International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2025/33 Volume 12, Issue 4, July – August (2025)

RESEARCH ARTICLE

Reinforcement Learning-Based AI-Powered Adaptive Routing Protocols for Urban VANETs

Jyoti Sunil More
Department of Computer Engineering, Fr. C. Rodrigues Institute of Technology, Vashi, Navi Mumbai, India.
jyoti.more@fcrit.ac.in

Vaishali V. Sarbhukan (Bodade)
Department of Information Technology, Fr. C. Rodrigues Institute of Technology, Vashi, Navi Mumbai, India.
Vaishali.bodade@fcrit.ac.in

Reena Lokare
Department of Information Technology, K.J. Somaiya Institute of Technology, Mumbai, India.
reena.lokare@gmail.com

Received: 11 May 2025 / Revised: 17 July 2025 / Accepted: 30 July 2025 / Published: 30 August 2025

Abstract – Due to the static and reactive nature of the traditional routing protocols like Ad-hoc On-Demand Distance Vector (AODV) and Dynamic Source Routing (DSR), theses protocols are not very efficient and face crucial challenges such as dynamic traffic patterns, frequent topology changes and high mobility, which are highly anticipated in the Urban Vehicular Ad Hoc Networks (VANETs). To overcome these issues, this paper proposes a reinforcement learning (RL) based adaptive routing protocol which is AI powered and customized according to the needs of modern VANETs. The model incorporates Q-learning, wherein each vehicle is treated as an autonomous learning agent. This agent selects optimal path at real time based on local feedback about vehicle mobility, neighbouring spots and traffic density. The model exhibits high simulation accuracy by integrating SUMO for real time mobility simulation and NS3 for network behaviour. To optimize the pivotal performance parameters such as routing overhead, packet delivery ratio, end to end delay, throughput, hop count a specialized reward function is designed. The proposed model is evaluated for the performance with previous popular models such as AODV and DSR across different modern traffic scenarios. Undoubtedly it has given improved responsiveness, scalability and routing optimality and reliability in complex modern traffic environments.

Index Terms – Intelligent Systems, Artificial Intelligence Reinforcement Learning, Q Learning, Modern Traffic, Adaptive Routing Protocols, VANETs.

1. INTRODUCTION

Vehicular Ad Hoc Networks (VANETs) are placed at the lead of contemporary urban mobility solutions due to the development and growing demand for intelligent transportation systems. Future intelligent transportation systems are built on top of Vehicular Ad Hoc Networks (VANETs)[1][2], which allow real-time communication between vehicles and with infrastructure elements. The increasing number of connected vehicles in metropolitan areas highlights the critical need for sophisticated and flexible routing systems. VANETs, in contrast to traditional mobile networks, function in a very dynamic environment that is marked by sporadic connectivity, fluctuating traffic intensities, and regular topological changes. VANETs support a diverse spectrum of navigation, traffic management and safety applications by providing a smooth and efficient communication across roadside infrastructure and vehicles. For real time exchange of data and optimization of routes dynamically, VANETs are critical especially in densely populated urban setups. Some complex challenges such as frequently changing topologies, intermediate connectivity, high speed movement of vehicles, frequently changing traffic density affects greatly the efficiency of VANETs.

Due to the simple structure and adaptability, the conventional routing protocols such as Dynamic Source Routing and Ad hoc On-Demand Distance Vector [3] are extensively used in mobile networks. Still these protocols exhibit some limitations when they are explored in urban VANET environments. They face challenges to maintain stable routes and to respond in real time to the frequent changes in network conditions. This ultimately results in route failures, poor packet delivery performance and increased delay. Due to these challenges, there is need for more context aware routing mechanisms which also are intelligent in terms of operational infrastructure.

Although in frequent dynamics of complex transport networks like changing traffic patterns in urban VANETs, the

RESEARCH ARTICLE

performance of these algorithms declines drastically. The inability of these protocols to respond to the frequently changing network conditions ultimately result into higher delay, route failures and bad packet delivery ratios. One of the solutions to address the challenges posed by these networks is the leveraging of AI approaches in the routing schemes of VANETs. By incorporating reinforcement learning (RL) [4], the vehicles can learn the best routing practices through communication with the environment. Eventually, the routing decisions can be enhanced by continuously adapting the dynamic network situations by leveraging RL based approaches.

In dynamic environments, for improving the routing decisions, there is need for a breakthrough solution, which is based on recent developments in artificial intelligence and reinforcement learning in particular. Based on the real time feedback in the dynamic environment, based on the interactions between vehicle and the transportation infrastructure, the system can allow the network nodes to adapt the optimal routes. This ultimately makes Reinforcement learning an apt solution for VANETs, to develop adaptive routing protocols.

The proposed approach is based on Q-learning [5] which is a model free RL technique. It enables the vehicles to mimic as intelligent agents, as the approach is based on AI powered routing protocol for urban VANETs. The proposed approach identifies the most efficient path in the network in real time. For this, it considers the local observations such as vehicle speed, traffic density and neighbourhood structure. This approach is tested using SUMO and NS-3 simulations which mimics the urban traffic and network scenarios. This study addresses the limitations of conventional protocols and also presents the base for devising an adaptive, scalable and intelligent routing systems to suit to a contemporary smart city setup.

1.1. Research Objectives

The research objectives are listed below:

1. To find and resolve the challenges faced by conventional routing protocols in modern VANET environments, especially in the context of metrics such as packet delivery ratio, latency, adaptability and routing overhead in real time and high-density traffic environment.
2. To explore Q learning to enable each vehicle to mimic as an intelligent agent having ability to make real time routing decisions based on the feedback obtained from the environment.
3. To guide the RL agents for optimal routing behaviour in real time urban environment using an intelligent reward function
4. To implement framework depicting the real time simulation by integrating SUMO and NS3 for dynamic modelling and network simulation.

To assess the performance of the proposed approach with conventional algorithms using significant parameters, across various urban traffic densities such as- low, medium and high.

VANETs serve as a crucial technology for supporting real time communication among infrastructure and vehicles in urban setups. A significant challenge for VANETs is posed by the dynamic topology, frequent link interruptions due to moving vehicles, congestion and varying traffic densities. The conventional protocols like AODV and DSR which are commonly used for routing in mobile ad hoc networks doesn’t adapt the complex dynamic situations posed by VANETs. These protocols get affected due to path instability, high delay and high control overhead specifically in environments with rapidly changing vehicle movements and high traffic density. This paper depicts following significant contributions in the above context:

- Proposes a novel adaptive algorithm for routing in dynamic environment where vehicle acts as a learning agent which has ability to take intelligent decisions by learning the situations from real time environment.
- It optimizes the crucial performance parameters such as end-to-end delay, packet delivery, routing overhead, etc by devising a special reward function which supports context awareness.
- To integrate SUMO and NS3 for supporting accurate modelling of urban traffic and network dynamics for carrying out rigorous performance testing.
- To analyze different traffic situations with low, medium and high density and compare the performance with conventional models like AODV and DSR based on metrics such as delay, routing overhead and throughput.
- To lay a foundation for deploying intelligent AI driven solutions for routing traffic in real-time urban VANET infrastructures, ultimately to contribute towards development of upcoming self-enhancing vehicular networks.

The research addresses crucial challenges of conventional routing algorithms in real time environments and also sets up an AI integrated and robust foundations for next generation VANET based intelligent transportation systems.

In conclusion, the routing issues in urban VANETs are addressed by this work in a reliable and expandable manner. In addition to improving vehicular networks' performance, our protocol establishes the foundation for upcoming AI-integrated transportation systems by utilizing reinforcement learning.

RESEARCH ARTICLE

1.2. Problem Statement

In dynamic urban traffic situation, conventional VANET routing protocols like AODV and DSR are not very effective to deal with high mobility, dynamic topology and dynamic traffic density, which results in increase in delay, decreased packet delivery rate and increased routing overhead; which eventually acts as the major driving force to design a novel approach to address these issues by using an intelligent adaptive routing solution that explores learning abilities of the entities to ensure best performance in complex real time environments.

1.3. Organization of Paper

The rest of the paper is organized as follows: A comprehensive review of literature is depicted in section 2 which focuses on existing challenges in VANET routing and recent approaches based on AI. The details of proposed approach based on Q learning including the architecture, algorithm and reward function is presented in section 3. The simulation setup and the dataset descriptions along with performance testing is described in section 4. The detailed result analysis and comparison with conventional approaches are also included in this discussion. The conclusion is shown in section 5 which outlines the conclusions and the scope for future work.

2. RELATED WORK

For urban VANETs, a significant gain in network performance is observed specifically after the incorporation of reinforcement learning approach into the modern developments in AI integrated adaptive routing algorithms. The major challenges posed by urban traffic incurred due to frequently changing traffic patterns and dynamic structures often called as traffic topologies are addressed by these methods.

