CIENTIFIC Publishing Limited

Journal of Intelligent Communication
https://ojs.ukscip.com/index.php/jic

Article

Dynamic Q-Learning-Based Handover in VANETs: An Approach for Li-Fi Based Handover Techniques

Chinmoy Sailendra Kalita * and **Maushumi Barooah**

Department of Computer Application, Assam Engineering College, Guwahati, Assam 781013, India
* Correspondence: chinmoykalita81@gmail.com

Received: 4 July 2025; **Revised:** 13 August 2025; **Accepted:** 16 August 2025; **Published:** 1 September 2025

Abstract: Vehicular Ad hoc Networks (VANETs) face significant challenges in maintaining seamless connectivity due to frequent handovers caused by high vehicular mobility, fluctuating signal strength, and dynamic traffic conditions. Conventional static threshold-based handover schemes are unable to adapt to these variations, often resulting in excessive latency, packet loss, handover failures, and poor reliability. Such drawbacks are particularly detrimental to critical applications like collision avoidance and real-time traffic management. To address these limitations, this paper proposes a dynamic handover framework based on Q-learning integrated with Li-Fi communication technology. The framework intelligently evaluates real-time parameters, such as vehicle speed, signal strength, and network occupancy, to make adaptive handover decisions rather than relying on preset thresholds. The proposed approach is modeled and tested using OMNeT++ and SUMO simulation platforms. Experimental results demonstrate that the proposed framework achieves a handover success rate of 90.0%, reduces failure rates to 9.8%, and limits ping-pong handovers to 4.7%. Moreover, throughput improves by 20.3%, and handover delay decreases by 30.2% compared to other state-of-the-art approaches. The packet delivery ratio is sustained at 95.6% even under high traffic density, indicating system robustness and scalability. These findings highlight that reinforcement learning-based handover management provides a promising solution for next-generation VANETs, offering adaptability, reliability, and efficiency for emerging intelligent transportation systems and critical vehicular applications, thereby ensuring safety.

Keywords: Adaptive Communication; Li-Fi; Handover Decision; Reinforcement Learning; Vehicular Networks; Q-Learning

1. Introduction

Vehicular Ad Hoc Networks (VANET) refer to the automobile part of the Intelligent Transport System (ITS), which enhances the intelligence on both the vehicle and infrastructure sides to communicate with each other towards the safety of roads, management of traffic flow, and real-time services like collision avoidance and optimal path. This system enables the sharing of critical safety information between Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications, providing potential collision warnings and real-time traffic information to support adequate and timely responses and informed decisions on the road. The high speed of vehicles, frequent change in network connectivity, and the varying availability of communication infrastructure make it especially difficult in VANET handover management [1].

Application requirements in VANETs are extremely demanding and need to be supported by handovers, including collision warnings, emergency alerts, and notifications of road hazards. All of these require high reliability with low latency; delay or packet loss caused by handover failures or delays leads to serious degradation in Quality of Service (QoS) in scenarios that involve split-second decision-making, such as highway driving or navigating busy urban centers. Highways, where vehicles travel at a high speed, and congested city centers, where traffic is usually locked up, represent environments for which efficient management of handover is critical [2].

Generally, conventional handover techniques used in VANETs are based on using signal strength thresholds and static criteria to decide when to execute a handover. They are mostly applied in stationary, low-speed environments where network conditions do not change often. However, in high mobility environments, static approaches have highly suboptimal performances [3]. These numerous handovers create overheads, and network performance degrades since several such handovers may be initiated prematurely or late, thus causing losses of packets, increased latencies, and generally poor communication quality under such safety-critical conditions [4].

Therefore, these critical challenges that exist within traditional handover management techniques call for developing adaptive and dynamic handover management techniques that may function in real-time response to alterations in network conditions.

Light Fidelity (Li-Fi) is another choice that can be utilised in Radio frequency (RF) based communication in VANETs, where vertical handovers occur in heterogeneous networks. It transmits data through the visible light spectrum. The advantages of Li-Fi over traditional RF communication include more bandwidth, less interference, and more security [5]. Kalita et al. [6] in 2020 proposed a handover mechanism of Li-Fi for VANETs using on vehicle Li-Fi sensors as well as an Anonymous Announcement System (AAS) on RSUs to enable Li-Fi VANET handover actively. They found that in comparison to systems based upon the RF techniques, their system outperforms, especially in terms of latency as well as the Packet data rate. Li-Fi-based systems rely upon line of sight (LOS) communication; they might face coverage gaps, particularly in the urban environment where obstacles are very likely. Therefore, integration of Li-Fi with adaptive decision-making algorithms like reinforcement learning is essential for taking full advantage of its capabilities in VANETs [7].

Despite significant advances in vehicular communication, efficient and reliable handover in high-mobility scenarios remains a bottleneck. Static threshold-based handover methods cannot adapt to dynamic conditions, such as varying signal quality, traffic density, and vehicular mobility. Additionally, while Li-Fi promises speedy and interference-free communication, its line-of-sight dependency presents another set of challenges, particularly in urban scenarios with obstacles. This study addresses the gaps by proposing Q-learning-based handover for Li-Fi-based VANETs. The dynamic learning-based approach utilizes real-time feedback to optimize handover decisions, enhancing adaptability and minimizing communication interruptions.

The contributions from this research are concisely enumerated as follows:

1. Existing methods largely rely on static, threshold-based criteria that fail to adapt to dynamic vehicular environments. To address this gap, a Q-learning-based framework is introduced that dynamically adjusts to real-time conditions, including signal strength, mobility, and network occupancy.
2. Previous studies have highlighted the potential of Li-Fi in vehicular networks but have not incorporated adaptive decision-making mechanisms. This work bridges that gap by combining Li-Fi communication with reinforcement learning, enabling reliable handovers under varying traffic and signal conditions.
3. Many existing solutions overlook dynamic traffic load and signal strength variability, leading to inefficiencies in decision making. To fill this gap, this study formulates a handover decision model based on queuing theory and entropy-based estimation, providing more robust input parameters for adaptive learning.
4. Deep reinforcement learning techniques demonstrate adaptability but face high computational complexity and scalability issues. This gap is addressed by implementing a tabular Q-learning algorithm with an ε-greedy policy, enabling lightweight, real-time decision-making suitable for large-scale vehicular networks.
5. Previous research has often emphasized security or theoretical models while overlooking quantitative performance benchmarks such as latency, throughput, and packet delivery ratio. To close this gap, extensive OMNeT++ and SUMO simulations were conducted, demonstrating measurable improvements in comparison to state-of-the-art research works such as DTe-DQN.

