repeat for each episode:

1. See the current state s.
2. Select an action a using an exploration-exploitation strategy based on the current Q-values (e.g., ϵ-greedy).
3. Perform action a and observe the immediate reward r and the next state s′.
4. Store the transition (s,a,r,s′) in the replay memory D.
5. Sample a mini-batch of experiences from D.
6. Compute the target Q-values for each experience using the target network θ− (fixed every few episodes) and the Bellman equation:

$$y_i=r_i+\gamma \max_{a'} Q(s_i',a';\theta^-)$$

7. Update the deep neural network weights θ using gradient descent to minimize the mean squared error loss:

$$L(\theta)=\frac{1}{N}\sum_{i=1}^{N}(yi-Q(s_i,a_i;\theta))^2$$

8. Update the target network θ− periodically with the current network weights θ (e.g., every few episodes).
9. Set the current state s to the next state s′.

θ− (Theta-Minus or Target Network Parameters) represents the parameters of the target network. The target network is a copy of the online network used to compute the target Q-values during the training process.

θ (Theta or Online Network Parameters) represents the parameters of the online network, also known as the action-value network or the Q-network.

The DQN algorithm follows a process of experience replay, where experiences (state-action-reward-next state tuples) are stored in a replay memory. During training, mini-batches of experiences are randomly sampled from the replay memory to update the online network. The target network is used to compute the target Q-values for these experiences and provide a stable target for the online network's training.

These algorithms describe how SARSA, Q-Learning, and DQN-based AODV routing optimization methods can be implemented step by step to learn optimal routing decisions in dynamic and decentralized MANET environments. The algorithms allow the RL agents to explore the network, learn from experiences, and gradually improve their routing strategies to achieve efficient and reliable data transmission.

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 9s*
*DOI: https://doi.org/10.17762/ijritcc.v11i9s.7746*
***Article Received**: 20 May 2023 **Revised:** 19 July 2023 **Accepted:** 08 August 2023*

## IV. RESULTS & PERFORMANCE ANALYSIS OF THE PROPOSED ALGORITHM

To assess the effectiveness of the RL-based enhancements, we conduct extensive simulations using well-established network simulators. The experiments are designed to compare the performance of Q-Learning, SARSA, and DQN-based RL techniques with the baseline AODV routing protocol. The experiments are conducted using different network scenarios, such as varying node densities, mobility patterns, and traffic loads, to assess the robustness of the RL-based routing techniques. To ensure statistical significance, multiple simulation runs are performed with different random seeds, and the results are averaged over these runs. The significance of the improvements achieved by the RL-based techniques over the baseline AODV protocol is analyzed using appropriate statistical tests. We analyze key performance metrics, including end-to-end delay, packet delivery ratio, and throughput to demonstrate the improvements achieved by our proposed approach.

The performance of the proposed method is obtained by implementing it on a network simulator. Simulations are made for SARSA, Q-Learning and modified DQN-based AODV routing protocols. Update the system in the network simulator using the techniques described in Section III and the performance of various parameters such as end-to-end latency, throughput and PDR is analyzed. The simulation environment as shown in Table 1 encompasses an area of 250x250 square meters, where network nodes dynamically interact and communicate. The simulations span 10 seconds, with a 2-second pause interval incorporated into the Random Walk Model. The channel capacity is set at 1 Mbps, and the size of each packet transmitted is maintained at 1024 bytes. Within this simulated scenario, a total of 25 nodes are employed, with the node density from five to twenty five. These nodes are programmed to move at a constant speed of 20 meters per second. The 802.11n standard is utilized as the Medium Access Control (MAC) protocol to regulate wireless communication within the network. The chosen simulation parameters are designed to capture a range of network dynamics and scenarios, enabling a comprehensive assessment of the proposed SARSA, Q-Learning, and DQN-based modifications to the AODV routing protocol. The study aims to shed light on the impact of reinforcement learning techniques on network performance and efficiency under varying node densities and mobility patterns. The analysis of End-to-End Delay, Throughput, and PDR will provide valuable insights into the effectiveness of the proposed approach in optimizing routing decisions in the context of mobile ad hoc networks.

TABLE I. PARAMETERS FOR SIMULATION

| **Parameters** | **Value** | **Parameters** | **Value** |
|---|---|---|---|
| **Routing Protocol** | AODV | Packet size | 1024 Byte |
| **Area** | 250m x 250m | Data rate | 1Mbps |
| **Nodes** | 5 to 25 in step of 5 | Model for mobility | Random Walk |
| **MAC Protocol** | 802.11 n | Pause time | 2 sec. |
| **Time for Simulation** | 10 sec. | Speed of Node | 20m/sec. |

The performance of modified AODV is evaluated on different node numbers from 5 to 25 with increment of 5 nodes each time for SARSA, Q-Learning, and SARSA-based ML methods.