In order to improve packet delivery and lower latency, Jiang et al. proposed the Environment-Aware Adaptive Reinforcement Routing (EARR) protocol, where the approach of reinforcement learning (RL) is integrated to dynamically make the routing decisions based on environmental parameters like signal strength and vehicle speed [1] This ultimately reduces the network delay and enhances the packet delivery rate. In dense vehicle networks, to elevate the efficiency of data transmission, a novel approach was proposed by Kim et al. who focused on resource allocation by using RL for dynamic channel allocation [2]. Liu et al. [3] exhibited results better than AODV and DSR through simulations by using Deep Q-Networks (DQN) to maximize packet delivery in urban traffic setups. To optimize the routing in urban networks and the related operations related to Flying ad hoc networks (FANETs), a hybrid model HIROL, that combines Artificial Bee Colony algorithms, DSR, OLSR, and ANNs was introduced by Reddy and Anusha [4]. In a hybrid RL-GPS routing algorithm proposed by Wang et al. [5] it was shown that scalability and reliability was improved by integrating the feedback from environment and node placements. Similarly, a scalable and robust reinforcement learning framework was created by Suh et al., that uses learning agents instead of having static setups to enhance efficiency and scalability in dynamic networks [6].

In VANETs, to improve routing performance and security it is important to identify hostile nodes and predict the best routes. For this, Hassan et al. designed a protocol based on Artificial Neural Network (ANN) that aggregates artificial neural network with modified AODV [7].

For Delay Tolerant Networks that considers real time density and social-tie strength to identify the best message replication, Yesuf and Prathap introduced CARL-DTN which is a context-aware RL-based routing. It results in enhancing packet delivery ratios and reduces the network overhead, [8]. In case of Dragonfly networks, a multi-agent RL routing scheme based on Q adaptive learning was presented by Kang et al. [9] to carry out self-learning of the routing strategies to reduce delay and improve efficiency. This approach exhibited better packet delivery rations and better efficiency.

Vishwanathrao and Vikhar developed a reinforcement machine learning-supported energy-efficient AODV protocol, RML-EEAODV, which dynamically makes routing choices to optimize energy and simultaneously preserves the considerable packet delivery ratios [10].

Bhavadharini et al. proposed an obstacle-aware routing protocol that enhances routing efficiency by predicting the link connections and by considering delay and energy consumption using multi-objective constraints and optimized Bi-LSTM [11]. An AI-enhanced Position Assisted Routing Protocol (PARP) that integrates deep reinforcement learning (RL) to enhance transmission efficiency was proposed by Anitha et al. [12]. It improves routing decisions based on placement of nodes and real time traffic situations. It minimizes packet loss and enhances transmission performance.

Researchers are more concerned about integrating RL based adaptive routing approaches to address the challenges posed by traditional routing approaches in the context of dynamic urban traffic environments which are frequently observed in recent developments of VANETs. Vongpasith et al. [13] introduced a routing approach based on Q learning which drastically minimizes end-to-end delay by dynamically changing routes depending on traffic density and vehicle velocity.

Additionally, a multi-agent reinforcement learning framework was proposed by Sun et al. [14] in which every vehicle learns the best routes to address the challenge of route breakages. The superiority of actor-critic models in reducing routing overhead, particularly in situations with heavy vehicle traffic, was shown in another noteworthy study by Upadhyay et al. [15]. In [16], Kandali et al. combined RL and vehicular mobility prediction to predict network stability, increasing the packet delivery ratio in the face of varying traffic. Additionally, Xie et al. [17] used SUMO-NS3 integration to assess reinforcement learning on grid-based city maps and discovered notable performance improvements over GPSR.

Nan et al. [18] presented an edge-based RL system that enables computation offloading from cars to RSUs, improving learning speed and lowering latency. Federated reinforcement learning was investigated by Arya et al. [19] in order to protect data privacy and facilitate cooperative route optimization between cars.

In order to achieve more realistic urban VANET performance, Wu et al. [20] lastly proposed a context-aware RL routing technique that made use of intersection delays and traffic light patterns. Numerous recent studies have demonstrated the efficacy of AI-integrated routing strategies as VANET research continues to advance. In order to reduce the overestimation bias and increase dependability in high-mobility situations, Zhu et al. [21] proposed a prioritized routing system that uses double Q-learning.

Bilban and Inan [22] presented a novel use of proximal policy optimization (PPO) in urban routing, showing quicker convergence in route selection. In the meantime, Li et al. [23] improved route stability by utilizing vehicle trajectories and urban layouts to integrate spatiotemporal variables into a DQN-based model. Another study by Wu et al. [24] presented a scalable reinforcement learning architecture that uses decentralized learning agents to adjust to vehicle flow. In order to mitigate the impact of rogue nodes, Felix et al. [25] also incorporated a trust-aware module into their RL-routing system. Their findings revealed fewer dropped packets and increased route validity. To address changing traffic dynamics, Khan et al. [26] also suggested an RL method with variable learning rates, which produced greater flexibility than fixed-rate learners. In order to increase the resilience of vehicle data transmission in a variety of urban situations, Walid et al. used ensemble learning in conjunction with reinforcement techniques in [27].

Furthermore, a meta-reinforcement learning framework that can generalize across various city maps was implemented by Sakr et al. [28], greatly lowering the need for retraining. In real-time video streaming use cases within VANETs, Emma et al. [29] presented a delay-sensitive Q-routing model that performed better than conventional protocols. Last but not least, Rajput et al. [30] concentrated on combining reinforcement-based decision-making with cloud-assisted learning for vehicles, greatly improving routing performance in situations with heavy traffic.

Modern developments in the domain of AI have proven to be very relevant to increase the routing performance of VANETs. To enhance the scalability in dense vehicle environments, Li et al. [31] proposed a hierarchical reinforcement learning (HRL) technique, which will be able to discriminate between macro and micro-routing decisions. Ji et al. [32] proposed an approach by implementing a Dueling DQN-based model to identify the valuable routing actions in unknown topologies. To analyze the topological dynamics, Bi et al. [33] created a reinforcement learning approach based on graph neural networks (GNNs). It provided better results as compared to conventional models with respect to metrics such as route prediction accuracy. To protect data privacy and lessen reliance on a central server, Wei et al. [34] included federated reinforcement learning into VANETs in a different study. To enhance vehicle coordination, Galvavo et al. [35] integrated road condition data with multi-agent reinforcement learning (MARL).

An optimal routing model with a low convergence time was presented by Singh et al. [36] using a hybrid RL and genetic algorithm. Choi et al. [37] proposed a predictive Q routing system wherein; the current routing choices were predicted by utilizing the past trends in traffic movement. Lee et al. presented an improved approach for conventional routing in [38], by integrating a transfer learning technique, which allowed existing RL policies to be deployed to new route maps with minor retraining. Qiao et al. [39] applied Curriculum learning in reinforcement-based VANET routing where the model slowly learns from fundamental to complex traffic patterns.

In time-critical VANET applications, Antonio et al. [40] achieved better results by proposing a delay-optimized RL framework that prioritized the high priority data packets without reducing overall performance. VANETs integrated with AI driven approaches have introduced novel reinforcement learning (RL) techniques in order to improve network performance. Cui et al. [41] suggested a topology-aware reliable routing protocol that utilizes adaptive Q-learning, to manage dynamic UAV network topologies and exhibit higher packet delivery rates and decrease overhead. For Dragonfly networks, Kang et al. [42] created Q-adaptive, a multi-agent RL-based routing method that showed notable gains in system throughput and latency reduction.

By incorporating multi-agent deep reinforcement learning into pre-existing Q-learning-based protocols, Kaviani et al. [43] presented DeepCQ+, which produces scalable and reliable routing in extremely dynamic networks. Emergency vehicle trip times were significantly reduced when Su et al. [44] introduced EMVLight, a decentralized multi-agent RL framework for traffic signal control and emergency vehicle routing. In order to reduce hostile activity in VANETs and improve communication reliability, Sarker et al. [45]

RESEARCH ARTICLE

suggested an RL-based neighbour selection architecture with adaptive trust management. In order to increase cache, hit rates and decrease latency in vehicle networks, Zhang et al. [46] created PRVC, a pre-trained RL-based caching technique. By combining deep reinforcement learning with policy gradient techniques for collision avoidance in VANETs, Ganesh and Ayyasamy [47] made it possible for cars to communicate with one another and maintain safe distances. To summarize, all the prominent approaches are presented in table 1 along with algorithms or methodology used, their advantages disadvantages.