This paper is organised as follows. Chapter 2 presents the related work. Chapter 3 describes the methodology, followed by the results and discussions in Chapter 4. Finally, the conclusion, along with future work is addressed in Chapter 5.

2. Related Works

Recent studies have presented numerous handover mechanisms to support VANETs. Duo et al. [8] introduced an SDN-based approach for hybrid networks, but had scalability issues. Dwivedi et al. [9] proposed the B-HAS protocol for safe handover, but it had a very high computational overhead. Aboud et al. [10] tried to minimize delays in 5G VANETs but had difficulties in heterogeneous cases. Xie et al. [11] and Son et al. [12] proposed blockchain-based lightweight protocols that address scalability issues. Alam et al. [13] presented reviews on handover techniques without the practical implementation. Rosli et al. [14] presented the issue of handovers in 5G with energy costs. Costa et al. [15] optimized video distribution in a handover process without touching security aspects. Oladosu et al. [16] provided a metaheuristic algorithm and neglected adaptability. Additionally, the works of Anilkumar and Rafeek [17] have proposed the ”Soteria” certificate-less mechanism, which could be subject to latency issues due to the implementation of blockchain in changing scenarios.

Recent developments in intelligent handover management systems, particularly the implementation of machine learning approaches, have shown promising solutions for addressing these deficiencies. One such promising approach is the application of the model-free reinforcement learning (RL) technique called Q-Learning in optimizing handover decisions within VANETs. Q-Learning enables the system to learn and adapt parameters such as vehicle velocity, network traffic, and signal strength real-time, thereby refining the decision-making process. Q-Learning-based systems constantly evolve their policies, deciding their future actions based on the changing network state, so more efficient and adaptive ways of handover management could be achieved as demonstrated by Liang et al. [18]. Overall, this results in a reduction of the handover frequency, while maintaining optimum system performance in terms of latency, throughput, and QoS as in Zhang et al. [19]. Kafle et al. [20] provided the first Q-Learning application for VANET, where decisions on handover are made based on real-time network conditions, informed by a Q-Learning evaluation. Proving their work, the authors demonstrated that the system based on RL techniques can outperform threshold-based handover schemes under network fluctuations. Subsequent studies, such as those carried out by Siriwardhana et al. [21], make Q-Learning for handover management more feasible by incorporating additional parameters, such as vehicle density, traffic load, and signal quality.

Shahwani et al. [22] developed an extensive comparison between traditional threshold-based handover methods and Q-Learning-based systems in VANETs. According to their results, machine learning-based handover mechanisms reduce latency by significant amounts and boost packet delivery ratios since they adapt to real-time network conditions dynamically. Wang and Li [23] proposed a double-deep Q-learning-based handover management system for mmWave heterogeneous networks with dual connectivity in order to enhance the efficiency of handovers and minimize latency in high-mobility scenarios. The approach could be computationally intensive as the scale of the network increases. Uppoor et al. [24] proposed a reinforcement learning-based handover parameter adaptation method using LSTM-aided digital twins for ultra-dense networks, which improves the accuracy of predictions and adaptability in dynamic environments. However, this method depends on a vast amount of data for training, and its applicability in real-time is still limited to fast-changing scenarios. Therefore, a hybrid technology might be necessary, combining Li-Fi with RF-based communication systems, in order to achieve seamless connectivity from one environment to another.

Unlike static handover methods, which depend on predefined thresholds, the integration of Q-Learning into the handover actually enhances the system's adaptability with respect to changes in network conditions, but it also opens the possibility of taking better advantage of network resources. A brief comparative summary of the above-discussed handover mechanism based on the focus area is given in **Table 1**.

Extensive research has been conducted on VANET handover management, utilizing various technologies such as Wi-Fi, blockchain, and 5G. However, a lot of gaps remain in the following aspects:

Adaptive Solutions: All blockchain and metaheuristic methods focus on static handover decisions, which are not adaptive to dynamic vehicular conditions, such as changing traffic patterns and mobility.

Limited Li-Fi Integration: Handover mechanisms of emerging Li-Fi technology are underexplored. Largely, not many researchers apply reinforcement learning or Q-learning in real-time, adaptive optimization of handover for handling the change in traffic density, strength of signals, or even congestion.

Performance Overlooked: Research focuses more on security (e.g., blockchain solutions) than on performance metrics such as latency, success rates, and throughput, especially in high-mobility scenarios.

This paper addresses the gap by proposing an adaptive Li-Fi handover technique that relies on a dynamic Q-learning algorithm. Inclusion of vehicular mobility, traffic patterns, network occupancy, and signal strength in its decision-making procedure will assist in significantly optimizing the performance of the handover in VANET, thereby reducing latency, increasing handover success rates, and improving overall network throughput. This is a new contribution to the domain of communication, especially in high-mobility scenarios where the traditional static methods are likely to fail.

Table 1. Comparative summary of existing handover mechanisms in VANETs.