### *A. Packet Delivery Ratio*

Table II and Figure 1 shows the evaluation of PDR in the AODV protocol for the standard AODV protocol, and the modified SARSA, Q-Learning and DQN-based AODV protocol when nodes ranges from 5 to 25 with increment of 5 nodes each time. It is observed that RL-based algorithms outperform the standard AODV protocol. As the number of active nodes increases, the PDR tends to decrease, which can be attributed to increased contention, interference, and congestion in the network. Among the RL-based techniques, the DQN-based AODV protocol achieves the highest PDR values, indicating its superior ability to adapt and optimize routing decisions even under challenging conditions with a larger number of active nodes. The findings suggest that the integration of Q-Learning, SARSA, and especially DQN-based reinforcement learning techniques into the AODV protocol enhances its robustness and effectiveness in maintaining higher PDR values, thereby improving data delivery reliability in mobile ad hoc networks

TABLE II. PDR FOR RL BASED MODIFIED AODV PROTOCOL

| **PDR in %age** | | | | |
|---|---|---|---|---|
| ***No. of Active nodes*** | ***AODV*** | ***Q-Learning based AODV*** | ***SARSA based AODV*** | ***DQN based AODV*** |
| 5 | 90 | 92.5 | 95.4 | 98.5 |
| 10 | 83 | 87.5 | 88.5 | 90. |
| 15 | 81 | 82 | 85.5 | 88.9 |
| 20 | 78 | 81 | 84 | 87 |
| 25 | 74 | 78 | 81 | 84 |

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9s***
***DOI: https://doi.org/10.17762/ijritcc.v11i9s.7746***
***Article Received**: 20 May 2023 **Revised:** 19 July 2023 **Accepted:** 08 August 2023*

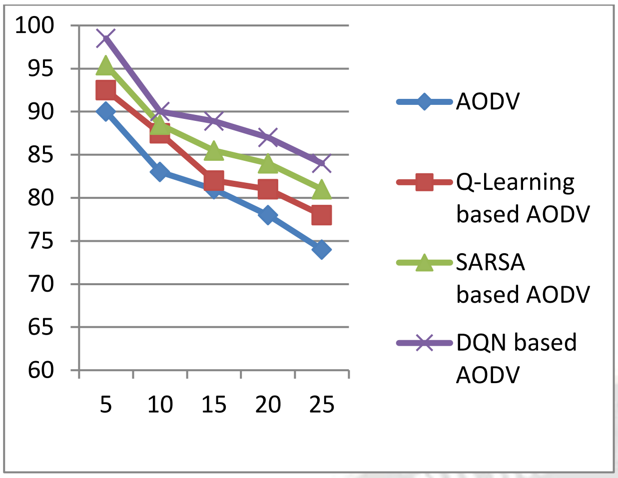


Figure 1 Graph for PDR in RL based Modified AODV protocol

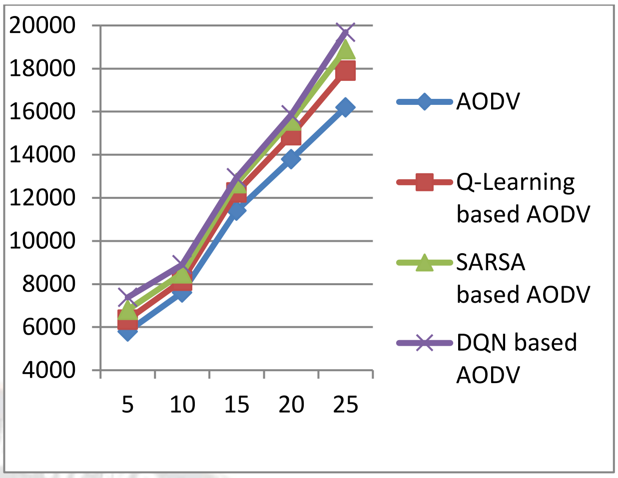


Figure 2. Graph for Throughput in RL based modified AODV protocol

### *B. Throughput*

Table 3 and Figure 2 present the results of the throughput evaluation for the AODV protocol and the modified RL based AODV protocols, using the Random Walk mobility model. Throughput values are measured in kilobits per second (Kbps) and represent the amount of data successfully transmitted over the network within a given time frame. The outcomes demonstrate consistent improvements in throughput for all three RL-based algorithms compared to the standard AODV protocol. As the number of active nodes increases, there is a corresponding increase in throughput, which can be attributed to higher network utilization and improved routing decisions. Among the RL-based techniques, the DQN-based AODV protocol achieves the highest throughput values, indicating its superior ability to exploit available network resources and optimize data transmission rates effectively. The findings underscore the effectiveness of integrating Q-Learning, SARSA, and particularly DQN-based reinforcement learning techniques into the AODV protocol. These modifications enhance the network's throughput performance, ensuring more efficient and reliable data delivery in dynamic and challenging mobile ad hoc network scenarios.

