Journal of Current Science and Technology, April-June 2026 Vol. 16 No. 2, Article 180
ISSN 2630-0656 (Online)

Cite this article: Kumar, A., Tyagi, S., Dixit, P., & Tyagi, S. S. (2026). Simulation-based evaluation of reinforcement learning-enhanced location-aware routing in urban vehicular ad-hoc networks (VANETs). *Journal of Current Science and Technology*, *16*(2), Article 180. https://doi.org/10.59796/jcst.V16N2.2026.180

Journal of Current Science and Technology
Journal homepage: https://jcst.rsu.ac.th

Simulation-Based Evaluation of Reinforcement Learning-Enhanced Location-Aware Routing in Urban Vehicular Ad-hoc Networks (VANETs)

Arvind Kumar[1,*], Shobha Tyagi[1], Prashant Dixit[2], and S.S. Tyagi[3]

[1]Manav Rachna International Institute of Research Studies, Faridabad 121004, India
[2]Galgotias University, Greater Noida 203201, India
[3]Gurugram University, Samaspur 122018, India

*Corresponding author; E-mail: er.kumararvind@gmail.com

Received 29 October 2025; Revised 27 November 2025; Accepted 24 December 2025; Published online 25 March 2026

Abstract

Vehicular Ad Hoc Networks (VANETs) need robust routing protocols to ensure rapid and reliable data transmission in urban environments characterized by high mobility and highly dynamic topologies. Traditional routing protocols lead to excessive routing overhead, increased hop counts, prolonged end-to-end delays, and reduced packet delivery ratios (PDR), which collectively hinder reliable and efficient data dissemination within intelligent transportation systems (ITS). This research presents a simulation-based evaluation of a Reinforcement Learning (RL)-enhanced Location-Aware Routing (LAR) protocol. By integrating RL with the traditional LAR protocol, the proposed framework dynamically adapts to network fluctuations, thereby minimizing routing overhead, hop counts, and end-to-end delay. Compared against classical routing protocols such as AODV, DSR, and LAR across sparse (50 vehicles), moderate (150), and dense (300) urban traffic scenarios using NS-3 and SUMO, RL-LAR demonstrates superior performance. Improvements ranging from 3% to 12% were observed in PDR, while average end-to-end delay was reduced by 9.7% to 13.8%. Additionally, routing overhead decreased by 4.3% to 8.7%, hop counts were reduced by 15% to 23% and throughput increased by 15% to 31% relative to baseline protocols. These gains were also validated by ANOVA ($p < 0.01$) and found to be suitable for routing in smart cities for future intelligent transportation systems.

Keywords: *location-aware routing; ns3 simulation; reinforcement learning; routing protocols; urban mobility; vehicular ad hoc networks*

1. Introduction

Rapid advancements in intelligent transportation systems (ITS) have ushered in a new era of interconnected vehicles. Communication between vehicles and infrastructure such as roadside units (RSUs) is fundamental to improving road safety, optimizing traffic flow, and enabling autonomous driving (Lee et al., 2024). Vehicular ad hoc networks (VANETs) form the foundation of this vision, providing high-mobility vehicles with the ability to exchange information dynamically and reliably within ever-changing road environments. However, routing in VANETs presents significant research challenges. Unlike traditional wireless networks, VANETs are characterized by high node mobility, frequent and unpredictable topology changes, varying node densities, and the presence of physical obstacles such as buildings and roadside structures. These factors result in rapid link breakages and unpredictable network partitions, making reliable and efficient data delivery a complex task. Urban environments, in particular, exacerbate these issues due to dense traffic, frequent intersections, and rapid fluctuations in vehicle speeds and directions.

Traditional routing protocols, such as ad hoc on-demand distance vector (AODV) (Perkins & Royer, 1999), dynamic source routing (DSR) (Johnson et al., 2001), and location-aided routing (LAR) (Ko & Vaidya, 2000), have been widely utilized for VANET routing. Among these, LAR improves efficiency by leveraging geographic information to limit the route discovery area. Nevertheless, it remains dependent on predefined heuristics and static decision-making. These mechanisms often struggle to adapt to the dynamic nature of urban traffic, leading to increased routing overhead, higher packet loss, and elevated end-to-end delays. Recent advances in artificial intelligence (AI) and machine learning (ML) have introduced new paradigms for addressing these challenges. Reinforcement learning (RL) is an ML method wherein an agent learns to take optimal actions based on experience obtained through trial and error within an environment. RL offers a powerful, model-free approach for dynamic and adaptive decision-making in uncertain environments like VANETs. By continuously interacting with the environment and learning from feedback, RL-based routing agents improve path selection, dynamically avoid congested or broken links, and adapt to real-time changes in network topology.

This paper proposes and evaluates a novel reinforcement learning-Enhanced location-aware routing (RL-LAR) protocol for urban VANETs. The proposed algorithm integrates Q-learning techniques into the traditional LAR framework, enabling intelligent, context-aware next-hop selection based on real-time network conditions within the LAR request zone. The RL agent observes the local network state, specifically neighbor density, link quality, and proximity to the destination node—to learn routing decisions that maximize long-term delivery success while minimizing overhead and latency. To validate the effectiveness of RL-LAR, simulations were conducted using NS-3 on a SUMO-generated urban road scenario of New Delhi, India, across various traffic densities. The performance of RL-LAR was compared against standard routing protocols (AODV, DSR, and standard LAR) in terms of packet delivery ratio (PDR), end-to-end delay, routing overhead, and hop count.

2. Related Work

Vehicular ad hoc networks (VANETs) are integral to smart transportation systems as they facilitate real-time communication between vehicles and infrastructure. The primary challenge in VANETs is the development of routing protocols capable of managing high mobility, rapidly changing network topologies, and urban obstacles. This section examines the transition from conventional to AI-driven routing protocols, emphasizing reinforcement learning (RL)-based methodologies and their evaluation in simulated environments.

Most early VANET routing protocols were derived from mobile ad hoc networks (MANETs). Notable examples include Ad hoc on-Demand distance vector (AODV) (Perkins & Royer, 1999), Dynamic source routing (DSR) (Johnson et al., 2001), and optimized link state routing (OLSR) (Clausen & Jacquet, 2003). While these protocols provided a foundation, their inability to adapt to VANET-specific mobility patterns and urban density resulted in high routing overhead and limited scalability, particularly in densely populated scenarios (Devi & Roopa, 2024). To address these issues, geographic routing protocols such as Greedy perimeter stateless routing (GPSR) (Karp & Kung, 2000) and Location-Aided Routing (LAR) (Ko & Vaidya, 2000) were proposed. These protocols utilize real-time location data to streamline the route discovery process and reduce unnecessary broadcasting. However, studies indicate that GPSR is susceptible to local maxima in urban canyons, and LAR's effectiveness diminishes when GPS data is inaccurate or obstructed by urban infrastructure (Wu et al., 2020). The limitations of static and heuristic-based protocols have prompted a shift toward Artificial Intelligence (AI) for VANET routing. Early machine learning (ML) approaches focused on traffic prediction, link reliability estimation, and congestion avoidance (Priya et al., 2024); however, these methods frequently required labeled data and were poorly suited to real-time, dynamic VANET environments.