Study	Year	Technology	Focus Area	Strength	Limitations
Duo et al. [8]	2020	SDN hybrid	Scalability	Provides flexible network management through centralized control	It finds difficulty in handling large-scale vehicular networks; control overhead increases significantly as vehicle density rises
Dwivedi et al. [9]	2023	Blockchain (B-HAS)	Authentication, security	Ensures secure authentication and prevents identity spoofing	Requires complex cryptographic operations; introduces high computational and communication overhead not suitable for real-time handovers
Aboud et al. [10]	2021	5G optimization	Delay reduction	Achieves reduced delay in homogeneous 5G environments	Works well in homogeneous 5G scenarios but has limited adaptability in heterogeneous multi-access environments
Xie et al. [11]; Son et al. [12]	2022, 2023	Blockchain	Lightweight protocols	Provide lightweight handover authentication with enhanced security	Still face scalability bottlenecks; latency may increase when vehicle numbers grow rapidly
Rosli et al. [14]	2023	5G	Energy efficiency	Reduces energy consumption during handover	Focuses on power saving but neglects reliability and latency under dense traffic conditions
Oladosu et al. [16]	2023	Metaheuristic	Intelligent handover	Uses intelligent search techniques for handover optimization	Lacks real-time adaptability; performance degrades in highly dynamic mobility scenarios
Liang et al. [18]	2014	Q-learning	Adaptive handover	Demonstrates adaptability to dynamic network conditions	Does not integrate Li-Fi, limiting bandwidth utilization and robustness in urban settings
Wang & Li [23]	2023	Double-deep Q-learning	mmWave handover	Provides strong adaptability in high-mobility scenarios	Requires heavy computation and memory, unsuitable for resource-constrained vehicular nodes
Uppoor et al. [24]	2023	RL + LSTM digital twin	Ultra-dense networks	Improves prediction accuracy and adaptability in dense environments	Relies on large datasets for training; real-time deployment remains challenging

3. Methodology

The proposed methodology consists of three main components: (i) estimation of signal strength using entropy, (ii) traffic load modeling through Little's theorem, and (iii) vehicular mobility analysis based on speed variation, all integrated into a Q-learning framework with an ε-greedy action selection policy. While the mathematical formulations describe each part in detail, the overall workflow can be better understood through system-level diagrams and algorithmic flowcharts.

The system supports VANETs running under heterogeneous communication environments, where Li-Fi and Wi-Fi are employed. Building on and enhancing existing works [6,21,25,26], the communication model, simulation structure, and mobility metrics have new contributions in state parameter integration and algorithm development. The system is designed to operate in a VANET environment using fixed wireless access (FWA) technology. These vehicles are equipped with on-board units consisting of their core wireless transceivers, sensors, and GPS systems. These will thus be able to communicate with the base stations. The bases have been mounted with Active Antenna Systems and multiple input multiple output (MIMO) technology so that they can work efficiently with the vehicles both by Li-Fi and Wi-Fi. The architecture is summarized in **Figure 1**.

The handover management technique proposed in this paper enhances the Li-Fi-based VANET handover model presented in Kalita and Barooah [6] by using a real-time Q-learning decision-making algorithm. Unlike fixed threshold-dependent handovers or single parameter-based handovers, the proposed technique uses three dynamically changing parameters: signal strength (SS), network occupancy (NO), and vehicular mobility (VMo) to make the handover decision more adaptive. The three parameters: signal strength (SS), network occupancy (NO), and vehicular mobility

(VMo)—form the state space of the Q-learning system. Each parameter is discretized into levels (e.g., high, medium, low), and every combination of these levels represents a unique state. At each state, the action space consists of two choices: either initiate a handover to a new base station or remain connected to the current one. The system is guided by a reward function that reflects the quality of each decision. The reward function reflects real network events by assigning values that guide optimal decisions: successful handovers receive a positive reward (+1), failed handovers a negative reward (−1), and ping-pong handovers a stronger penalty (−2). Maintaining a stable connection with good QoS is given a smaller positive reward (+0.5), encouraging reliability while discouraging unnecessary switches, as detailed in **Figure 2**.

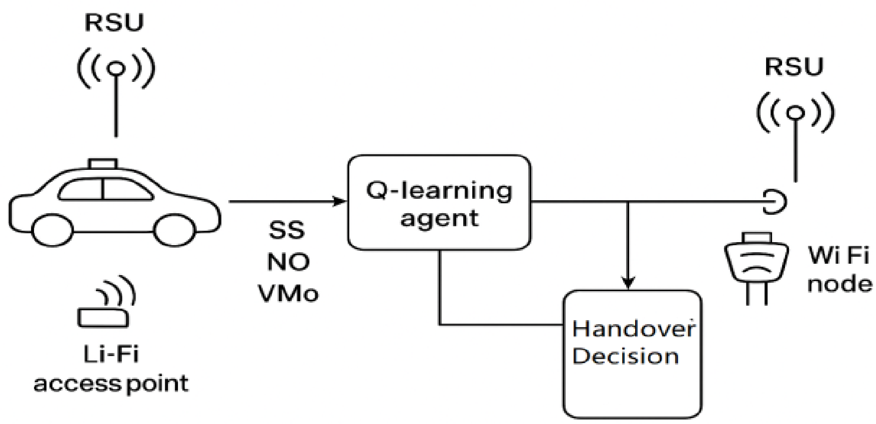

Figure 1. VANET system architecture.

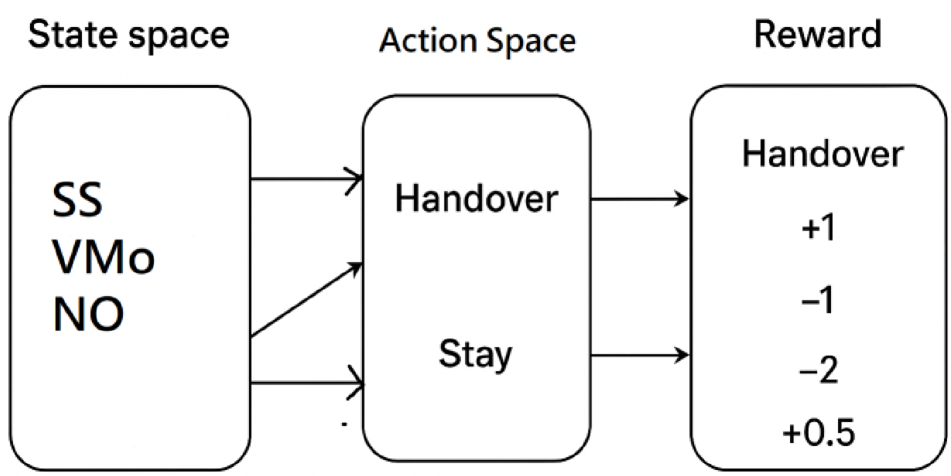

Figure 2. Diagram of state, action, and reward for the handover mechanism.