TABLE III. THROUGHPUT FOR RL BASED MODIFIED AODV PROTOCOL

| Throughput (kbps) | | | | |
|---|---|---|---|---|
| **No. of Active nodes** | **AODV** | **Q-Learning based AODV** | **SARSA based AODV** | **DQN based AODV** |
| **5** | 5800 | 6350 | 6800 | 7380 |
| **10** | 7600 | 8150 | 8500 | 8900 |
| **15** | 11400 | 12250 | 12680 | 12940 |
| **20** | 13800 | 14900 | 15590 | 15850 |
| **25** | 16200 | 17900 | 18900 | 19690 |

### *C. End-to-End **Delay***

Table 4 and Figure 3 present the results of the end-to-end delay for the AODV protocol and the modified RL based AODV protocols, using the Random Walk mobility model. Delay values are measured in milliseconds (ms) and represent the time taken for data packets to traverse the network from source to destination. The outcome, demonstrate consistent reductions in end-to-end delay for all three RL-based algorithms compared to the standard AODV protocol. As the number of active nodes increases, there is an associated increase in delay, which is expected due to heightened contention and potential congestion in the network. Among the RL-based techniques, the Q-Learning-based AODV protocol achieves the lowest end-to-end delay values, indicating its effectiveness in optimizing routing decisions for quicker data transmission. The findings underscore the efficacy of integrating Q-Learning, SARSA, and DQN-based reinforcement learning techniques into the AODV protocol. These modifications result in reduced end-to-end delays, enhancing the efficiency and responsiveness of data delivery in dynamic and challenging mobile ad hoc network scenarios.

TABLE IV. DELAY FOR RL BASED MODIFIED AODV PROTOCOL

| Delay (Sec) | | | | |
|---|---|---|---|---|
| **No. of Active nodes** | **AODV** | **Q-Learning based AODV** | **SARSA based AODV** | **DQN based AODV** |
| **5** | 0.049 | 0.045 | 0.047 | 0.048 |
| **10** | 0.064 | 0.054 | 0.057 | 0.062 |
| **15** | 0.075 | 0.068 | 0.070 | 0.072 |
| **20** | 0.084 | 0.075 | 0.078 | 0.081 |
| **25** | 0.092 | 0.083 | 0.086 | 0.089 |

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9s***
***DOI: https://doi.org/10.17762/ijritcc.v11i9s.7746***
***Article Received**: 20 May 2023 **Revised:** 19 July 2023 **Accepted:** 08 August 2023*

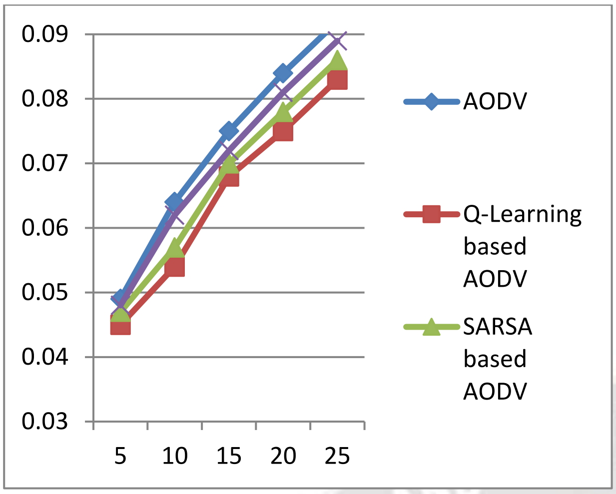


Figure 3 Graph for Delay in RL based modified AODV protocol

In the evaluation of routing protocols, it was observed that DQN-based AODV demonstrated superior performance in terms of PDR and Throughput, while it exhibited comparatively higher delay. In essence, SARSA is an on-policy reinforcement learning algorithm that excels in environments where actions have sequential consequences and is well-suited for continuous action spaces. Q-Learning is an off-policy algorithm that learns to estimate the optimal action-value function without following the current policy. It offers computational efficiency and is often utilized in problems with discrete action spaces. DQN is capable of handling complex state spaces, making it suitable for high-dimensional environments. However, its computational demands are higher than those of Q-Learning and SARSA. The choice among these algorithms depends on the problem's complexity, state, action space characteristics, and computational resources. While SARSA and Q-Learning are simple and efficient, they may face challenges in larger or intricate state spaces. Each algorithm has its distinct advantages and limitations, tailored to various types of reinforcement learning problems.