Table 1 Prominent Approaches Reviewed

Methodology	References	Algorithms / Models	Pros	Cons
RL-Based Routing	[1, 2, 5, 13–15, 18, 20, 22, 24, 47]	Q-Learning, DQN, PPO, Basic RL, RSU-RL, Mobility-Aware RL	Adaptive, low overhead, suitable for dynamic networks	Slow learning, lacks scalability in dense networks
Deep RL (DRL) & GNN Variants	[3, 15, 23, 31, 33, 39]	DRL, DRL+Graph, Hierarchical DRL, Curriculum RL	Topology-aware, supports urban/logistics use cases	Computationally intensive, needs training data
Multi-Agent RL (MARL)	[6, 9, 32, 34, 35, 43, 44]	DeepCQ+, Q-Adaptive, DDQN, EMVLight, MARL-VLC	Robust, scalable, distributed control	Coordination overhead, training complexity
Hybrid/Heuristic Learning	[4, 10, 11, 36]	HIROL, Bi-LSTM, Genetic + Firefly	Energy-efficient, obstacle-aware	High algorithm complexity, limited generalization
Trust, Federated & Secure RL	[7, 19, 25, 42]	AODV+ANN, FL-IDS, Trust-RL	Privacy-preserving, secure communication	High local processing demand
Meta-RL, Transfer, TL-RL	[8, 29, 38, 40]	CARL-DTN, Meta-RL Caching, TL-RL, AIM5LA	Faster adaptation, edge-ready, latency-aware	Needs pretrained models, memory intensive
Caching & Content-Aware Models	[16, 26, 27, 45]	Segment Routing, Ensemble Classifier, PRVC	Improves content delivery, jamming resilience	Limited routing intelligence
Game Theory / Others	[28, 30, 41, 46, 12]	Game Models, Delay QoS, Adaptive Q-RL	Encourages cooperation, QoS enhancement	Complex logic, narrow use-case focus

The literature reviews mentioned above have shown a number of important research gaps that highlight the necessity of a novel AI-driven routing algorithm for urban VANETs. Although reinforcement learning (RL) algorithms have been used to VANET routing in many recent research, most of them still rely on oversimplified network assumptions or mainly validate theoretical models without being deployed in dense, realistic urban mobility environments. Current RL-based protocols, such those that use Q-learning or Deep Q-Networks (DQN), frequently test in isolated situations devoid of network congestion, infrastructure heterogeneity, or dynamic vehicle dynamics, disregard real-time scalability, or fail to incorporate real traffic simulation data.

Furthermore, the majority of studies only assess performance on a small number of metrics, including packet delivery ratio or delay, and fail to take into account a variety of QoS factors, such as routing overhead, dependability, and flexibility. The restricted comparison with well-known classical protocols such as AODV and DSR under identical simulation conditions is another significant gap that makes it challenging to evaluate the practical advantages of AI-based techniques. To address these gaps, this paper presents a novel idea based on adaptive routing protocol leveraging Q learning approach where every vehicle is interpreted as an intelligent agent. The model selects the optimal routes based on the information about environmental parameters such as neighbour count, traffic density and mobility topology.

This approach has shown a great performance elevation validated by simulation using SUMO for mobility and NS3 for network framework realization. The paper's contribution is significant in terms of creating a real time routing model, specially curated for grid based urban VANETs, building a reward function to optimize multiple routing metrics and later validating the model with existing traditional algorithms such as DSR and AODV across variety of datasets. This work will prove to be a significant paradigm shift in the implementations of AI-integrated routing in smart city transport systems.

RESEARCH ARTICLE

3. PROPOSED MODELLING

The proposed approach integrates a Q-learning approach for adaptive routing protocol. In this, each node in the VANET is mimics as an intelligent agent based on reinforcement learning (RL). The goal of the agent is to choose best possible next hop path to make sure that the data transmission is efficient, in the context of urban traffic having dynamic behavior and varying network environment. Each node records to local settings pertaining to the traffic density, current position and the status of neighboring nodes (e.g. mobility and position). These collected recordings are mapped into state space which represents the surroundings in which the agents interact.

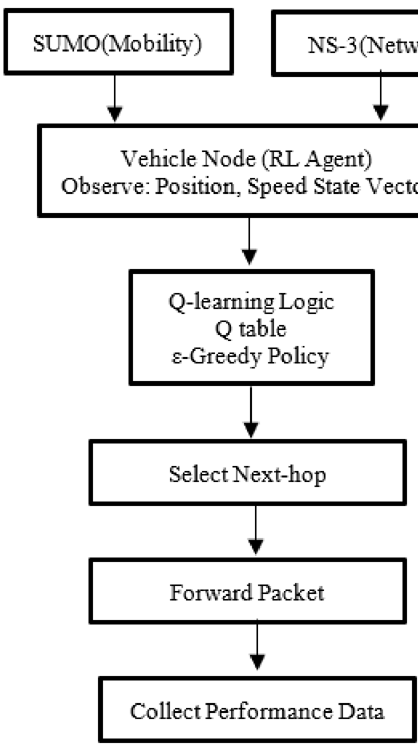

Figure 1 Data Flow Diagram for the Proposed RL-Based Model for Modeling Urban Traffic Using VANETs

The action space holds information about the feasible next hop node that the vehicle can go next. To implement the optimal routing strategy, the RL agent uses the Q learning algorithm. Depending upon the rewards received after making a routing decision, the agent modifies its Q value at every step.

The reward function is designed with a goal to reduce latency, improve reliable paths by considering parameters like successful packet delivery and end to end latency. The nodes continuously adapt to the dynamism in the traffic environment and network topologies and optimizes the route with the help of learning process.

Every node has ability to select its next hop node depending on its learned experience. It improves the scalability and permits the network to deal with dynamic vehicular environment. The overhead of centralized control and route selection at regular intervals is reduced by using Q learning by proposed protocol.

An urban traffic environment is simulated using a grid-based setup simulated in SUMO and NS3, for testing the performance of the model. The performance is compared with conventional algorithms like AODV and DSR.

Figure 1 depicts the data flow diagram for the proposed RL-based model for modeling urban traffic using VANETs. The key features provided by a Q learning based routing protocol, where every node is interpreted as RL agent are as follows:

State Representation: It includes vehicle speed, number of neighbors, traffic density and vehicle's current position.

Action Space: Set of feasible next-hop nodes.

Reward Function: Devised to ensure high reliability paths and low-latency by considering latency, packet delivery success and hop count.

Learning Strategy: The nodes learn and adapt the features over the time from the Q-values, based on feedback from successful transmissions.

The strategy presented below describes the Q-learning-based adaptive routing protocol designed for the urban VANETs:

Strategy: Q-learning-based Adaptive Routing Protocol

Input: A set of vehicles (nodes) in the VANET with their initial positions and mobility parameters, Traffic density and other relevant network conditions, Predefined action space (possible next-hop nodes), Reward function based on packet delivery success, latency, and hop count, Q-table initialized with zero values. The expected optimal routing decisions for each node in the network.

Step 1: Initialize Parameters

- Initialize the Q-table for each node:

Q (state, action) = 0 for all possible states and actions.

Define the learning rate (α), discount factor (γ), and exploration factor (ε).

- Set the maximum number of episodes (iterations) for training.
- Define state space: Each state consists of local observations (e.g., position, speed, number of neighbors, traffic density).
- Define action space: Each action corresponds to selecting a next-hop node for data forwarding.

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2025/33 Volume 12, Issue 4, July – August (2025)

- Set the reward function:

Reward for a successful packet delivery: reward = +1.

Penalty for packet loss, high latency, or unsuccessful forwarding: reward = -1.

Consider hop count and delay to refine rewards.

Step 2: Q-learning Process (for each node)

For each episode (representing a round of interactions):

- Initialization: Randomly select a starting node.

Set the initial state: state = current_observations.

- Decision-making (Exploration or Exploitation):

Exploration: With probability ε, select a random action (next-hop node).

Exploitation: With probability 1-ε, select the action that maximizes the Q-value (action = argmax Q(state, action)).

- Action Execution:

Perform the selected action (forward data to the chosen next-hop node).

Collect feedback from the environment (reward, next state).

- Update Q-table:

Update the Q-value for the chosen state-action pair.

- Transition to Next State:

Move to the next state based on the network and traffic conditions.

- Termination Condition:

If the maximum number of episodes is reached or convergence is achieved (i.e., the Q-values stabilize), stop the learning process.

Step 3: Optimal Routing Decision

- After the learning phase, each node will use the learned Q-table to make real-time routing decisions:

At each time step, the node selects the next-hop based on the highest Q-value for the current state as stated in equation (1):

$$\text{next\_hop} = \arg\max Q(\text{state}, \text{action}) \quad (1)$$

- The chosen next-hop node becomes the forwarding destination for the packet.

Step 4: Performance Evaluation

Evaluate the protocol's performance based on key metrics such as Packet Delivery Ratio (PDR), Average End-to-End Delay, and Routing Overhead.

Compare the results with traditional routing protocols (e.g., AODV, DSR).

To improve the overall network efficiency in dynamic urban traffic setup, the proposed algorithm ensures that every node learns the optimum routing strategy using Q-learning, and adapts to the local environment. The mathematical model designs the problem as RL task. The model can be analyzed into simpler components such as- state space, action space, reward function and Q learning update strategy.

3.1. State Space (S)

All the feasible conditions a node can enter into the VANET setup is collected in a set called state space. Every node in the network exhibits following metrics to build its state:

Si=(pi, vi, di, Ni), pi: Position of node i in the network (e.g., coordinates), Ni: Number of neighboring nodes in the range of node I, di: Traffic density in the area surrounding node i. vi: Velocity (speed) of node i.

Thus, the state Si for node i is a vector of these environmental parameters, and the set of all possible states is S={S1,S2,.,Sn}.

3.2. Action Space (A)

The action space represents all possible decisions that the node can make. Each action corresponds to selecting a next-hop node to forward a data packet to. Let the action space for node ii be Ai={a1,a2,,ak}, where a_j represents selecting the j-th next-hop neighbor.