Three parameters considered in designing the Q learning handover system are explained using mathematical correlations as follows:

3.1. Signal Strength (SS)

The signal coverage area (SCA) is represented in our work by a circular region. Therefore, the SS is defined on the distance (di) between the 2 base stations, which is given as from Rappaport [27]:

$$di = 4\int_{d/2}^{x} \sqrt{r^2 + x^2}dx \quad (1)$$

$$ss = k \cdot 1/di \quad (2)$$

Where, k = no of network phases

$$S(ss) = [-log(P(ss))] \quad (3)$$

Where, $S(ss)$ is the Shannon entropy [25] for signal strength.

3.2. Network Occupancy (NO)

The network occupancy is monitored by using Traffic Load (TL). The traffic load is dependent on two factors: vehicle traffic (VT) in the network, defined by the average queue size ($qavg$) and the number of vehicles (m). The value of m is defined for the vehicles in its 1-hop distance.

$$TL = qavg \cdot m \quad (4)$$

For calculating the $qavg$, the average number of vehicles in a SCA (N_p) is dependent on the arrival rate of vehicles into the SCA (λ) and the average amount of time a vehicle spends in the SCA(T) given by Little [26], defined by Little's theorem.

$$N_p = \lambda T \quad (5)$$

If the leaving rate of vehicles from the SCA can be denoted by μ, N_p and T can be formulated as:

$$N_p = \frac{\lambda}{\mu - \lambda} \quad (6)$$

$$T = \frac{1}{\mu - \lambda} \quad (7)$$

If we consider T to include the queuing delay plus the service time Ts, the total time spent in the queue (Tt) can be calculated as:

$$Ts = \frac{1}{\mu} \quad (8)$$

$$Tt = T - 1/\mu \quad (9)$$

The $qavg$ can be obtained from Equations (5), (7), and (9) as:

$$\begin{aligned} qavg &= \lambda Tt \\ &= \frac{\lambda}{\mu - \lambda} - \frac{\lambda}{\mu} \\ &= N_p - \beta \end{aligned} \quad (10)$$

Here, β is the optimal traffic transfer ratio.

Equations (5)–(10) build step by step on traffic load modeling using Little's law. First, Equation (5) defines the average number of vehicles in the signal coverage area as the product of the arrival rate and the residence time. Equation (6) then breaks residence time into queueing delay and service time, with Equation (8) expressing service time as the reciprocal of the departure rate. Substituting this value yields Equation (9), which isolates the queueing delay. Finally, Equation (10) applies Little's law again to express average queue length as the total vehicles minus those served, linking directly to network occupancy. This progression clarifies how traffic load and queueing behavior are incorporated into the Q-learning state space.

3.3. Vehicular Mobility (VMo)

The vehicular mobility (VMo) of each node represents the movement of vehicles, and how it changes over time. In order to measure such VMo, the difference between the average speed (Avs) of the nodes in their final and initial locations is estimated in 't' time units. Dist is the distance between 1-hop vehicles. This can be presented as follows from Uppoor et al. [24]:

$$Avs = \frac{Dist}{t} \tag{11}$$

$$Vmo = Avs(final) - Avs(initial) \tag{12}$$

The Active network lifetime (ANL) can be obtained by the minimum value of the weight (Wt) associated with the vehicles in a SCA. This weight (Wt) parameter is given as:

$$Wt = w1 \cdot ss + w2 \cdot TL + w3 \cdot VMo \tag{13}$$

In Equation (13), $w1$, $w2$, $w3$ are represented as weight factors and the sum of these weight factors value is equal to 1 i.e., $w1 + w2 + w3 = 1$.

3.4. Q-Learning Framework

The parameters, NO and SS are the states that are taken in account to take an action for handover. Let (S,) represent state S and action A based on the Q values. Each state S will have four parameters and this (S, A) is determined and updated in the rule using Watkins [28].

$$Q(S,A) + \alpha(R + \gamma Q(S',A') - Q(S,A)) \rightarrow Q(S,A) \tag{14}$$

The term (S', A′) defines next state and action R is the reward given by the agent, γ is the discount factor and α is the learning rate in the range from 0 to 1, i.e., it denotes the step length to estimate the (S, A).

The action is taken using ϵ −greedy policy, the ϵ represents epsilon. In ϵ −greedy policy, when the probability is (1−ϵ), then the action will be taken as per the value in the Q-table. If the handover request is agreed and the action is yes, then it will select a network. The algorithm for the handover execution is depicted using the following flow chart, as shown in **Figure 3**.

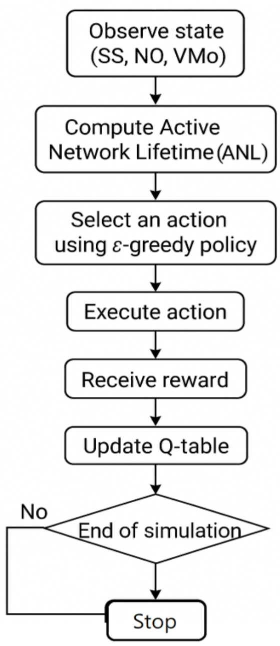

Figure 3. Flowchart for the handover process.

4. Results and Discussion

This chapter presents a comparative study of the performance parameters of our proposed Q-Learning-based handover technique with the DTe-DQN approach [23] for handling handovers in VANETs. For the simulation of realistic automotive environments, OMNeT++ at the network layer and SUMO for traffic mobility modelling, with the Veins framework, is used for the integration. Simulation was performed with different vehicle densities and fading scenarios. Li-Fi communication was simulated as an LOS-limited optical system, with DSRC serving as the baseline RF, and default propagation losses were simulated using the Two-Ray Ground and Log-normal models. The simulation parameters used for the proposed work are summarized in **Table 2**.

Table 2. Simulation parameters.