Reinforcement learning (RL), particularly Q-learning and deep Q-networks (DQN), has emerged as a compelling solution for dynamic routing. RL formulates routing as a sequential decision-making problem wherein agents (vehicles or infrastructure nodes) learn optimal routing decisions by interacting with the environment and receiving feedback via rewards or penalties (Zhang et al., 2021; Mnih et al., 2015). For instance, Sun et al (2018) proposed a Q-learning-based routing protocol that selects next-hop nodes based on historical delivery success, demonstrating improved PDR and reduced delays in urban scenarios. Similarly, Li et al. (2014) introduced QGrid, which partitions the city into grids with RL-driven next-grid selection, outperforming traditional protocols under high-density conditions. Beyond single-agent RL, Multi-Agent Reinforcement Learning (MARL) has gained traction for its ability to coordinate decisions among clusters of vehicles or between vehicles and RSUs (Zhang et al., 2021; Khan et al., 2022). Trinh et al. (2024) employed a Multi-agent deep deterministic policy gradient (MADDPG) approach to optimize rerouting and congestion management, showing significant improvements in throughput and latency. Hybrid frameworks combining RL with other AI paradigms have been explored to further enhance routing reliability and adaptability. Fuzzy logic has been integrated with RL to manage the uncertainty and imprecision inherent in VANET environments (Suwodjo & Ibrahim, 2024; Yang & Yoo, 2022; Zhang et al., 2021). Swarm intelligence techniques, such as Particle Swarm Optimization (PSO), have been shown to expedite path discovery and improve load balancing. For example, Kumar et al. (2023) proposed a PSO-based data dissemination protocol that dynamically adjusts to vehicle density and mobility patterns. Furthermore, deep learning models such as graph neural networks (GNNs) and Recurrent Neural Networks (RNNs) have been utilized to extract spatiotemporal features for routing decisions, especially in highly dynamic or dense networks (Hu & Lee, 2022; Rajesh et al., 2023; Rao et al., 2024). Bartwal et al. (2025) demonstrated a cluster-based, trustworthy multipath routing protocol using deep RNNs, which improved security and Quality of Service (QoS) in urban scenarios. The integration of edge computing and federated learning into VANET routing is a burgeoning research area. Distributed RL agents deployed on RSUs can collaboratively optimize routing decisions, reducing the computational burden on individual vehicles and lowering latency (Wu, 2023). Alqubaysi et al. (2025) deployed federated RL agents across RSUs, demonstrating improved scalability, privacy, and adaptability. As RL-based protocols become more prevalent, concerns regarding security, privacy, and computational complexity have also been addressed. Blockchain-enabled RL routing (Faisal, 2024) and privacy-preserving federated learning (Hunt et al., 2021) aim to ensure data integrity and user anonymity. Lightweight RL models and safe RL techniques have been designed for resource-constrained vehicular environments (Van Otterlo & Wiering, 2012; Sankar et al., 2023), while the need for real-world validation remains an open challenge (Husain & Sharma, 2018). Most recent studies employ simulation tools such as NS-3, SUMO, and Veins to validate protocol performance under realistic mobility and network conditions (Lansky et al., 2022; Li et al., 2014; Kumar et al., 2023; Hunt et al., 2021). Performance metrics typically include PDR, end-to-end delay, throughput, routing overhead, and scalability, measured across sparse, moderate, and dense traffic settings.

Reinforcement Learning and hybrid AI-based routing protocols have advanced the state of the art in VANET routing, offering significant improvements in delivery reliability, scalability, and adaptation to urban environments. However, challenges remain in reducing computational costs, ensuring security, and bridging the gap between simulation and real-world deployment. This paper builds upon these developments, focusing on a comprehensive, simulation-based evaluation of RL-enhanced location-aware routing under realistic urban scenarios. Based on the preceding review, a comparative summary is presented in Table 1.

The comparison presented in Table 1 highlights the application of advanced RL and AI-based routing in VANETs. However, existing methodologies often lack robust integration with geographical routing and realistic urban simulations. The presented work addresses these gaps by proposing a rigorously evaluated reinforcement learning enhanced location-aware routing (RL-LAR) protocol, designed specifically for dynamic urban VANET scenarios.

3. Methodology

RL-LAR enhances the conventional LAR protocol by embedding a Q-learning agent in each vehicle. This agent continuously learns and adapts its routing policy based on real-time local network conditions and historical experience, aiming to maximize data packet delivery while minimizing delay and routing overhead. The main innovation lies in the dynamic selection of the most reliable and efficient next-hop nodes using RL, instead of relying solely on static location-based heuristics.

Table 1 Comparison table of the related work

Authors / Year	Methodology / Protocol	Key Features	Performance Highlights	Limitations
Perkins & Royer (1999)	AODV (Ad hoc on-demand distance vector)	Reactive routing, route discovery on demand	Simple and scalable for small networks	High overhead and delays in dense urban environments
Johnson et al. (2001)	DSR (Dynamic source routing)	Source routing, route caching	Effective in low-mobility scenarios	Inefficient in high-mobility/dense
Karp & Kung (2000)	GPSR (Greedy perimeter stateless routing)	Geographic routing, greedy forwarding	Reduces flooding, efficient in open areas	Fails in urban canyons (local maxima)
Ko & Vaidya (2000)	LAR (Location-aided routing)	Uses location info to restrict route discovery	Lowers overhead vs. AODV/DSR	Relies on accurate GPS, static heuristics
Sun et al. (2018)	Q-learning-based routing	RL agent selects next hop based on delivery history	Improved PDR, reduced delay in urban test	Single-agent RL, limited scalability
Li et al. (2014)	QGrid (Q-learning grid-based routing)	City grid partitioning, RL-driven next-grid selection	Outperforms LAR in high-density scenarios	Grid mapping complexity
Zhang et al. (2021)	RL + Fuzzy logic	Combines RL with fuzzy logic for uncertainty management	Higher adaptability, robust to imprecision	Increased computational requirements
Kumar et al. (2023)	Particle swarm optimization (PSO)-based protocol	Swarm intelligence for path finding	Adapts to density, improves load balancing	Overhead of swarm processing
Trinh et al. (2024)	Multi-agent DDPG (MADDPG) for routing	Multi-agent RL, deep policy gradient, rerouting	Higher throughput, lower latency	Coordination complexity
Bartwal et al. (2025)	Deep RNN-based secure multipath routing	RNN for spatiotemporal learning, trust/security focus	Increased security, improved QoS in clusters	High resource consumption, simulation only
Alqubaysi et al. (2025)	Federated RL-based routing via RSUs	Collaborative RL, privacy-preserving, federated learning	Improved scalability and privacy	Real-world deployment pending
Faisal (2024)	Blockchain-enabled RL Routing	RL with blockchain for secure authentication and routing	Ensures data integrity, security	Added blockchain and RL overhead
Wu et al. (2020)	RSU-assisted RL-based traffic-aware routing	RSUs assist in RL-based route optimization	Reduced vehicle-side computation, better delay	Infrastructure required

Protocol operation involves state representation (s), action space (a), reward function (r) and Q-learning algorithm. State representation is constituted by the number of neighboring vehicles, their relative position and velocity, signal strength of the available links, distance between the current node and the destination node and traffic density. The set of possible actions space consists of selecting one of the neighboring nodes (within the LAR request zone) as the next hop for the data packet. Each routing action yields a reward, defined as

i. +1 for successful packet delivery to the destination
ii. −1 for packet loss or excessive delay
iii. +0.1 for each hop taken (to encourage shorter paths)

iv. −0.1 as additional penalties for link breakages or retransmissions

The Q-learning algorithm is adopted for its simplicity and effectiveness in model-free environments. The Q-value table is updated after each routing decision based on the observed reward.