The action taken at each state is the selection of one of the available next-hop nodes.

3.3. Reward Function (R)

The reward function provides feedback to the RL agent based on the quality of its routing decision. The reward function is designed to incentivize efficient routing, minimizing packet delivery delays and avoiding congested or unreliable routes. It can be defined as equation stated in equation (2):

$$R(\mathrm{Si},\mathrm{aj}) \begin{cases} +1 \;\; if\ packet\ delivery\ is\ successful, low\ latency\ and\ few\ hops, \\ -1 \;\; if\ packet\ is\ lost\ or\ delayed\ significantly, \\ -0.5 \;\; if\ route\ incurs\ high\ overhead\ or\ congestion \end{cases} \quad (2)$$

The reward is calculated based on:

- Packet delivery success.
- End-to-end delay.
- Number of hops.
- Routing stability (i.e., avoiding route breakages).

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2025/33 Volume 12, Issue 4, July – August (2025)

3.4. Q-learning Update Rule

The Q-learning algorithm learns the optimal action policy by updating the Q-values based on the feedback from the environment (reward). The Q-value represents the expected future reward of taking a certain action from a given state. The update rule is given by the equation stated by equation (3):

$$Q(Si,aj) = Q(Si,aj) + \alpha[R(Si,aj) + \gamma \max_{a'} Q(Si+1,a') - Q(Si,aj)] \quad (3)$$

Where:

- $Q(S_i, a_j)$ is the Q-value for the current state Si and action aj.
- α is the learning rate, controlling how much the newly received reward affects the Q-value.
- γ is the discount factor, which determines the importance of future rewards.
- R(Si,aj) is the reward obtained from performing action aj in state Si.
- max a′Q(Si+1,a′) is the maximum Q-value for the next state Si+1 after taking action aj, ensuring the agent focuses on the most rewarding future actions.

3.5. Policy

The policy π/pi defines the decision-making rule for selecting the next-hop node based on the current state. After the Q-values have been learned, the policy is derived by choosing the action that maximizes the Q-value for the current state as shown in equation (4):

$$\pi(Si) = \arg\max_{a'} Q(Si,aj) \quad (4)$$

Thus, the policy selects the next-hop node aj that maximizes the expected long-term reward for node i, ensuring efficient data routing.

3.6. System Dynamics

The state of the network evolves dynamically as vehicles move through the urban grid, changing the network topology and traffic conditions. The system's state transition is influenced by: Vehicle mobility: pi(t+1)=f(pi(t),vi(t)), where the position pi of vehicle i at time t depends on its speed vi, by network topology i.e. Changes in the number of neighbors N_i and link stability and by traffic density i.e. Changes in d_i based on the location and speed of surrounding vehicles.

The protocol continuously updates the Q-table as nodes experience and adapt to new network conditions, optimizing the routing decisions over time.

3.7. Performance Metrics

To evaluate the effectiveness of the proposed model, the following key performance metrics are used:

Packet Delivery Ratio (PDR): The percentage of packets successfully delivered to the destination.

PDR=Packets Delivered/Packets Sent×100

Average End-to-End Delay: The average time it takes for a packet to travel from source to destination.

Routing Overhead: The total number of control messages exchanged in the network divided by the total data delivered.

The major steps are depicted in algorithm 1.

Input: Set of nodes (vehicles) in a VANET

State space S = {position, speed, traffic density, neighbors}

Action space A = {next-hop candidates}

Learning rate α, discount factor γ, exploration rate ε

Reward function R(s, a) based on packet delivery, delay, hop count

Output:

Optimal next-hop routing decisions per node

Steps are as follows:

1. Initialize Q-table Q(s, a) to zero for all states s ∈ S and actions a ∈ A.
2. For each episode (simulation iteration):
3. For each node i in the network:
4. Observe current state s from local environment.
5. With probability ε, explore: select random action a.
6. Otherwise, exploit: select a = argmax Q(s, a).
7. Execute action: forward packet to next-hop node.
8. Observe next state s' and receive reward r = R(s, a).
9. Update Q-value using:
10. Q(s, a) ← Q(s, a) + α [r + γ * max Q(s', a') − Q(s, a)].
11. 10. Set s ← s' and repeat until packet delivery or episode ends.
12. 11. Repeat steps 3–10 for all nodes in the network.
13. 12. After learning, each node uses argmax Q(s, a) for routing in real time.
14. 13. Monitor and periodically retrain Q-values based on network dynamics.

RESEARCH ARTICLE

15. 14.Evaluate performance: compute PDR, delay, overhead, throughput.
16. 15.Compare results with conventional protocols (AODV, DSR).

Algorithm 1 Q-Learning-Based Adaptive Routing for Urban VANETs

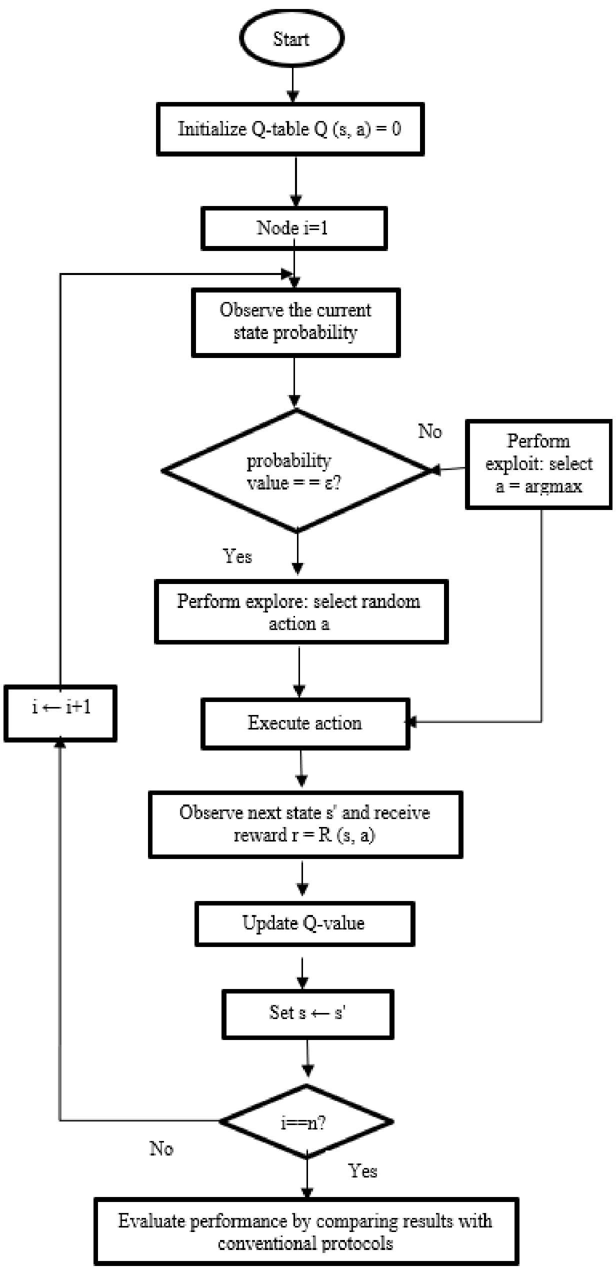

Figure 2 Flowchart for the RL Based Proposed Model

The uniqueness of this model as compared to the previous models is tabulated in table 2.

Table 2 Comparison of Proposed Model with Previous Models on Various Aspects

Aspect	Previous Models	Proposed RL-based Model
Learning Approach	Many used static or pre-defined rules (e.g., AODV, DSR), or centralized RL	Uses distributed Q-learning, where each node is an independent RL agent
Decision Scope	Central controller or static tables decide routes	Nodes learn routing policies locally via real-time experience
Adaptability	Limited adaptability; routes may not react to fast-changing urban dynamics	Continuously adapts to changing conditions (mobility, density, topology)
State Parameters	Often limited to position or neighbour info	Includes position, speed, traffic density, number of neighbours
Environment Understanding	Partial or abstract	Full context-aware modelling, enabling nuanced decisions
Routing Flexibility	Reactive; slow to adapt to new routes	Proactive and adaptive in real time
Scalability	High overhead in dense networks	Low routing overhead due to decentralized learning
Congestion Management	Poor under high vehicle density	Handles congestion-aware routing

4. RESULTS AND DISCUSSIONS

The three artificial datasets utilized in the performance comparison of AI-driven adaptive routing with AODV and DSR, which is detailed in table 3, are as follows:

The Urban_LowDensity dataset is perfect for evaluating baseline protocols because it replicates a low-traffic urban setting. The configuration has low intersection congestion and moderate-speed mobility with about 50 vehicles moving in a 5x5 city grid.

In this dataset, the data transmission is low and also the infrequent route disruptions are observed; hence the communication load is also low. The performance of protocol under low load can be observed and analysed using this dataset which proves to be helpful for evaluating the performance in the environment where traffic is predictable and controlled.