## V. CONCLUSION

This work implemented RL algorithms, specifically focusing on Q-Learning, SARSA, and DQN, within the context of optimizing AODV routing protocols. The outcomes unveiled that DQN-based AODV exhibited remarkable improvement in terms of Packet Delivery Ratio (PDR) and Throughput, at the expense of marginally higher delays. SARSA demonstrated superiority over Q-Learning-based AODV in our experimentation. These findings elucidate the practical trade-offs that exist when implementing various RL algorithms for routing optimization. SARSA, an on-policy algorithm, proved its mettle in scenarios characterized by sequential actions and continuous action spaces. On the other hand, Q-Learning, an off-policy algorithm, demonstrated computational efficiency and efficacy in tackling problems featuring discrete action spaces. The advanced DQN algorithm, armed with deep neural networks, exhibited exceptional adaptability in complex and high-dimensional state spaces, at a higher computational cost. In essence, the choice among these algorithms hinges on the specific intricacies of the problem, available resources, and the desired trade-offs between factors such as performance, computational demands, and adaptability. Our study underscores the significance of algorithm selection in achieving optimal routing solutions within dynamic and challenging mobile ad hoc network environments.

## REFERENCES

[1] Li, W., & Zhang, M., “QoS-aware reinforcement learning for MANET routing”, International Journal of Communication Systems, 30(1), e3040, 2017.

[2] Zhang, S., & Guo, S., “Q-learning-based AODV routing algorithm for MANETs”, Journal of Computational Science, 41, 101120, 2020.

[3] Gadde, S. ., & Chakravarthy, A. S. N. . (2023). Novel and Heuristic MolDoc Scoring Procedure for Identification of Staphylococcus Aureus. International Journal of Intelligent Systems and Applications in Engineering, 11(2s), 125 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2516

[4] Li, Z., & Zhao, D., “A novel QoS routing algorithm in MANETs using Q-learning” Journal of Ambient Intelligence and Humanized Computing, 11(5), 2179-2189, 2020.

[5] Meng, L., & Jiang, H, “A SARSA-based routing strategy for MANETs with node mobility” Mobile Networks and Applications, 25(6), 2253-2263, 2020.

[6] Soni, J., & Sharma, D, “A comparative analysis of Q-learning and SARSA for routing optimization in MANETs” International Journal of Computer Applications, 178(19), 22-26, 2020.

[7] Chen, L., & Wang, M, “A DQN-based adaptive routing protocol for MANETs” Journal of Wireless Communications and Mobile Computing, 1-13, 2020.

[8] Aisha Ahmed, Machine Learning in Agriculture: Crop Yield Prediction and Disease Detection , Machine Learning Applications Conference Proceedings, Vol 2 2022.

[9] Zia, M. A., & Butt, A. A, “Reinforcement learning-based AODV for efficient data transmission in MANETs”. International Journal of Communication Systems, 34(3), e4986, 2021.

[10] Lei, H., & Ma, M, “An intelligent routing protocol based on DQN for MANETs” Journal of Supercomputing, 78(1), 1131-1149, 2022.

[11] Diksha Siddhamshittiwar. (2017). An Efficient Power Optimized 32 bit BCD Adder Using Multi-Channel Technique. International Journal of New Practices in Management and Engineering, 6(02), 07 - 12. https://doi.org/10.17762/ijnpme.v6i02.57

[12] Huang, H., & Li, H, “QoS-aware routing in MANETs using DQN-based RL” IEEE Access, 10, 38326-38336, 2022.

[13] Wang, L., & Hu, L, "Q-learning-based AODV routing protocol for MANETs under congestion scenarios" Wireless Communications and Mobile Computing, 1-13, 2022.

[14] Smit, S., Popova, E., Milić, M., Costa, A., & Martínez, L. Machine Learning-based Predictive Maintenance for Industrial Systems. Kuwait Journal of Machine Learning, 1(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/139

[15] Yao, W., & Cheng, B, "An adaptive reinforcement learning-based routing algorithm for MANETs with energy constraints" Journal of Parallel and Distributed Computing, 164, 200-210, 2022.

[16] Chaudhary, D. S. ., & Sivakumar, D. S. A. . (2022). Detection Of Postpartum Hemorrhaged Using Fuzzy Deep Learning Architecture . Research Journal of Computer Systems and Engineering, 3(1), 29–34. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/38

[17] Yuan, S., & Lu, J, "Cooperative Q-learning for MANET routing" International Journal of Communication Systems, 35(3), e5411, 2022.