When a source vehicle needs to transmit a packet, it uses the LAR request zone to limit the broadcast area. All neighboring vehicles within this zone are potential candidates for the next hop, as shown in Figure 1. The RL agent observes the current state within the LAR request zone, evaluates possible next-hop nodes, and selects the one with the highest Q-value. Then the packet is forwarded to the selected next-hop node. Upon successful delivery or failure, the RL agent receives a reward and updates the Q-table accordingly. Finally, the Q-values are adjusted using the standard Q-learning update rule. Pseudocode for RL-LAR Protocol is presented below:

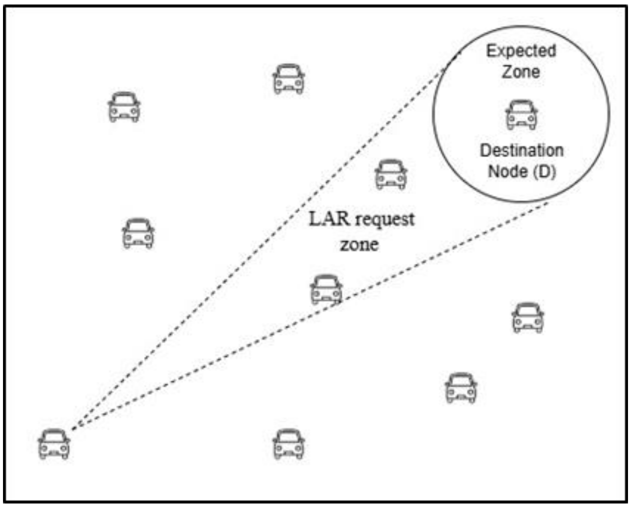

Figure 1 Illustration of the LAR request zone and network area

Algorithm RL-LAR (Reinforcement learning-enhanced location-aided routing)

```
Initialize Q-table Q(s, a) arbitrarily for each state-action pair
Set learning rate α ∈ (0,1), discount factor γ ∈ (0,1), exploration rate ε ∈ (0,1)
For each data packet to be sent from source node S to destination D:
  Initialize state s ← observe_state(S)
  While current node ≠ D do
    Identify neighbors within LAR request zone: N_neighbors
    For each neighbor n ∈ N_neighbors:
      Construct possible action a_n = "forward to neighbor n"
    With probability ε:
      Randomly select action a from possible actions (exploration)
    Otherwise:
      Select action a = argmax_a' Q(s, a') (exploitation)
    Execute action a: forward packet to chosen neighbor n*
    Observe new state s', reward r
    Update Q-table:
      Q(s, a) ← Q(s, a) + α * [r + γ * max_a' Q(s', a') − Q(s, a)]
    s ← s'
    If packet is delivered to D:
      Break
    If forwarding fails or exceeds max hops:
      Break
End
```

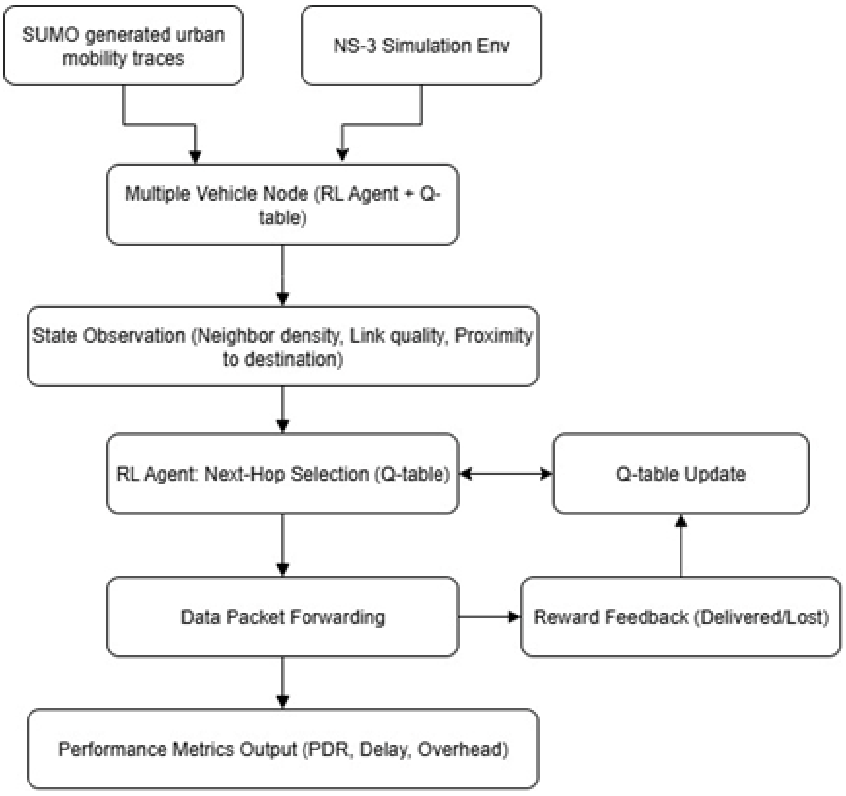

Figure 2 Overall Architecture

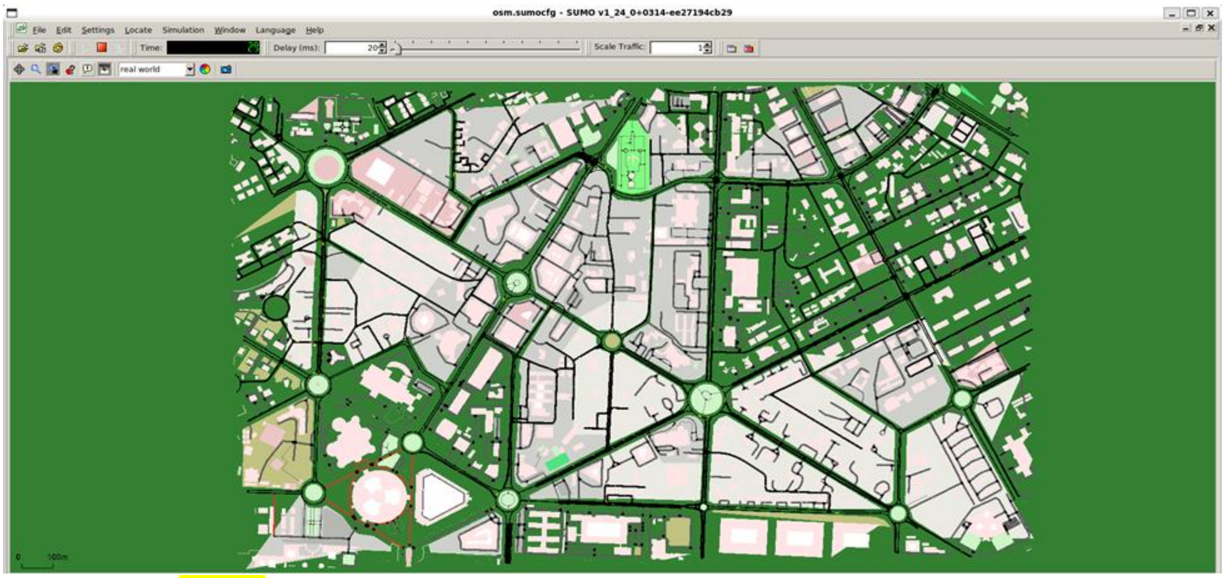

Figure 3 Section of the SUMO-generated scenario of New Delhi used for the simulation

In the proposed algorithm, link failure is handled reactively through the Q-learning framework's reward-penalty system and episode termination logic, emphasizing learning to avoid such links rather than immediate recovery. During packet forwarding, if a link breakage occurs due to mobility-induced disconnection, it triggers a "forwarding failure" condition in the pseudocode loop. The agent observes the new state s' and receives a composite negative reward of -1 for packet loss or excessive delay, plus an additional -0.1 penalty specifically for link breakages or retransmissions. This reward signal updates the Q-table via the standard formula $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$, devaluing the faulty next-hop action and reinforcing avoidance of unstable neighbors in future selections within the LAR request zone.