In contrast to this, a more dynamic and real-time urban environment can be mimicked by simulating the movement of about 100 cars, across 5x5 grid which helps in formation of the Urban_MediumDensity dataset. Here, in this environment, the vehicle mobility frequently changes with moderate speeds and exhibits occasional intersection congestion. As and when the communication load increases with traffic, the routing calls and data communication becomes more frequent. These setups help to assess the protocol in the situations where the routing methods respond to changing network conditions and manage mild stress. It depicts actual traffic throughout regular city hours.

Lastly, the complexity of peak urban traffic with approximately 200 vehicles in a dense grid layout is captured by the Urban_HighDensity dataset. Significant traffic, particularly at junctions, frequent vehicle stops, and extremely unpredictable mobility patterns are all included. Due to congestion and mobility, the network has a high routing overhead and a higher probability of connection failures. The reliability and scalability of routing technologies are carefully tested using this dataset. It is essential for verifying performance in the most demanding and authentic urban settings.

Table 3 Description of Datasets

Dataset Name	Description	Number of Vehicles	Traffic Density	Grid Size	Mobility Pattern	Scenario Type
Urban-LowDensity	Sparse vehicle deployment in a 5x5 city grid. Low probability of congestion.	50	Low	5x5	Uniform random movement	Light Urban Traffic
Urban-MediumDensity	Moderate vehicle density simulating average city traffic.	100	Medium	5x5	Stop-and-go traffic flow	Moderate Urban Traffic
Urban_HighDensity	High vehicle density with frequent traffic jams and route interruptions.	200	High	5x5	Dense with frequent stops	Heavy Urban Traffic

4.1. Description of Metrics

The five key performance metrics in VANET routing are: PDR, which measures packet delivery reliability; End-to-End Delay, indicating the time taken for data to reach its destination; and Routing Overhead, reflecting the control packet burden. Throughput measures the network's data handling efficiency, while Hop Count assesses path optimality based on the number of intermediate nodes. Together, these metrics evaluate a protocol's reliability, efficiency, and scalability under dynamic traffic conditions. These metrics are described in table 4.

Table 4 Description of Performance Metrics

Metric	Description	Importance
Packet Delivery Ratio (PDR)	The ratio of successfully delivered packets to the total number of packets sent.	Measures reliability and success rate of the routing protocol.
End-to-End Delay	Average time taken for a packet to travel from the source to the destination.	Indicates the latency and responsiveness of the protocol.
Routing Overhead	Total number of control messages or routing-related packets generated, relative to data packets delivered.	Reflects the efficiency and scalability under dynamic conditions.
Throughput	The amount of data successfully delivered over the network per unit time (typically measured in kbps).	Evaluates the data-carrying capacity of the protocol under traffic load.
Path Optimality / Hop Count	Average number of hops (intermediate nodes) taken to deliver a packet from source to destination.	Indicates the route efficiency; fewer hops often mean lower delay and energy.

4.1.1. Packet Delivery Ratio (PDR %)

Packet Delivery Ratio (PDR) is a crucial metric for routing reliability. It reflects the percentage of clearly delivered packets against the total number of packets sent by the source. The reliability and robustness of model is assured by a greater value of PDR for varying traffic conditions. In every dataset, the RL-Based approach performs better than both AODV and DSR as depicted in figure 3 RL can adjust to changing traffic situations since it maintains a high PDR of 96% in low-density settings, which gradually drops as the density rises. As density rises, both AODV and DSR exhibit a discernible drop in PDR, with AODV typically outperforming DSR by a small margin. In the dataset Urban_LowDensity, it exhibits a PDR of 96%, and 91% for AODV and 89% for DSR. As traffic density rises, PDR reduces for all the models. The RL based model depicts higher adaptability with PDR of 87% in the High-Density dataset and clearly outperforms AODV 76% and DSR 72%. This indicates the proficiency of RL model to dynamically cope up to network parameters such as packet loss or changes in topology.

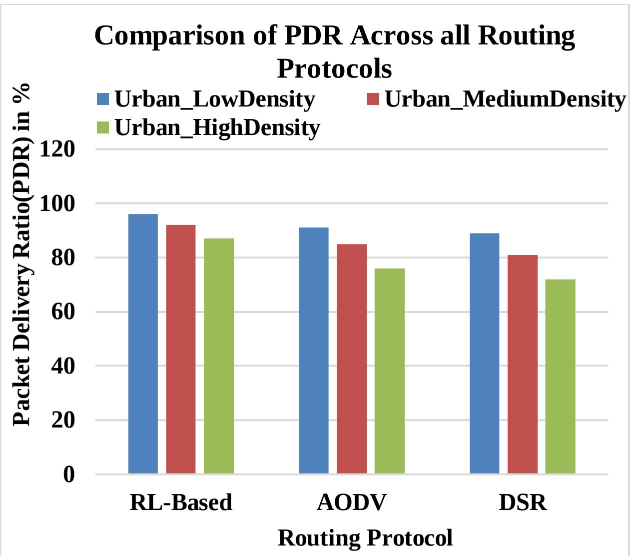

Figure 3 Performance Comparison for the Metric Packet Delivery Ratio PDR

4.1.2. End-to-End Delay (in ms)

In a network, the average time taken by the packet to get delivered form source to destination is often termed as End-to-End Delay. It includes latencies due to route finding, transmission, waiting time in queue, and transfer time.

Low latency is very crucial for real time and context aware applications such as emergency response and autonomous driving. Amongst all datasets used for comparison, the proposed RL-Based protocol provides the best performance in terms of least end-to-end delay, especially in the high-density traffic setups as depicted in figure 4. It adapts to the network more efficiently by preventing the delay due to traffic congestion. The traditional algorithms AODV and DSR exhibit more delays and thus show their inefficiency to perform in fast communication and congested settings, especially in the Urban_HighDensity scenario. In Urban_MediumDensity, the delay for RL-based is 67ms, which is very less as compared to AODV's 81ms and DSR's 89ms. For heavy traffic, the latency increases across all protocols, but RL-based remains comparatively lesser (optimum) at 90ms, while AODV and DSR exhibits 108ms and 115ms. This is because of learning based approach of RL based protocol for route selection, avoiding congested or dynamically changing paths.

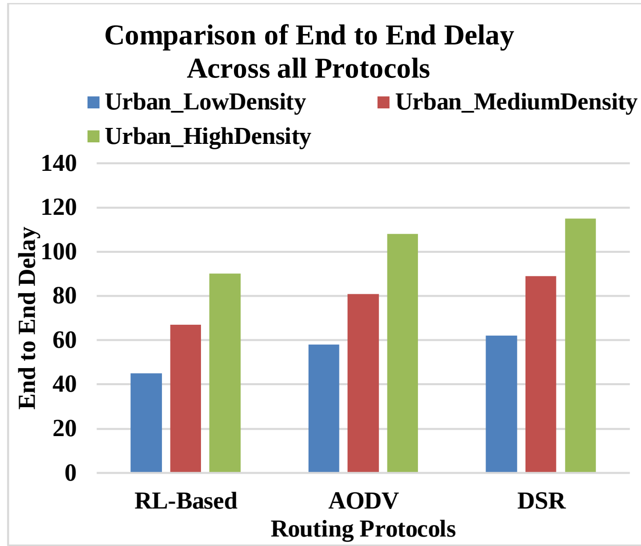

Figure 4 Performance Comparison for the Metric End to End Delay

4.1.3. Routing Overhead (Packets)

To maintain the network routing information, the number of control packets are required, this decides the Routing Overhead. The bandwidth requirement increases to cater this overhead, due to this, the available bandwidth reduces and congestion increases.

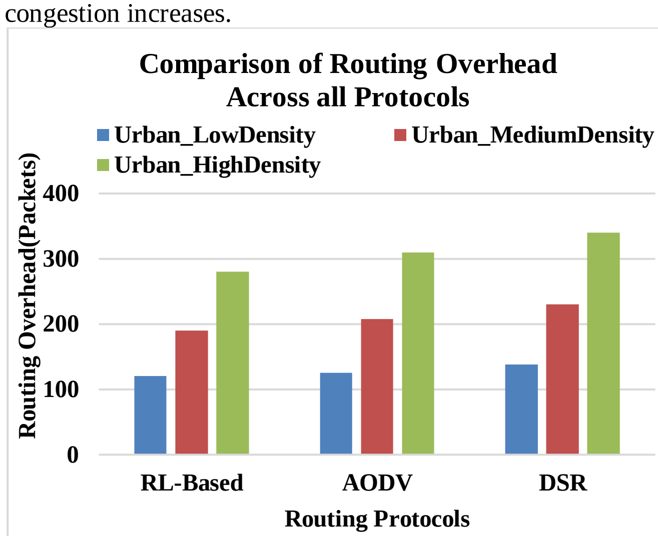

Figure 5 Performance Comparison for the Metric Routing Overhead

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2025/33 Volume 12, Issue 4, July – August (2025)

RESEARCH ARTICLE

Thus, lower overhead automatically leads to a more efficient model. RL-based protocol shows a decrease routing overhead than the AODV and DSR protocols, in all the datasets, particularly in the high-density traffic situations as shown in figure 5. This clearly implies that the RL based approach is effective in the context of the situations where fewer routing updates are available.