Category	Parameter	Value / Setting	Units
Environment setup	Number of vehicles	100	vehicles
	Maximum speed	120	km/h
	Communication types	DSRC, LTE, Li-Fi	–
	Simulation duration	300	s
	Fading levels	Low, Moderate, High	–
	Propagation model	Two-Ray Ground + Log-normal Shadowing	–
	Vehicle density	Low: 10, Medium: 50, High: 100	vehicles/km
	Radio communication range	Li-Fi: 50 (LOS), DSRC: 250 (NLOS)	m
	Simulation frameworks	OMNeT++ (network), SUMO (traffic)	–
Q-learning setup	Learning rate (α)	0.1	–
	Discount factor (γ)	0.9	–
	Exploration rate (ε)	1.0 →0.05 (decayed)	–
	State parameters	SS, NO, VMo (each discretized: Low/Med/High)	–
	Action space	{Handover, Stay connected}	–
	Reward values	+1 (success), −1 (failure), −2 (ping-pong), +0.5 (stable QoS)	–
Execution details	Training episodes	1000	episodes
	Simulation runs per scenario	5	runs

The Q-learning framework was configured with learning rate (α) = 0.1, discount factor (γ) = 0.9, and an ε-greedy exploration strategy where ϵ decays linearly from 1.0 to 0.05 during training. State variables: SS, NO, and VMo were discretized into three levels each (low, medium, high). Reward values were defined as +1 for successful handover, −1 for failure, −2 for ping-pong, and +0.5 for maintaining stable QoS. These values were selected based on prior reinforcement learning studies (cite) and tuned through preliminary experiments to ensure stable convergence.

To ensure the robustness of the reported findings, each simulation scenario was executed five times using different random seeds. The performance metrics presented in the figures represent the mean values across these runs. For instance, the observed throughput improvement showed a standard deviation of less than 2.5% of the mean, confirming that the reported performance gains are statistically reliable and not due to random variations. Where relevant, 95% confidence intervals were also computed to highlight the statistical reliability of the observed improvements. The results obtained are summarized in the below sub sections.

4.1. Probability of Handover Success

The probability of successful handover is defined as the ratio of successful handovers to the total number of attempted handovers. **Figure 4** summarizes the results, which show that the Q-Learning-based approach achieved a handover success probability of more than 90% in high-traffic scenarios. However, under similar conditions, DTe-DQN has a reported maximum success rate of 83.1%. The result of improving this system is due to the adoption of an adaptive learning mechanism based on Q-Learning, which is specifically tailored to make adaptations of handovers based on real-time network conditions.

4.2. Handover Failure Rate

The failure rate of handover refers to the number of handovers that fail per total number of handovers made in establishing a connection. The proposed method always keeps the failure rate lower than 10%. As shown in **Figure 5**, the system has an approximately 16% failure rate with DTe-DQN. The reason for this decrease in failure rate is because of Q-Learning enabled continuous learning process, which can account for the timely adjustment of handover strategies according to time-varying vehicular conditions.

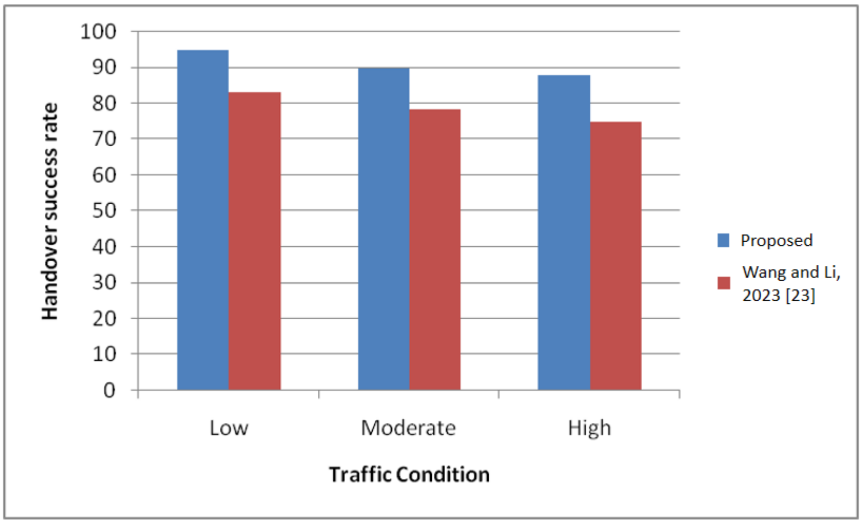

Figure 4. Handover success rate (%) comparison between proposed and Wang and Li [23].

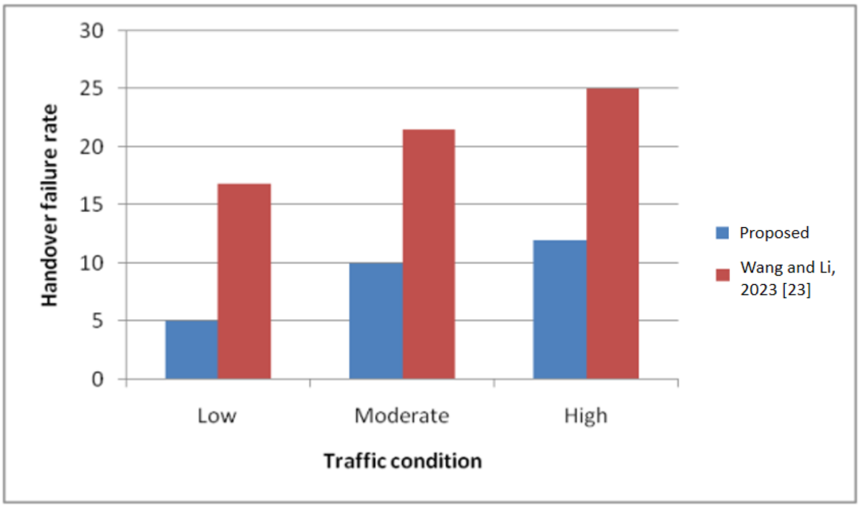

Figure 5. Handover failure rate comparison between proposed and Wang and Li [23].

4.3. Ping-Pong Rate

The ping-pong handovers occur when the vehicle rapidly switches between two base stations, and hence the network resources utilized are not fully efficient. The Q-Learning method achieves a ping-pong rate of less than 5%, as shown in **Figure 6.** This is far below the 8–9% that the DTe-DQN method achieves. Such a reduction, therefore, determines the efficiency of improving network performance since it minimizes unnecessary signaling overhead and enhances the user experience at the occurrence of events during handover.

4.4. Throughput and Delay

Throughput and delay are two important performance metrics of the network. The proposed method succeeded in making an average improvement in throughput by approximately 20% compared to traditional methods, as can be seen from **Figure 7.** From **Figure 8**, it can be seen that the average delay on handovers decreases by about 30%. In other words, the transition time between base stations was relatively shorter. The DTe-DQN method, despite being competitive, had higher delays due to its more conservative approach toward handover, especially at higher mobilities.