In this protocol, Decisions are based on real-time network state such as neighbour density and link stability, rather than static heuristics. The RL agent continually improves its routing policy based on cumulative experience, enabling adaptation to dynamic urban VANET conditions. By focusing route discovery within the LAR request zone and learning optimal paths, RL-LAR reduces unnecessary broadcasts and retransmissions. The reward structure incentivizes both successful delivery and minimized number of hops, directly impacting end-to-end delay and efficiency. RL agents self-optimize in response to traffic changes, node failures, and varying densities. Thereby efficiently supporting dense networks through reduced overhead and intelligent next-hop selection. This framework can be extended to incorporate new state variables or adapt to other routing objectives such as energy efficiency or security. An overall architecture of the proposed method is depicted in Figure 2 using a block diagram.

4. Simulation Scenario and Setup

For a rigorous and realistic evaluation of the RL-LAR protocol, vehicular mobility traces were generated using simulation of urban mobility (SUMO), a widely adopted, open-source microscopic traffic simulation tool. The Manhattan-style grid, representing a part of New Delhi, India, was chosen to emulate frequent intersections, traffic lights, and high vehicular density. An area of 1000 m × 1000 m (1 km²) was considered, comprising bidirectional roads, each randomly spaced apart. Each road consists of two lanes per direction with fully functional traffic lights at each intersection. Each light has a fixed cycle of 30 seconds for green and 30 seconds for red. Vehicles were assigned random sources and destinations. A simulation duration of 300 seconds was used for each run. For traffic densities, three traffic density scenarios were evaluated: sparse (50 vehicles), moderate (150 vehicles), and dense (300 vehicles). Vehicles were randomly assigned speeds between 10–50 km/h. The SUMO generated scenario used for the simulation is shown in figure 3. The simulation was performed using NS-3, which allows the direct import of a SUMO-generated mobility trace files for the node movement. The following network parameters were considered for this purpose:

i. Communication standard: IEEE 802.11p (WAVE) at 5.9 GHz.
ii. Transmission range: 300 meters.
iii. Channel bandwidth: 10 MHz.
iv. Data rate: 6 Mbps.
v. Propagation model: Two-ray ground model, suitable for medium to long distances in urban scenarios.
vi. Traffic pattern: Constant bit rate (CBR) and variable bit rate (VBR) UDP flows.
vii. Packet size: 512 bytes (CBR), 1024 bytes (VBR).
viii. Each vehicle initiates periodic UDP packets to randomly chosen destinations.

For each scenario, the following metrics were collected and averaged over 30 simulation runs for statistical significance:

i. **Packet delivery ratio (PDR):** PDR is a key performance metric in networking, particularly for evaluating the reliability of data transmission in ad hoc networks. It quantifies the fraction of packets that are successfully delivered from the source to destination relative to the total packets sent:

$$\text{PDR} = \left(\frac{\text{Number of packets received } (N_r)}{\text{Number of packets sent } (N_s)}\right) \text{x } 100\% \qquad (1)$$

where, N_s is total number of data packets transmitted by the sender and N_r is the number of those packets successfully received and acknowledged by the receiver

ii. **Average end-to-end delay:** It is a performance metric used in networking to measure the average time taken for a data packet to travel from the source to the destination across the network. It is calculated as:

$$\text{Average End-to-End Delay} = \frac{\sum_{i=1}^{N_r} D_i}{N_r} \qquad (2)$$

where, D_i is the end-to-end delay for the i-th successfully delivered packet (in milliseconds).

iii. **Routing overhead:** It is a performance metric in networking protocols that measures the additional network traffic generated by routing control messages (like route requests, replies, and errors) relative to the total traffic. It is typically expressed as a percentage and calculated as:

$$\text{Routing Overhead} = \left(\frac{N_{rt}}{N_{total}}\right) \times 100\% \quad (3)$$

where, N_{rt} is the total number of routing (control) packets transmitted during the simulation, $N_{total}=N_{rt}+N_{data}$, with N_{data} being the number of data packets.

iv. **Average hop count:** It is a performance metric in routing protocols (e.g., in MANETs or WSNs) that quantifies the average number of intermediate nodes (hops) a packet traverses from source to destination. It is calculated as:

$$\text{Average Hop Count} = \frac{\sum_{i=1}^{N_r} H_i}{N_r} \quad (4)$$

where, H_i is the number of intermediate relays or edges in the path count for the i -th successfully delivered packet and N_r is the number of successfully delivered packets.

v. **Throughput:** It is a performance metric in networking that measures the effective data rate of successful transmission from source to destination, typically expressed in bits per second (bps) or kilobits per second (kbps). It is calculated as:

$$\text{Throughput} = \frac{N_r \times S}{T} \quad (5)$$

where, S is the size (in bits) of the successfully delivered packet, N_r is the number of successfully delivered packets, and T is the total time duration of the simulation (in seconds).

Based upon the above scenario, 30 simulations each for sparse, moderate, and dense conditions were run using different random seeds for robust statistical analysis. The results were averaged for use in this research paper. For comparison purposes, AODV, DSR, LAR and RL-LAR protocols were considered. AODV is a standard reactive routing protocol, DSR is a source-routing, on-demand protocol for dynamic networks, LAR is Location-Aided Routing as a geographic baseline, and RL-LAR is the proposed RL-enhanced protocol.

5. Results and Discussion

The results were obtained from the simulation-based evaluation of the proposed RL-LAR protocol in comparison to three baseline routing protocols: AODV, DSR, and standard LAR. The evaluation considered diverse urban traffic densities, sparse, moderate, and dense. It focuses on five core metrics: packet delivery ratio (PDR), average end-to-end delay, routing overhead, average hop count, and throughput. All results were averaged over 30-episode simulation runs for statistical robustness. Tables 2 to 6 presents a comparative summary of the results obtained.

This research paper employs one-way analysis of variance (ANOVA) as the primary statistical test to validate the performance improvements of the proposed RL-LAR protocol compared to baseline protocols (AODV, DSR, and standard LAR). This test was applied to the simulation results across all five key performance metrics, i.e. packet delivery ratio (PDR), average end-to-end delay, routing overhead, average hop count, and throughput. ANOVA uses df = 3,116, critical F = 3.96 at $\alpha = 0.01$ for validation. Resulting F-statistic and p-value are incorporated within the tables 2-6.