AODV and DSR rely highly on reactive routing algorithms, due to which they reflect higher overhead, especially in higher density areas in metropolitan grids. The RL-based model exhibits significantly lower routing overhead, especially in the Urban_HighDensity scenario (280 control packets) as compared to AODV and DSR which shows 310 and 340 control packets. This efficiency is due to the distributed learning mechanism in RL which avoids frequent route discoveries.

4.1.4. Throughput in kbps

Throughput is a measure of the total amount of successfully transferred data over the given time interval. It depicts the ability to handle data in a network.

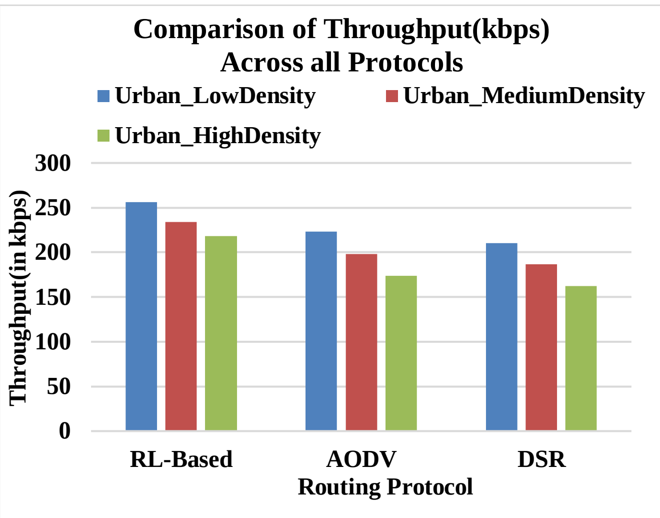

Figure 6 Performance Comparison for the Metric Throughput

For all the datasets, the RL-based approach shows a significantly high throughput, especially in Urban_LowDensity and Urban_MediumDensity setup where there is less congestion as shown in figure 6. The throughput for AODV and DSR reduces with increase in network density, whereas the RL-Based approach produces higher throughput due to its adaptive nature.

The RL-based protocol exhibits the highest throughput for all datasets. In Urban_MediumDensity, it depicts as 234 kbps, whereas it is 198 kbps (AODV) and 187 kbps (DSR). This is due to RL-based agents, who select paths with less congestion and more stable routes.

4.1.5. Average Hop Count

Hop Count defines the number of intermediate nodes a packet traverses from source to destination. Lesser hops ensure faster delivery and decrease the packet loss or path breaks.

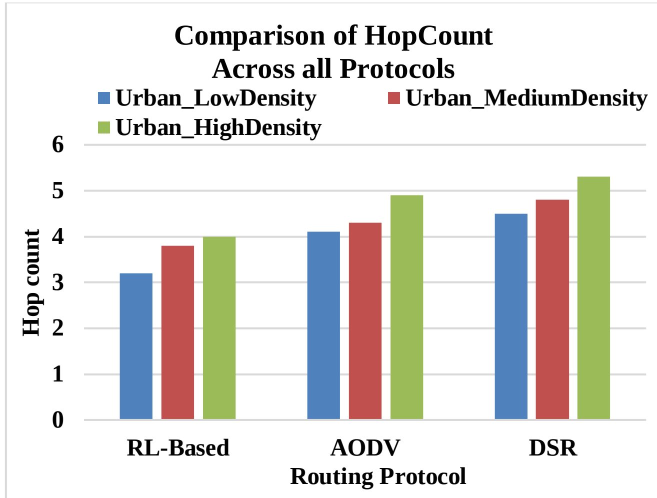

Figure 7 Performance Comparison for the Metric Average Hop Count

The RL-Based protocol maintains an average hop count of 3.8, which is less than AODV (4.3) and DSR (4.8) as shown in figure 7. Shorter path lengths imply more direct routes, minimizing delays and reducing the probability of link failures.

RL-based routing approach provides a less hop count for all datasets. For example, it averages 3.2 hops in Urban_LowDensity and 4.0 in Urban_HighDensity, while AODV and DSR range between 4.1–5.3 hops. This optimization is by virtue of learning-based route selection avoids long or unstable routes.

The result readings shown in table 5 clearly show that for all the datasets, the performance of RL based adaptive routing approach is better than traditional routing protocols like AODV and DSR. Its ability to adjust runtime to the frequently changing network conditions is a major benefit. This is particularly evident in the Urban_HighDensity setup, where the RL protocol performs far better than AODV and DSR particularly for the metrics such as Packet Delivery Ratio (PDR) and End-to-End Delay. The traditional algorithms rely on static and reactive route discovery, whereas the RL based adaptive routing method constantly learns from the environment, modifies the routing decisions based on real-time traffic and network topology.

The impact of RL based protocol for managing the network resources is obvious by its performance showed in terms of throughput and lower routing overhead. RL based approach has exhibited higher throughput, particularly in setups with

RESEARCH ARTICLE

heavy network density by reducing congestion and the need for not so effective route changes required dynamically.

AODV and DSR rely on static data and exhibit increase routing overhead. They perform worse because of their struggle to keep the routes study and effectively managing the dynamically growing traffic. They depend on frequent route finding and maintenance, so they show poor performance in highly dynamic setups.

Table 5 Performance Comparison for the Datasets for Various Protocols

Dataset	Protocol	PDR (%)	Delay (ms)	Overhead (Packets)	Throughput (kbps)	Average Hop Count
Urban_LowDensity	RL-Based	96	45	120	256	3.2
	AODV	91	58	125	223	4.1
	DSR	89	62	138	210	4.5
Urban_MediumDensity	RL-Based	92	67	190	234	3.8
	AODV	85	81	208	198	4.3
	DSR	81	89	230	187	4.8
Urban_HighDensity	RL-Based	87	90	280	218	4
	AODV	76	108	310	174	4.9
	DSR	72	115	340	162	5.3

It is obvious from the results that amongst all the algorithms, RL based adaptive routing protocol performs better as compared to the traditional counterparts AODV and DSR across all the datasets, the same is exhibited. The ability of RL based approach to dynamically adapt the shifting network situations is the main advantage of it. In Urban_HighDensity scenario, where the RL protocol performs better than AODV and DSR in terms of Packet Delivery Ratio (PDR) and End-to-End Delay, it is particularly noticed. RL based approach continuously learns from the environment and modifies the routing decisions by adapting the real time traffic situations and the changing network topologies.

On the contrary, the conventional protocols depend fully on the static data to find the optimum routes. The RL based protocol exhibits its superiority in managing resources and proves it through its elevated performance in throughput and lowering the routing overhead.

In the environments with heavy network traffic, the RL based approach reduces congestion and decreases the need for effective less route modifications thus it improves the throughput. The traditional protocols perform poorly in overall setups because they struggle to keep the routes stable and effectively to manage the increasing traffic and congestion situations.

Thus, the RL based approach is very effective in congested and dynamic urban VANETs, due to its ability to learn from the environment and make proactive choices. It has capacity to completely transform routing approach in modern vehicle infrastructures by reducing delay, increasing efficiency, lowering routing overhead and maintain a higher packet delivery rate.

The real world urban VANETs get benefited significantly by the proposed Q-learning-based adaptive routing algorithm. Its capacity to learn from the surroundings and make distributed, context aware decisions to optimize the route discovery makes it highly effective in dynamic high density city traffic. The algorithm assures reliable and low delay communication for real time service needs such as ambulance or fire trucks, by making a real time optimum route discovery. In smart city setups, this model can improve adaptive traffic management by allowing vehicles to avoid the highly congested routes and thus enhancing the overall traffic routes.

The distributed nature of the approach reduces the dependency on the infrastructure and makes it effective for areas which needs limited RSU coverage. The algorithm adapts by scaling effectively with node density and by maintaining the performance even under heavy traffic.

RESEARCH ARTICLE

For vehicular IoT and electrical vehicles, it is vital to have lower routing overhead and reduced communication paths, which leads to decreased bandwidth usage and energy consumption. By leveraging the performance benefits of providing higher throughput and lower delay, it also helps n deploying real-time multimedia and safety-critical applications, such as collision avoidance and video transmission.

To summarize, The RL-based model achieves better results because it combines the best of the aspect like Learning from experience (Q-learning), Real-time local sensing, Efficient route selection, Low overhead communication, High adaptability to changing traffic and topology. These factors lead to improved PDR, lower delays, higher throughput, lower hop counts, and minimal control overhead, especially in dense and dynamic urban environments.

Furthermore, the approach acts as a base model for future AI-integrated vehicle networks, including autonomous driving and 5G-based V2X systems. It can be integrated with deep reinforcement learning, edge computing, or federated learning for further deployment across intelligent transportation systems.