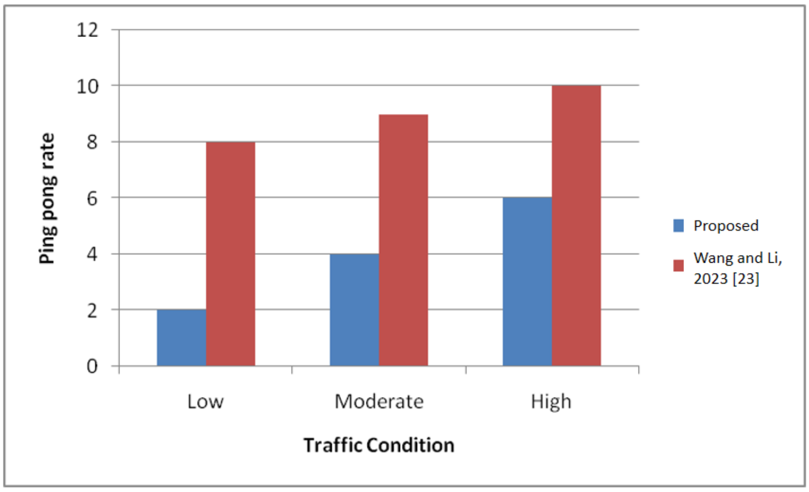

Figure 6. Ping pong rate comparison between proposed and Wang and Li [23].

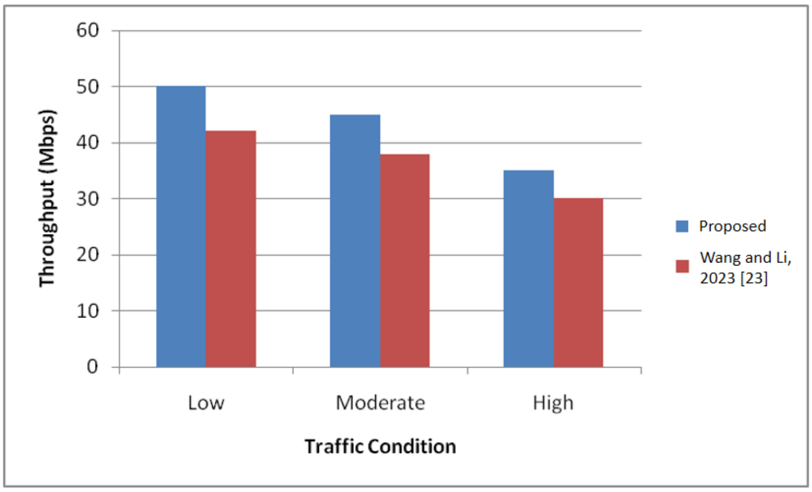

Figure 7. Throughput (Mbps) comparison between proposed and Wang and Li [23].

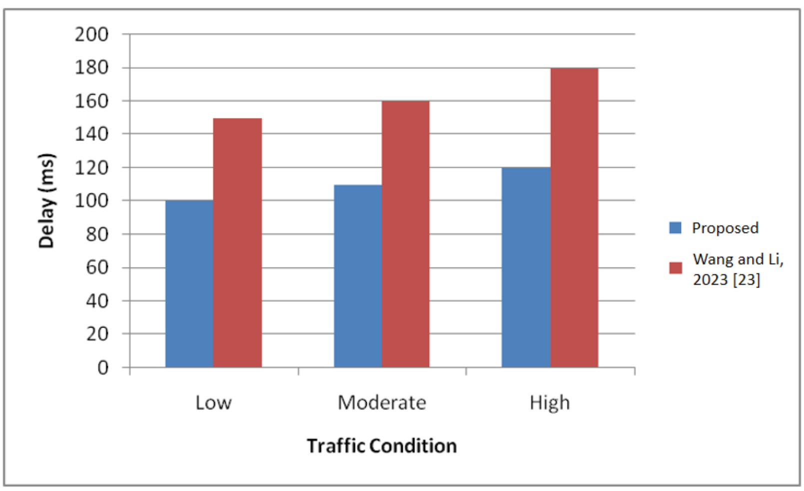

Figure 8. Delay (ms) comparison between proposed and Wang and Li [23].

4.5. Packet Delivery Ratio (PDR)

Figure 9 illustrates the PDR for various traffic conditions. The Q-learning model outperforms the DTe-DQN approach at all times, registering a PDR of more than 96% in low and medium traffic, and 93% in high-density traffic. The DTe-DQN model falls to less than 89% in high-traffic situations. This enhancement showcases the reliability of the proposed model in guaranteeing communication during handovers, even under dynamic scenarios.

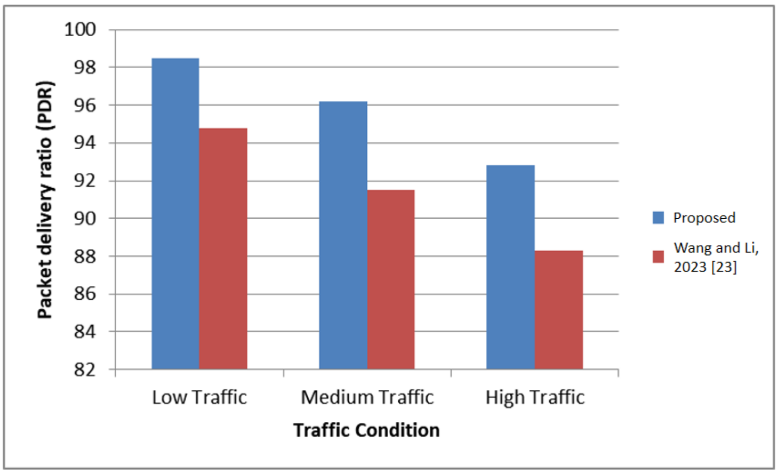

Figure 9. Packet delivery ratio (%) comparison between proposed and Wang and Li [23].