Table 2 Packet delivery ratio (PDR) versus traffic density

Metric	Traffic Density	RL-LAR	LAR	DSR	AODV	F-Statistic	p-Value
Packet Delivery Ratio (%)	Sparse (50)	96.2 (±0.5)	92.7 (±1.2)	91.5 (±1.5)	89.8 (±1.8)	122.62	<0.0001
	Moderate (150)	94.8 (±0.8)	89.9 (±1.5)	87.3 (±1.8)	85.4 (±2.0)	196.68	<0.0001
	Dense (300)	92.4 (±1.0)	86.7 (±1.8)	84.5 (±2.0)	83.1 (±2.2)	153.78	<0.0001

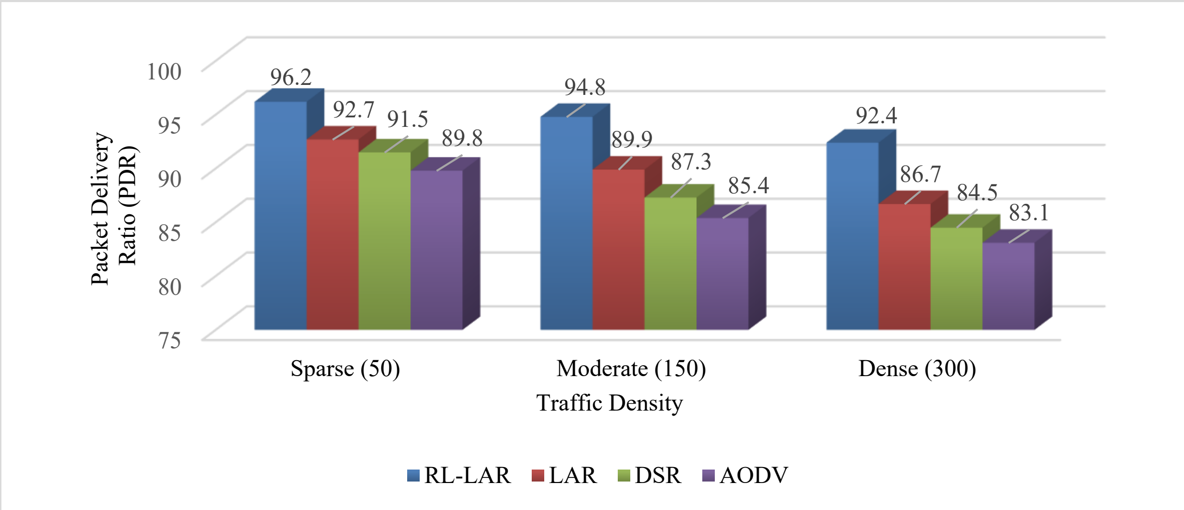

Figure 4 Packet delivery ratio (PDR) versus traffic density

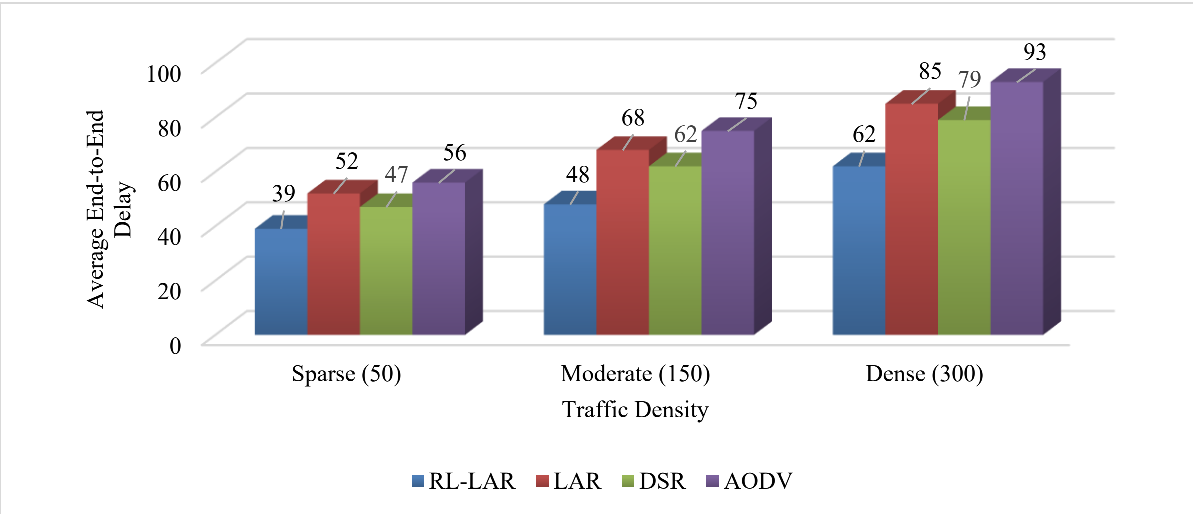

Figure 5 Average end-to-end delay vs. traffic density

Table 2 and Figure 4 show that the RL-LAR protocol consistently achieves a higher PDR across all traffic densities compared to the baseline protocols. In sparse scenarios (50 vehicles), RL-LAR attains an average PDR of 96.2%, outperforming LAR (92.7%), DSR (91.5%), and AODV (89.8%). The performance gap widens in moderate (150 vehicles) and dense (300 vehicles) networks, with RL-LAR maintaining PDRs of 94.8% and 92.4%, respectively. In contrast, LAR, DSR, and AODV display significant drops in PDR under dense conditions, with AODV falling to 83.1%. The improvement in PDR for RL-LAR is attributed to its adaptive routing decisions powered by reinforcement learning. By dynamically evaluating neighbor density, link stability, and real-time network states, RL-LAR effectively avoids congested or unstable links, thereby minimizing packet losses that often occur due to frequent topology changes in urban VANETs. In denser networks, traditional protocols struggle with increased collisions and link breakages, while the learning-driven approach of RL-LAR sustains higher delivery reliability.

Table 3 Average end-to-end delay versus. traffic density

Metric	Traffic density	RL-LAR	LAR	DSR	AODV	F-statistics	*p*-value
Avg. end-to-end delay (ms)	Sparse (50)	39 (±2.0)	52 (±5.0)	47 (±4.5)	56 (±6.0)	75.54	<0.0001
	Moderate (150)	48 (±3.0)	68 (±7.0)	62 (±6.5)	75 (±8.0)	96.13	<0.0001
	Dense (300)	62 (±4.0)	85 (±9.0)	79 (±8.5)	93 (±10.0)	77.07	<0.0001

Table 4 Routing overhead versus traffic density

Metric	Traffic density	RL-LAR	LAR	DSR	AODV	F-statistic	*p*-value
Routing Overhead (%)	Sparse (50)	4.3 (±0.3)	7.1 (±0.7)	8.2 (±0.8)	9.7 (±1.0)	281.22	<0.0001
	Moderate (150)	6.2 (±0.4)	10.5 (±1.0)	12.3 (±1.2)	13.9 (±1.4)	290.24	<0.0001
	Dense (300)	8.7 (±0.5)	13.8 (±1.3)	15.7 (±1.5)	16.2 (±1.6)	208.71	<0.0001

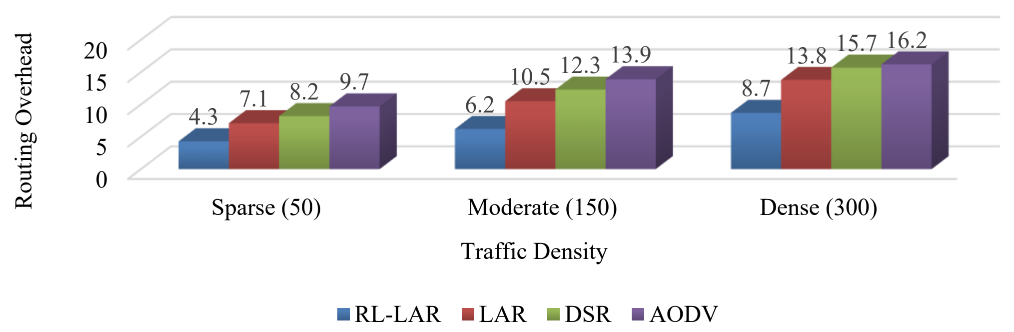

Figure 6 Routing overhead versus traffic density

Table 3 and Figure 5 show that RL-LAR demonstrates a substantial reduction in average end-to-end delay across all scenarios. For sparse traffic, RL-LAR achieves an average delay of 39 ms, compared to 52 ms (LAR), 47 ms (DSR), and 56 ms (AODV). This delay advantage becomes more pronounced in dense traffic, where RL-LAR records 62 ms, while LAR, DSR, and AODV experience delays of 85 ms, 79 ms, and 93 ms, respectively. The reduced delay in RL-LAR is a direct consequence of its reward structure, which penalizes excessive hops and delays, thereby incentivizing the agent to learn shorter and more stable routes. Unlike heuristic-based protocols that can be trapped in suboptimal paths or local maxima, especially in urban topologies with frequent intersections RL-LAR adapts in real time to select routes that minimize overall transmission time.