5. CONCLUSION

The research focused on improving real-time route planning for Urban Vehicular Ad-hoc Networks (VANETs). It developed an AI-powered, adaptive routing protocol using reinforcement learning. This new method consistently outperformed traditional routing approaches, like AODV and DSR. This was especially true in high-density traffic situations where conventional systems struggle to maintain reliable communication. When tested across diverse city traffic scenarios, the protocol showed significant improvements in key performance indicators such as Packet Delivery Ratio (PDR), End-to-End Delay, Routing Overhead, and Throughput. A core advantage of the Reinforcement Learning (RL) model was its inherent capacity to learn and automatically adapt to fluctuations in the network environment. This led to the establishment of more robust routes, a reduction in traffic bottlenecks, and an overall increase in network throughput. Moreover, the efficiency of the RL method in navigating intricate and dynamic vehicular networks was highlighted by its ability to decrease routing overhead and provide faster responses. The findings indicate that RL-driven protocols are well-suited to provide a dependable and adaptable solution for next-generation Urban VANET applications, especially in metropolitan settings characterized by variable traffic conditions. To sum up, this study shows how AI methods like reinforcement learning may be included into car networks to improve their functionality and open the door to the creation of smarter, self-optimizing transportation systems. For additional validation and improvement, future studies should concentrate on integrating deep reinforcement learning techniques and evaluating the suggested approach in extensive city simulations.

REFERENCES

[1] Jiang Y, Zhu J, Yang K. Environment-Aware Adaptive Reinforcement Learning-Based Routing for Vehicular Ad Hoc Networks. Sensors. 2024;24(1):40. MDPI https://doi.org/10.3390/s24010040

[2] Kim J-W, Kim J-W, Lee J. Intelligent Resource Allocation Scheme Using Reinforcement Learning for Efficient Data Transmission in VANET. Sensors. 2024; 24(9):2753. MDPI https://doi.org/10.3390/s24092753

[3] Boya Liu, Guoai Xu, Guosheng Xu, Chenyu Wang, Peiliang Zuo Deep Reinforcement Learning-Based Intelligent Security Forwarding Strategy for VANET, Sensors 2023, 23(3), 1204; MDPI https://doi.org/10.3390/s23031204

[4] Ch. Naveen Kumar Reddy, M. Anusha, "Hybrid Intelligent Routing with Optimized Learning (HIROL) for Adaptive Routing Topology management in FANETs," SSRG International Journal of Electrical and Electronics Engineering, vol. 11, pp. 30-43, 2024. Crossref, https://doi.org/10.14445/23488379/IJEEE-V11I11P104

[5] Yiwei Wang, Pengbo Si, Miao Zhang, Meng Li, Ruizhe Yang A Q-Learning Based Routing Algorithm for Software Defined Vehicular Network. 7th International Conference on Computer and Communications, 2022 IEEE.

[6] Suh, B.; Akobir, I.; Kim, J.; Park, Y.; Kim, K.-I. A Resilient Routing Protocol to Reduce Update Cost by Unsupervised Learning and Deep Reinforcement Learning in Mobile Ad Hoc Networks. Electronics 2025, 14, 166. https://doi.org/10.3390/electronics14010166

[7] Hassan MU, Al-Awady AA, Ali A, Sifatullah, Akram M, Iqbal MM, Jahangir K.,Yahya A. ANN-Based Intelligent Secure Routing Protocol in Vehicular Ad Hoc Networks (VANETs) Using Enhanced AODV. Sensors. 2024;24(3):818. MDPI https://doi.org/10.3390/s24030818

[8] Yesuf FY, Prathap M. CARL-DTN: Context Adaptive Reinforcement Learning based Routing Algorithm in Delay Tolerant Network. arXiv preprint arXiv:2105.00544v1.2 May 2021. https://doi.org/10.48550/arXiv.2105.00544

[9] Kang Y, Wang X, Lan Z. Q-adaptive: A Multi-Agent Reinforcement Learning Based Routing on Dragonfly Network. arXiv preprint arXiv:2403.16301v2[cs.NI], 3 April 2024. https://doi.org/10.48550/arXiv.2403.16301

[10] Biradar A V., Vikhar P A. Reinforcement Machine Learning-based Improved Protocol for Energy Efficiency on Mobile Ad-Hoc Networks. International Journal of Intelligent Systems and Applications in Engineering, 2024, 12(8s), 654–670.

[11] Bhavadharini R. M., Mercy Rajaselvi Beaulah P., Om Kumar C.U., Krithiga R. An Obstacle Aware Efficient MANET Routing with Optimized Bi-LSTM and Multi-Objective Constraints on Improved Heuristic Algorithm. Volume 15, Issue 4, DOI https://doi.org/10.3233/AIS-220369, SAGE Journals

[12] Anitha V, Umamaheswari TS, Saravanan B, Abhang S, Ramagundam S. AI-Enhanced Routing Protocols for Efficient Data Transmission in Wireless Networks. ICTACT Journal on Communication Technology, An International Publication of ICT Academy. September 2024; 15(3):3300-3306, DOI: 10.21917/ijct.2024.0490

[13] Phouthone Vongpasith, Soukpaseuth Banchong, Vilaysack Keosouttha, Bouaketh Vannachit, Lathsamy Chidtavong, "Q-Learning Based Routing Algorithm for Urban Vehicular Networks", International Journal of Scientific Development and Research (IJSDR), Volume 8 Issue 6, June 2023

[14] Sun Yanglong, Yiming Lin, Yuliang Tang. A Reinforcement Learning-Based Routing Protocol in VANETs January 2019 Lecture Notes in Electrical Engineering In book: Communications, Signal Processing, and Systems (pp.2493-2500) DOI:10.1007/978-981-10-6571-2_303

[15] Pratima Upadhyay, Venkatadri Marriboina, Samta Jain Goyal, Sunil Kumar, El Sayed M. El Kenawy, Abdel hameed Ibrahim, Amel Ali Al

hussan & Doaa Sami Khafaga. An improved deep reinforcement learning routing technique for collision free VANET. Scientifc Reports | (2023) 13:21796 www.nature.com/scientificreports

[16] K. Kandali, L. Bennis, O. E. Bannay and H. Bennis, "An Intelligent Machine Learning Based Routing Scheme for VANET," in IEEE Access, vol. 10, pp. 74318-74333, 2022, doi: 10.1109/ACCESS.2022.3190964.

[17] Hongbin Xie, Ge Song, Zhuoran Shi, Jingyuan Zhang, Zhenjia Lin, Qing Yu, Hongdi Fu, Xuan Song, Haoran Zhang. Reinforcement learning for vehicle-to-grid: A review. Advances in Applied Energy. Volume 17, 2025, 100214. https://doi.org/10.1016/j.adapen.2025.100214

[18] Zhaojun Nan, Yunjian Jia, Zhengchuan Chen et al. Delay-Aware Content Delivery with Deep Reinforcement Learning in Internet of Vehicles June 2021 IEEE Transactions on Intelligent Transportation Systems PP (99):1-12. DOI:10.1109/TITS.2021.3087833

[19] Arya, M.; Sastry, H.; Dewangan, B.K.; Rahmani, M.K.I.; Bhatia, S.; Muzaffar, A.W.; Bivi, M.A. Intruder Detection in VANET Data Streams Using Federated Learning for Smart City Environments. Electronics 2023, 12, 894. https://doi.org/10.3390/electronics12040894

[20] Jinqiao Wu, Min Fang, Haikun Li, And Xiao Li. RSU-Assisted Traffic-Aware Routing Based on Reinforcement Learning for Urban Vanets. Volume 8, 2020, IEEE Access.

[21] Zhengwei Zhu, Can Hu Chenyang Zhu, Yanping Zhu and Yu Sheng. An Improved Dueling Deep Double-Q Network Based on Prioritized Experience Replay for Path Planning of Unmanned Surface Vehicles J. Mar. Sci. Eng. 2021, 9(11), 1267; MDPI https://doi.org/10.3390/jmse9111267

[22] Mehmet Bilban and Onur İnan, Optimizing Autonomous Vehicle Performance Using Improved Proximal Policy Optimization. Sensors 2025, 25(6), 1941; MDPI, https://doi.org/10.3390/s25061941

[23] Yuanyuan Li, Qingfeng Guan, Junfeng Gu & Xintong Jiang. A deep reinforcement learning with dynamic spatio-temporal graph model for solving urban logistics delivery planning problems. International Journal of Digital Earth Volume 17, 2024 - Issue 1 Article:2376273 https://doi.org/10.1080/17538947.2024.2376273.

[24] Jinqiao Wu, Min Fang, Xiao li. Reinforcement Learning Based Mobility Adaptive Routing for Vehicular Ad-Hoc Networks. August 2018 Wireless Personal Communications 101(2) DOI: 10.1007/s11277-018-5809-z.

[25] Felix O. Olowononi, Danda B. Rawat & Chunmei Liu. Trust-Based Adversarial Resiliency in Vehicular Cyber Physical Systems Using Reinforcement Learning. Security in Computing and Communications Springer Nature Conference paper First Online: 10 February 2021 pp 139–151.

[26] Saifullah Khan a, Muhammad Alam b, Martin Fränzle a, Nils Müllner c, Yuanfang Chen. A Traffic Aware Segment-based Routing protocol for VANETs in urban scenarios. Computers & Electrical Engineering Volume 68, May 2018, Pages 447-462 https://doi.org/10.1016/j.compeleceng.2018.04.017.