4.6. Average Handover Decision Time

Figure 10 indicates the average time taken to make a handover decision in different traffic conditions. The Q-learning model exhibits less decision latency, on average around 20% on average DTe-DQN. This is mainly because of the fast and light-weight adaptation properties of tabular Q-learning, as it estimates state-action values very quickly and reacts in real time. Smaller decision time is vital for minimizing disturbance in vehicular systems.

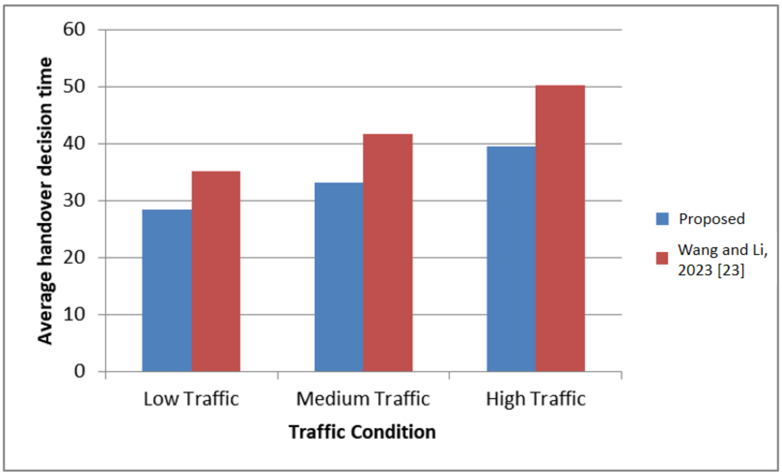

Figure 10. Average handover decision time (ms) comparison between proposed and Wang and Li [23].

Let E represent the number of episodes, T the number of time steps (iterations) per episode, and A the number of possible actions. In each iteration, the computationally dominant steps are selecting an action. Therefore, the per iteration complexity is approximately $O(A)$, and over the course of one episode with T iterations, the complexity becomes $O(T \times A)$. Extending this to E episodes, the total computational complexity of the algorithm is $O(E \times T \times A)$. In terms of space, the Q-table must store values for every possible state-action pair, giving a storage complexity of $O(|StateSpace| \times A)$.

4.7. Scalability Analysis

The proposed Q-learning-based handover algorithm possesses superior scalability for large-scale network topologies and denser vehicle populations. Scalability, for this study, is defined as the capacity of the algorithm to exert efficient and adaptive handover performance with increasing numbers of vehicles, communication nodes, and traffic variability. The algorithm was tested under varied vehicle densities (10, 50, and 100 vehicles/km) to mimic low, medium, and high traffic conditions under grid-like urban topologies as well as linear highway environments. The learning framework adjusts well because of the lightweight character of Q-learning, which only needs a finite state-action space and is independent of deep models or high-dimensional function approximates. To minimize computational complexity, state variables like signal strength, network occupancy, and mobility are discretized into easy-to-handle ranges to enable high-speed Q-table lookups and updates. The method ensures that convergence can be achieved within a realistic time period, even as the number of vehicles increases. Moreover, the modularity of the algorithm enables it to work independently on every vehicle node. The distributed learning process avoids centralized bottlenecks and enables concurrent system-wide handover decisions, thereby supporting increased network capacity without incurring processing overhead costs.

The results of the experimental evaluation of the proposed Q-Learning-based handover management strategy clearly indicate that, compared to existing approaches, particularly DTe-DQN, there are considerable improvements in these approaches. The analysis shows that the Q-Learning framework performs well in different traffic conditions, resulting in higher handover success rates and a greater probability of effectively delivering packets to their destinations. The Q-learning-based approach registered a 90% handover success rate for moderate traffic load, which is more efficient than the 83.1% success rate by the DTe-DQN method. The reason for this enhancement is mostly because of the real-time adaptability of the Q-learning algorithm to dynamic conditions of vehicular speeds and signal strengths. The ongoing update of handover decisions in real-time network situations males the method proposed here more accurate and timely, thereby improving the reliability and responsiveness of vehicular communication. As it would seem, throughput improvements are also noteworthy in the proposed approach since average throughput of 50 Mbps is achievable in low-traffic scenarios, while the DTe-DQN method reported its maximum throughput of 42 Mbps. For real-time applications such as video streams and navigation, this increase in throughput can make a significant difference for end users in terms of network performance quality, which in turn impacts their user experience. The strength of the method is also established using PDR and Average Handover Decision Time. The model holds a PDR over 95%, but DTe-DQN falls below 91% under traffic congestion. The average handover decision time is also minimized by around 20%, which is very important in order to keep latency to a minimum in high-speed vehicular scenarios.

5. Conclusions

This paper addressed key research gaps in VANET handover management by proposing an adaptive Li-Fi–assisted Q-learning framework. Unlike prior static threshold-based and blockchain-driven methods that lacked adaptability, the proposed model successfully integrated signal strength, traffic load, and vehicular mobility into the decision process. By doing so, it directly tackled the problems of poor adaptability, underutilization of Li-Fi, and neglect of performance metrics identified in the introduction. Simulation results confirmed that the approach achieved a handover success rate above 90%, a failure rate reduction of 6%, a ping pong reduction of 3%, a 20% throughput improvement, and a 30% delay reduction, thereby closing the loop between identified challenges and achieved outcomes.

For future work, research will extend into deep reinforcement learning architectures tailored for Li-Fi handover. In particular, exploring Dueling DQN and Actor–Critic frameworks could improve decision-making in high-dimensional continuous state spaces. Potential challenges include computational complexity on resource-constrained vehicular nodes and ensuring real-time responsiveness. Application scenarios such as dense urban traffic with frequent LOS–NLOS transitions will be targeted to evaluate scalability and practicality.

Author Contributions

Conceptualization, C.S.K. and M.B.; methodology, C.S.K.; software, C.S.K.; validation, C.S.K. and M.B.; formal analysis, C.S.K.; investigation, C.S.K.; resources, C.S.K.; data curation, C.S.K.; writing—original draft preparation, C.S.K.;

writing—review and editing, C.S.K. and M.B.; visualization,M.B. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

No public data was used for this research.

Conflicts of Interest

The authors declare that there is no conflict of interest.