Table 4 and Figure 6 show that routing overhead, measured as the proportion of control packets to total network traffic, is significantly lower in RL-LAR. In moderate-density scenarios, RL-LAR incurs an overhead of 6.2%, compared to 10.5% for LAR, 12.3% for DSR, and 13.9% for AODV. Under dense conditions, the overhead for RL-LAR remains below 8.7%, while the other protocols observe sharp increases, exceeding 16% in the case of AODV.

This reduction in overhead is primarily due to the intelligent next-hop selection of RL-LAR and restricted route discovery within the LAR request zone. The ability of the RL agent to recognize and avoid unreliable nodes reduces the need for repeated route discoveries and retransmissions, thereby conserving bandwidth and improving scalability, a factor that is especially vital in congested urban environments.

Table 5 Average hop count vs. traffic density

Metric	Traffic density	RL-LAR	LAR	DSR	AODV	F-statistic	*p*-value
Average Hop Count	Sparse (50)	2.4 (±0.2)	2.8 (±0.3)	2.9 (±0.3)	3.1 (±0.4)	27.37	<0.0001
	Moderate (150)	2.7 (±0.3)	3.1 (±0.4)	3.3 (±0.4)	3.7 (±0.5)	31.52	<0.0001
	Dense (300)	3.3 (±0.4)	3.8 (±0.5)	4 (±0.5)	4.3 (±0.6)	20.78	<0.0001

Table 6 Throughput versus Traffic density

Metric	Traffic density	RL-LAR	LAR	DSR	AODV	F-statistic	*p*-value
Throughput (kbps)	Sparse (50)	410 (±12)	377 (±20)	360 (±22)	345 (±25)	56.58	<0.0001
	Moderate (150)	392 (±15)	342 (±25)	325 (±27)	310 (±30)	61.52	<0.0001
	Dense (300)	370 (±18)	315 (±28)	298 (±30)	283 (±32)	57.11	<0.0001

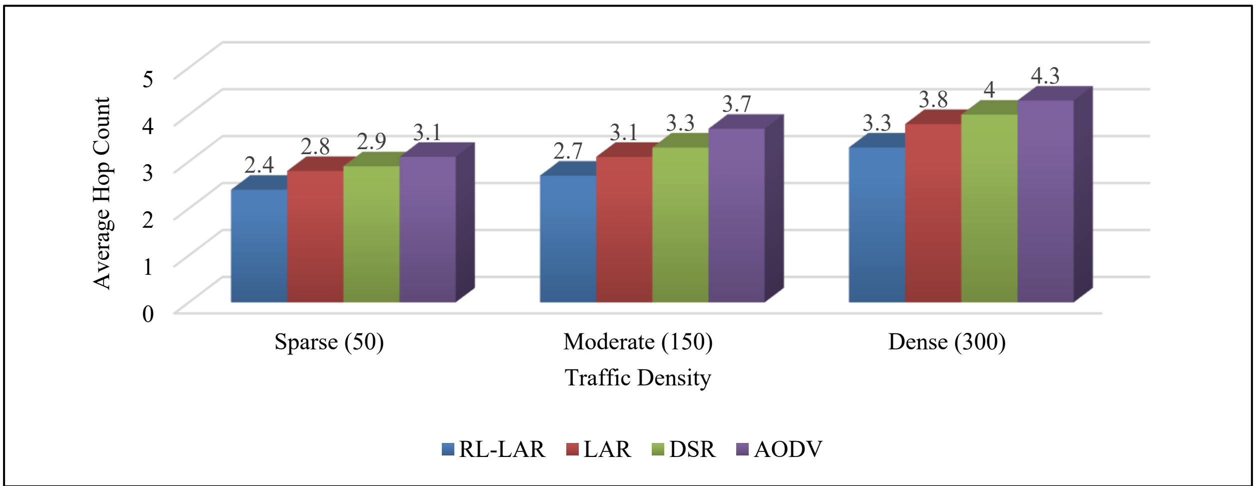

Figure 7 Average hop count vs. traffic density

Table 5 and Figure 7 show that RL-LAR consistently achieves lower average hop counts in all traffic settings. For example, in sparse networks, the average hop count for RL-LAR is 2.4, while LAR, DSR, and AODV register hop counts of 2.8, 2.9, and 3.1, respectively. In dense conditions, RL-LAR maintains an average hop count of 3.3, compared to 3.8 for LAR, 4.0 for DSR, and 4.3 for AODV.

Fewer hops are indicative of more direct and efficient routing, which not only reduces delay but also lessens the likelihood of packet loss due to intermediate node failures. The learning-driven approach of RL-LAR enables the selection of routes with optimal trade-offs between reliability and path length, further enhancing network performance.

Table 6 and Figure 8 show that throughput, measured in kbps, is consistently higher for RL-LAR. In dense urban scenarios, RL-LAR achieves an average throughput of 370 kbps, outperforming LAR (315 kbps), DSR (298 kbps), and AODV (283 kbps). Higher throughput in RL-LAR is a direct outcome of its superior packet delivery and lower end-to-end delays, which collectively facilitate a greater volume of successful data transmission within the simulation period. Efficient bandwidth utilization and minimized retransmissions further contribute to the observed throughput gains.

Figure 9 depicts the convergence of the Q-learning agent over 600 simulated episodes of packet routing, demonstrating the reinforcement learning dynamics underlying the improved performance of RL-LAR. The 50-episode moving average (red line) highlights policy stability, while instantaneous rewards (blue line) capture experimental volatility. Each episode represents a delivery attempt within the LAR request zone.