[27] Walid El-Shafai, Ahmad Taher Azar, and Saim Ahmed. AI-Driven Ensemble Classifier for Jamming Attack Detection in VANETs to Enhance Security in Smart Cities. IEEE Access, Volume 13, 2025 /Digital Object Identifier 10.1109/ACCESS.2025.3552544.

[28] Sakr, H., Elsabrouty, M. Meta-reinforcement learning for edge caching in vehicular networks. J Ambient Intell Human Comput 14, 4607–4619 (2023). https://doi.org/10.1007/s12652-023-04583-z.

[29] Emna Bouzid,Sonia Gaied Fantar, Habib Youssef. Delay and Quality of Link Aware Routing Protocol Enhancing Video Streaming in Urban VANET, January 2020, Advances in Intelligent Systems and Computing DOI: 10.1007/978-3-030-16660-1_54.

[30] Nitin Singh Rajput, Rahul Banerjee, Dheeraj Sanghi, Chitra Kalyansundaram. Evolution of cooperation in vehicular cloud assisted networks for ITS services: A hunt game-based approach, in Future Generation Computer Systems, Volume 146, September 2023, Pages 62-77, https://doi.org/10.1016/j.future.2023.04.013

[31] F. Li, X. Song, H. Chen, X. Li and Y. Wang, “Hierarchical Routing for Vehicular Ad Hoc Networks via Reinforcement Learning,” in IEEE Transactions on Vehicular Technology, vol. 68, no. 2, pp. 1852-1865, Feb. 2019, doi: 10.1109/TVT.2018.2887282.

[32] Yuxin Ji; Yu Wang; Haitao Zhao; Guan Gui; Haris Gacanin; Hikmet Sari. Multi-Agent Reinforcement Learning Resources Allocation Method Using Dueling Double Deep Q-Network in Vehicular Networks. Volume: 72 Issue: 10, Page(s): 13447 – 13460 12 May 2023, IEEE Transactions on Vehicular Technology DOI: 10.1109/TVT.2023.3275546

[33] Xiang Bi, Lingjie Huang, Benhong Zhang, Zhen Chen and Zengwei Lyu. A dynamic routing algorithm for VANET with graph neural networks and deep reinforcement learning, International Journal of Sensor Networks. Vol. 48, No. 2, June 17, 2025 pp 119-134. https://doi.org/10.1504/IJSNET.2025.146775

[34] Dawei Wei; Junying Zhang; Mohammad Shojafar; Saru Kumari; Ning Xi; Jianfeng Ma. Privacy-Aware Multiagent Deep Reinforcement Learning for Task Offloading in VANET.IEEE Transactions on Intelligent Transportation Systems Volume: 24, Issue: 11, November 2023, Page(s): 13108 – 13122, DOI: 10.1109/TITS.2022.3202196

[35] Gonçalo Galvão; Manuela Vieira; Paula Louro; Manuel Augusto Vieira; Mário Véstias; Pedro Vieira. Multi Agent Reinforcement Learning System for Vehicular and Pedestrian Traffic Control with Visible Light Communication,8th International Young Engineers Forum on Electrical and Computer Engineering (YEF-ECE), Caparica /Lisbon, Portugal, 05 July 2024. DOI: 10.1109/YEF-ECE62614.2024.10625165.

[36] Gagan Deep Singh; Manish Prateek; Sunil Kumar; Madhushi Verma; Dilbag Singh; Heung-No Lee. Hybrid Genetic Firefly Algorithm-Based Routing Protocol for VANETs, IEEE Access (Volume: 10), Page(s): 9142 – 9151, 13 January 2022.DOI: 10.1109/ACCESS.2022.3142811.

[37] Samuel P.M. Choi, Dit-Yan Yeun. Predictive Q-Routing: A Memory-based Reinforcement Learning Approach to Adaptive Traffic Control. NIPS’95: Proceedings of the 9th International Conference on Neural Information Processing Systems 1995, Pages 945 – 951.

[38] Donghoun Lee. Transfer Learning-Based Deep Reinforcement Learning Approach for Robust Route Guidance in Mixed Traffic Environment, IEEE Access (Volume: 12), 30 April 2024, Page(s): 61667 – 61680.

[39] Zhiqian Qiao; Katharina Muelling; John M. Dolan; Praveen Palanisamy; Priyantha Mudalige. Automatically Generated Curriculum based Reinforcement Learning for Autonomous Vehicles in Urban Environment.2018 IEEE Intelligent Vehicles Symposium (IV),26-30 June 2018. DOI: 10.1109/IVS.2018.8500603

[40] Guillen-Perez Antonio, Cano Maria-Dolores. AIM5LA: A Latency-Aware Deep Reinforcement Learning-Based Autonomous Intersection Management System for 5G Communication Networks, Sensors 2022, 22(6). https://doi.org/10.3390/s22062217

[41] Yanpeng Cui, Qixun Zhang, Zhiyong Feng, Zhiqing Wei, Ce Shi, Heng Yang. Topology-Aware Resilient Routing Protocol for FANETs: An Adaptive Q-Learning Approach. IEEE Internet of Things Journal, vol. 9, no. 19, pp. 18632-18649, 1 Oct.1, 2022, https://doi.org/10.1109/JIOT.2022.3162849

[42] Y. Kang, X. Wang and Z. Lan, “Study of Workload Interference with Intelligent Routing on Dragonfly,” SC22: International Conference for High Performance Computing, Networking, Storage and Analysis, Dallas, TX, USA, 2022, pp. 1-14, doi: 10.1109/SC41404.2022.00025.

[43] Saeed Kaviani and Bo Ryu and Ejaz Ahmed and Kevin A. Larson and Anh Le and Alex Yahja and Jae H. Kim, Robust and Scalable Routing with Multi-Agent Deep Reinforcement Learning for MANETs Networking and Internet Architecture 2021, (cs.NI) https://arxiv.org/abs/2101.03273.

[44] Su H, Zhong YD, Chow JYJ, Dey B, Jin L. EMVLight: a Multi-agent Reinforcement Learning Framework for an Emergency Vehicle Decentralized Routing and Traffic Signal Control System. arXiv preprint arXiv:2206.13441. 2022 Jun 27. https://doi.org/10.1016/j.trc.2022.1039.

RESEARCH ARTICLE

[45] Sarker O, Shen H, Babar MA. Reinforcement Learning Based Neighbor Selection for VANET with Adaptive Trust Management. 22nd IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) 2023. https://doi.org/10.48550/arXiv.2310.00635.

[46] Zhang Z, Guan L, Li Y, Johnson SG, Meng Q. PRVC: A Novel Vehicular Ad-Hoc Network Caching Based on Pre-trained Reinforcement Learning International Conference on Advanced Information Networking and Applications Advanced Information Networking and Applications (AINA), April 2024, Lecture Notes on Data Engineering and Communications Technologies, vol 199. Springer, Cham; 2024. P. 43–54.DOI: 10.1007/978-3-031-57840-3_5.

[47] A. Ganesh & S. Ayyasamy Enhanced Approach in VANETs for Avoidance of Collision with Reinforcement Learning Strategy. International Conference on Artificial Intelligence for Smart Community, November 2022, pp 419–428.

Authors

Dr. Jyoti Sunil More pursued B.E. Computer Science Engg. (2003) from Shivaji University, M. Tech. in Computer Engineering (2006) from Dr. B.A.T.U., Lonere and Ph. D. in Computer Engg. From University of Mumbai in 2019. She has more than 20 years of Teaching experience and is currently affiliated to FCRIT, Vashi, affiliated to Mumbai University. She has published several research papers in International conferences and journals. Her area of expertise is Computer Networks, Databases, Data mining, Machine Learning, Data Science, Artificial Intelligence, etc.

Dr. Vaishali V. Sarbhukan (Bodade) has completed her Ph. D (IT) in 2019 from Mumbai University. She has completed her M. E. (Computer) and B. E. (CSE) in 2013 and 2002 from Mumbai University and Amaravati University respectively. Currently she is working as Associate Professor, Department of Information Technology, FCRIT, Vashi, Navi Mumbai. She has 19 years of teaching experience. She has published 26 papers in various Scopus, SCI and international journal and 13 papers in national and international conferences. Her area of interest is security, Machine learning, Data science, AI etc.

Dr. Reena Lokare holds a Doctor of Philosophy in Computer Engineering degree from Mumbai University, India in 2023. She is working as an assistant professor at K. J. Somaiya Institute of Technology, Mumbai, India. She has nearly 21 years of teaching experience. Her areas of expertise include data structures, algorithms, automata theory, soft computing, machine learning and deep learning, bioinformatics, etc. She has nearly 20 research publications and 2 collaborative research projects in the healthcare domain. She can be contacted on: reena.l@somaiya.edu.

How to cite this article:

Jyoti Sunil More, Vaishali V. Sarbhukan (Bodade), Reena Lokare, "Reinforcement Learning-Based AI-Powered Adaptive Routing Protocols for Urban VANETs", International Journal of Computer Networks and Applications (IJCNA), 12(4), PP: 538-553, 2025, DOI: 10.22247/ijcna/2025/33.