References

1. Saini, M.; Mann, S. Handoff Schemes for Vehicular Ad-Hoc Networks: A Survey. In Proceedings of the 1st International Conference on Advancement in Electronics & Communication Engineering (ICAECE), Tebessa, Algeria, 15–16 May 2023.
2. Zeadally, S.; Hunt, R.; Chen, Y.; et al. Vehicular Ad Hoc Networks (VANETs): Status, Results, and Challenges. *Telecommun. Syst.* **2012**, *50*, 217–241.
3. Pack, S.; Choi, Y. Fast Handoff Scheme Based on Mobility Prediction in Public Wireless LAN Systems. *IET Proc. Commun.* **2004**, *151*, 489–495.
4. Chang, Y.T.; Ding, J.W.; Ke, C.H.; et al. A Survey of Handoff Schemes for Vehicular Ad-Hoc Networks. In Proceedings of the IWCMC' 10: 2010 International Wireless Communications and Mobile Computing Conference, Caen, France, 28 June–2 July 2010.
5. Wang, Y.; Haas, H. Dynamic Load Balancing with Handover in Hybrid Li-Fi and Wi-Fi Networks. *J. Lightwave Technol.* **2015**, *33*, 4671–4682.
6. Kalita, C.S.; Barooah, M. Li-Fi Based Handoff Technique in VANET. In Proceedings of the 2020 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2–4 July 2020.
7. Li, L.; Wen, D.; Yao, D. A Survey of Traffic Control With Vehicular Communications. *IEEE Trans. Intell. Transp. Syst.* **2013**, *15*, 425–432.
8. Duo, R.; Wu, C.; Yoshinaga, T.; et al. SDN-Based Handover Scheme in Cellular/IEEE 802.11p Hybrid Vehicular Networks. *Sensors* **2020**, *20*, 1082.
9. Dwivedi, S.; Amin, R.; Vollala, S.; et al. B-HAS: Blockchain-Assisted Efficient Handover Authentication and Secure Communication Protocol in VANETs. *IEEE Trans. Netw. Sci. Eng.* **2023**, *10*, 3491–3504.
10. Aboud, A.; Touati, H.; Hnich, B. Handover Optimization for VANET in 5G Networks. In Proceedings of the 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 9–12 January 2021.
11. Xie, Q.; Ding, Z.; Tang, W.; et al. Provable Secure and Lightweight Blockchain-Based V2I Handover Authentication and V2V Broadcast Protocol for VANETs. *IEEE Trans. Veh. Technol.* **2023**, *72*, 15200–15212.
12. Son, S.; Lee, J.; Park, Y.; et al. Design of Blockchain-Based Lightweight V2I Handover Authentication Protocol for VANET. *IEEE Trans. Netw. Sci. Eng.* **2022**, *9*, 1346–1358.
13. Alam, S.; Sulistyo, S.; Mustika, I.; et al. Review of Potential Methods for Handover Decision in V2V VANET. In Proceedings of the 2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), Jember, Indonesia, 16–17 October 2019.
14. Rosli, M.; Razak, S.; Yogarayan, S. 5G Handover Issues and Techniques for Vehicular Communications. *Indones. J. Electr. Eng. Comput. Sci.* **2023**, *32*, 1442–1450.

15. Costa, A.; Pacheco, L.; Rosário, D.; et al. Skipping-Based Handover Algorithm for Video Distribution Over Ultra-Dense VANET. *Comput. Netw.* **2020**, *176*, 107252.
16. Oladosu, G.; Tu, C.; Owolawi, P.; et al. Intelligent Metaheuristic-Based Handover Algorithm for Vehicular Ad Hoc Networks. *J. Commun.* **2023**, *18*, 589–598.
17. Anilkumar, S.; Rafeek, J. Soteria: A Blockchain Assisted Lightweight and Efficient Certificateless Handover Authentication Mechanism for VANET. In Proceedings of the 2023 3rd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), Ernakulam, India, 18–20 May 2023.
18. Liang, W.; Li, Z.; Zhang, H.; et al. Vehicular Ad Hoc Networks: Architectures, Research Issues, Methodologies, Challenges, and Trends. *Int. J. Distrib. Sens. Netw.* **2015**, *11*, 745303.
19. Zhang, Z.; Boukerche, A.; Pazzi, R. A Novel Multi-Hop Clustering Scheme for Vehicular Ad-Hoc Networks. In Proceedings of the MSWiM' 11: The 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Miami, FL, USA, 31 October–4 November 2011.
20. Kafle, V.P.; Kamioka, E.; Yamada, S. CoMoRoHo: Cooperative Mobile Router-Based Handover Scheme for Long-Vehicular Multihomed Networks. *IEICE Trans. Commun.* **2006**, *89*, 2774–2785.
21. Siriwardhana, A.C.P.K.; Yuan, J.; Shen, Z. Optimizing Handover Mechanism in Vehicular Networks Using Deep Learning and Optimization Techniques. *Comput. Netw.* **2025**, *270*, 111488. [CrossRef]
22. Shahwani, H.; Shah, S.A.; Ashraf, M.; et al. A Comprehensive Survey on Data Dissemination in Vehicular Ad Hoc Networks. *Veh. Commun.* **2022**, *34*, 100420.
23. Wang, H.; Li, B. Double-Deep Q-Learning-Based Handover Management in mmWave Heterogeneous Networks with Dual Connectivity. *Trans. Emerg. Telecommun. Technol.* **2023**, *35*, 1–15.
24. Uppoor, S.; Trullols-Cruces, O.; Fiore, M.; et al. Generation and Analysis of a Large-Scale Urban Vehicular Mobility Dataset. *IEEE Trans. Mob. Comput.* **2014**, *13*, 1061–1075.
25. Cincotta, P.; Giordano, C.; Silva, R.; et al. The Shannon Entropy: An Efficient Indicator of Dynamical Stability. *Phys. D Nonlinear Phenom.* **2020**, *417*, 132816.
26. Little, J.D.C. A Proof for the Queueing Formula: $L = \lambda W$. *Oper. Res.* **1961**, *9*, 383–387.
27. Rappaport, T.S. *Wireless Communications: Principles and Practice*, 2nd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2002.
28. Watkins, C.J.C.H. Learning From Delayed Rewards. PhD Thesis, University of Cambridge, King's College, Cambridge, UK, May 1989.