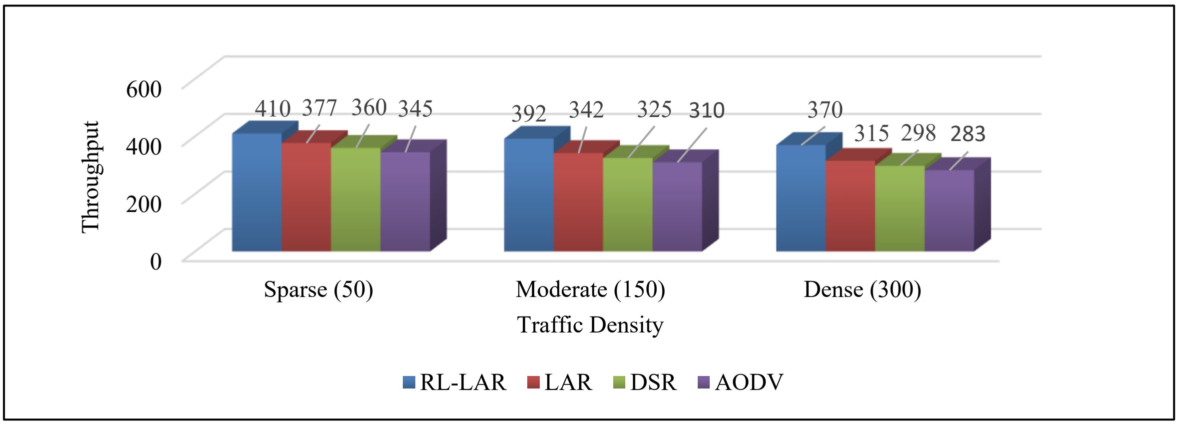

Figure 8 Throughput versus traffic density

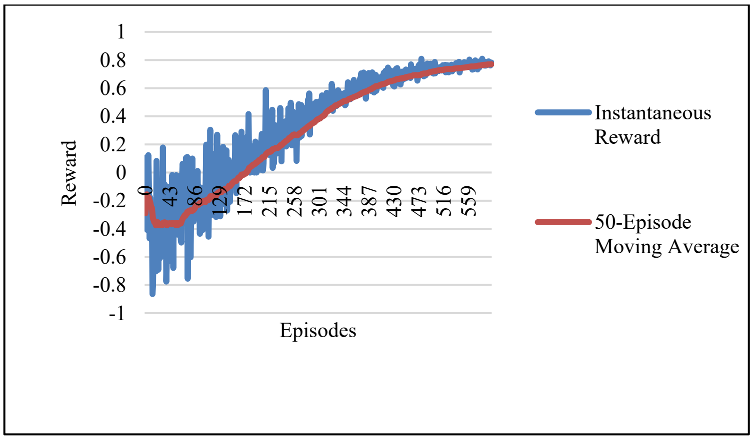

Figure 9 RL-LAR convergence

All performance improvements of RL-LAR over baseline protocols were validated using one-way ANOVA tests, which revealed p-values < 0.01 for all key metrics, indicating strong statistical significance. Additionally, RL-LAR exhibited lower variance in results across different random seeds and simulation runs, underscoring its robustness and stability in dynamic urban traffic environments as compared to the results reported Sun et al. (2018). Furthermore, compared to QGrid protocol by Li et al. (2014), which leverages Q-learning for adaptive VANET routing in urban settings, RL-LAR demonstrates clear superiority through real-time, per-vehicle online learning versus the static, offline Q-table used by QGrid, which remains fixed post-training. This enables dynamic adaptation to topology changes without grid partitioning constraints. While QGrid selects next-grids based on historical delivery success and unaddressed broadcasting, RL-LAR refines states in a refined request zone and actions with nuanced rewards achieving improved PDR, reduced delays and gained throughput, outpacing baselines by 3-31% with the use of modern NS-3/SUMO and validations with ANOVA-backed statistical validation (p<0.01).

It was observed that RL-LAR rapidly adapts to sudden topological changes, such as road blockages or variable traffic signals, demonstrating real-world applicability for smart city deployments. The protocol also scales efficiently with increasing network size and density, preserving performance without excessive overhead. While RL-LAR offers significant advantages, its performance is sensitive to the learning rate and reward structure. Furthermore, real-world deployment may require additional optimization for computational efficiency and safe exploration. The RL-LAR protocol offers substantial improvements in reliability, efficiency, and scalability over classical routing protocols in urban VANET environments. Its reinforcement learning foundation enables real-time adaptation to network dynamics, making it highly suitable for future intelligent transportation systems.

6. Conclusion

The proposed RL-LAR protocol effectively integrates reinforcement learning with location-aware routing, allowing for dynamic, intelligent next-hop selection in urban VANETs. Simulation results demonstrate that this protocol consistently outperforms conventional protocols (AODV, DSR, and LAR) in terms of packet delivery ratio (PDR), end-to-end delay, routing overhead, average hop count, and throughput, with the most significant advantages observed under dense urban traffic. The protocol adapts well to frequent topology changes and scales efficiently as network density increases, indicating strong potential for practical deployment in real-world urban environments. While this research utilizes realistic mobility and network models, it is limited to simulations. Real-world factors such as GPS inaccuracies, signal interference, and hardware constraints are not fully captured. The deployment of reinforcement learning agents across all vehicles may introduce computational and memory overhead, especially for resource-constrained devices. Furthermore, the performance of RL-LAR is dependent on precise tuning of RL parameters, such as the learning rate, discount factor, and reward structure. Suboptimal tuning could lead to degraded performance or instability. Excessive exploration during learning phase can temporarily degrade network performance, whereas insufficient exploration may prevent the discovery of optimal routes.

Computational and memory overhead of Q-learning agents, particularly in dense networks, needs further optimization. Future iterations may integrate lightweight RL variants, such as model-based approximations or federated learning to enhance privacy-preserving adaptations. Extending the protocol to multi-agent RL (MARL) frameworks could enable coordinated routing among vehicle clusters. Hybrid integrations with emerging paradigms, including fuzzy logic for uncertainty management, swarm intelligence for dynamic path discovery, or deep RL for spatiotemporal prediction, may be explored for further optimization, thus paving the way for seamless smart city integration.

The state-aware next-hop selection of RL-LAR, based on neighbor density and link quality, promotes scalable, bandwidth-efficient networks under IEEE 802.11p. This conserves resources in dense smart cities. Societally, it promises environmental benefits through optimized vehicle emissions and monetary benefits in logistics. However, resource-constrained vehicles may face additional strain from RL computational overhead, potentially widening technological gaps in emerging markets.

7. Abbreviations

Abbreviation	Full Term
AODV	Ad hoc On-Demand Distance Vector
AI	Artificial Intelligence
ANOVA	Analysis of Variance
BSM	Basic Safety Message
CBR	Constant Bit Rate
DDPG	Deep Deterministic Policy Gradient
DQN	Deep Q-Network
DSR	Dynamic Source Routing
GNN	Graph Neural Network
GPS	Global Positioning System
GPSR	Greedy Perimeter Stateless Routing
ITS	Intelligent Transportation System
LAR	Location-Aided Routing
MADDPG	Multi-Agent Deep Deterministic Policy Gradient
MANET	Mobile Ad Hoc Network
MARL	Multi-Agent Reinforcement Learning
ML	Machine Learning
NS-3	Network Simulator 3
OLSR	Optimized Link State Routing
PDR	Packet Delivery Ratio
PSO	Particle Swarm Optimization
QoS	Quality of Service
RL	Reinforcement Learning
RL-LAR	Reinforcement Learning–Enhanced Location-Aware Routing
RNN	Recurrent Neural Network
RSU	Roadside Unit
SUMO	Simulation of Urban Mobility
UDP	User Datagram Protocol
VANET	Vehicular Ad Hoc Network
VBR	Variable Bit Rate
WAVE	Wireless Access in Vehicular Environments

8. CRediT Statement

Arvind Kumar: Conceptualization, Methodology, Writing – Original Draft, Writing – Review & Editing, Visualization.
Shobha Tyagi: Supervision, Formal Analysis.
Prashant Dixit: Supervision, Resources, Software.
S. S. Tyagi: Supervision, Resources.

9. References

Alqubaysi, T., Asmari, A. F. A., Alanazi, F., Almutairi, A., & Armghan, A. (2025). Federated learning-based predictive traffic management using a contained privacy-preserving scheme for autonomous vehicles. *Sensors*, *25*(4), Article 1116. https://doi.org/10.3390/s25041116

Bartwal, H., Sivaraman, H., & Kumar, J. (2025). A cluster-based trusted secure multipath routing protocol for mobile ad hoc networks. *International Journal of Computer Networks & Communications*, *17*(3), 89–109. https://doi.org/10.5121/ijcnc.2025.17306

Clausen, T., & Jacquet, P. (2003). Optimized Link State Routing Protocol (OLSR) RFC 3626. *Network Working Group. Internet Engineering Task Force (IETF).* https://doi.org/10.17487/rfc3626

Devi, Y. S., & Roopa, M. (2024). *Performance analysis of routing protocols in vehicular ad hoc networks* [Conference presentation]. *Evolution in Signal Processing and Telecommunication Networks. ICMEET 2023.* Lecture Notes in Electrical Engineering, Singapore. https://doi.org/10.1007/978-981-97-0644-0_36

Faisal, S. M. (2024). Blockchain-enabled security model for VANETs: A novel approach to authentication and position. *Research Square.* https://doi.org/10.21203/rs.3.rs-4859344/v1

Hu, H., & Lee, M. J. (2022). Graph neural network-based clustering enhancement in VANET for cooperative driving [Conference presentation]. *2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIC).* IEEE, Jeju Island, Korea. https://doi.org/10.1109/ICAIIC54071.2022.9722625

Hunt, N., Fulton, N., Magliacane, S., Hoang, T. N., Das, S., & Solar-Lezama, A. (2021). Verifiably safe exploration for end-to-end reinforcement learning [Conference presentation]. *Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control*, New York, US. https://doi.org/10.1145/3447928.3456653

Husain, A., & Sharma, S. C. (2018). Implementation of geographical location based routing protocols in vehicular environment. *International Journal of System Assurance Engineering and Management*, *9*(1), 18-25. https://doi.org/10.1007/s13198-016-0425-3

Johnson, D. B., Maltz, D. A., & Broch, J. (2001). DSR: The dynamic source routing protocol for multi-hop wireless ad hoc networks. *Ad Hoc Networking*, *5*(1), 139-172.

Karp, B., & Kung, H. T. (2000). GPSR: Greedy perimeter stateless routing for wireless networks [Conference presentation]. *Proceedings of the 6th annual international conference on Mobile computing and networking*. New York, US. https://doi.org/10.1145/345910.345953

Khan, M. U., Hosseinzadeh, M., & Mosavi, A. (2022). An intersection-based routing scheme using Q-learning in vehicular ad hoc networks for traffic management in the intelligent transportation system. *Mathematics*, *10*(20), Article 3731. https://doi.org/10.3390/math10203731

Ko, Y., & Vaidya, N. H. (2000). Location-aided routing (LAR) in mobile ad hoc networks. *Wireless Networks*, *6*(4), 307–321. https://doi.org/10.1023/a:1019106118419

Kumar, A., Dixit, P., & Tyagi, S. S. (2023). Particle swarm optimization for efficient data dissemination in VANETs [Conference presentation]. *International Conference on Artificial-Business Analytics, Quantum and Machine Learning.* Springer Nature Singapore, Singapore. https://doi.org/10.1007/978-981-97-2508-3_4

Lansky, J., Rahmani, A. M., & Hosseinzadeh, M. (2022). Reinforcement learning-based routing protocols in vehicular ad hoc networks for intelligent transport system (ITS): A survey. *Mathematics*, *10*(24), Article 4673. https://doi.org/10.3390/math10244673

Lee, S., Oh, J., Kim, M., Lim, M., Yun, K., Yun, H., Kim, C., & Lee, J. (2024). A study on reducing traffic congestion in the roadside unit for autonomous vehicles using BSM and PVD. *World Electric Vehicle Journal*, *15*(3), Article 117. https://doi.org/10.3390/wevj15030117

Li, R., Li, F., Li, X., & Wang, Y. (2014). QGrid: Q-learning based routing protocol for vehicular ad hoc networks [Conference presentation]. *2014 IEEE 33rd international performance computing and communications conference (IPCCC)*. IEEE, Austin, TX, USA. https://doi.org/10.1109/PCCC.2014.7017079

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis,

D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529-533. https://doi.org/10.1038/nature14236

Perkins, C., & Royer, E. (1999). Ad-hoc on-demand distance vector routing [Conference presentation]. *IEEE Workshop on Mobile Computing Systems and Applications (WMCSA).* New Orleans, LA, USA. https://doi.org/10.1109/MCSA.1999.749281

Priya, K. C., Sharma, S., Kumar, M. S., Gangwar, P. K., Aarthi, R., & Kumar, R. S. (2024). Optimizing urban mobility in smart cities through deep learning-based traffic management [Conference presentation]. *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, Kamand, India. https://doi.org/10.1109/ICCCNT61001.2024.10724627

Rajesh, P., Shajin, F. H., Ansal, V., & Kumar, B. V. (2023). Enhanced artificial transgender longicorn algorithm and recurrent neural network–based enhanced DC–DC converter for torque ripple minimization of BLDC motor. *Journal of Current Science and Technology*, *13*(2), 182–204. https://doi.org/10.59796/jcst.V13N2.2023.1735

Rao, P. S., Polisetty, V. R. M., Jayanth, K. K., Manoj, N. S., Mohith, V., & Kumar, R. P. (2024). Deep adaptive algorithms for local urban traffic control: Deep reinforcement learning with DQN [Conference presentation]. *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*. IEEE, Bengaluru, India. https://doi.org/10.1109/IDCIoT59759.2024.10467630

Sankar, P., Joel, M. R., & Husain, A. J. (2023). DIB-a novel optimized VANET traffic management using a deep neural network. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(10s), 484–491. https://doi.org/10.17762/ijritcc.v11i10s.7657

Sun, Y., Lin, Y., & Tang, Y. (2018). A reinforcement learning-based routing protocol in VANETs [Conference presentation]. *International Conference in Communications, Signal Processing, and Systems.* Springer Singapore, Singapore. https://doi.org/10.1007/978-981-10-6571-2_303

Suwodjo, R. A., & Ibrahim, Z. B. (2024). Modeling of an adaptive HHO gas controller based on fuzzy logic and polynomial function controls to improve engine torque of gasoline engine. *Journal of Current Science and Technology, 14*(3), Article 72. https://doi.org/10.59796/jcst.V14N3.2024.72

Trinh, H. T., Bae, S. H., & Tran, D. Q. (2024). Combining multi-agent deep deterministic policy gradient and rerouting technique to improve traffic network performance under mixed traffic conditions. *SIMULATION*, *100*(10), 1033-1051. https://doi.org/10.1177/00375497241237831

Van Otterlo, M., & Wiering, M. (2012). Reinforcement learning and Markov decision processes. *Adaptation, Learning, and Optimization*. Berlin, Germany. https://doi.org/10.1007/978-3-642-27645-3_1

Wu, J., Fang, M., Li, H., & Li, X. (2020). RSU-assisted traffic-aware routing based on reinforcement learning for urban vanets. *IEEE Access*, *8*, 5733-5748. https://doi.org/10.1109/ACCESS.2020.2963850

Wu, Z. (2023). Design of dynamic traffic information prediction system based on multi-source information fusion [Conference presentation]. *2023 2nd International Conference on Artificial Intelligence and Intelligent Information Processing (AIIIP).* IEEE, Hangzhou, China. https://doi.org/10.1109/AIIIP61647.2023.00020

Yang, Q., & Yoo, S. J. (2022). Grouped intersection-based routing using reinforcement learning for urban VANETs [Conference presentation]. *2022 13th international conference on information and communication technology convergence (ICTC).* IEEE, Jeju Island, Korea. https://doi.org/10.1109/ICTC55196.2022.9952627

Zhang, W., Yang, X., Song, Q., & Zhao, L. (2021). V2V routing in VANET based on fuzzy logic and reinforcement learning. *International Journal of Computers Communications & Control*, *16*(1). Article 4123. https://doi.org/10.15837/ijccc.2021.1.4123