

A Survey of Reinforcement Learning Approaches for Traffic Signal Control

Shrinesh S. Rawool¹, Sanchita S. Patil¹, Siddhesh C. Kadam¹, Gauri S. Bolave¹, Ms P. C. Jasud¹, Mrs A. M. Kate¹, Mr A. P. Redekar²

¹Department of AIML

²Department of Electrical Engineering

²Assistant Professor

^{1,2}Dr Babuji Salunkhe Institute of Engineering and Technology (BSIET)
Kolhapur, India

Abstract—Traffic congestion is a major problem in urban transportation systems, leading to increased travel time, fuel consumption, and environmental pollution. The current traffic control system operates using an open-loop control strategy with a fixed operating time at road junctions. It operates in a sequence, stopping and allowing traffic from each lane one at a time based on predefined timing. However, it does not consider real-time traffic congestion at the junction and therefore fails to adapt to varying traffic conditions, such as dense or sparse traffic. To overcome these issues, various optimisation approaches have been developed that utilise real-time traffic data to control traffic signals and adapt to the current traffic scenario. One of the key approaches is using Reinforcement Learning (RL), which enables traffic control agents to learn optimal signal policies through interaction with the traffic environment. This review presents various reinforcement learning approaches used for traffic signal control, including Q-Learning, Deep Q-Network (DQN), and Proximal Policy Optimisation (PPO), while highlighting their advantages, limitations, and challenges in real-world traffic systems. The optimisation techniques, reinforcement learning methods, and simulation environments used for evaluation are considered in this comparative study. The outcomes of this review guide researchers working on intelligent traffic signal control.

Keywords—Traffic Signal Control, Reinforcement Learning, Deep Q-Network, Proximal Policy Optimisation, Intelligent Transportation Systems

I. INTRODUCTION

Rapid urbanisation is leading to an increasing number of vehicle owners. This means the number of vehicles on roads increases, resulting in traffic congestion problems and highlighting the need for efficient systems to handle traffic. The largest bottleneck for traffic is the road junction. Traditional systems at junctions can lead to increased commute time, fuel consumption, and pollution. The most important drawback of such systems is the delay in emergency services like ambulances, fire trucks, VIP convoys, etc. These vehicles can get stuck in traffic queues, which can lead to serious consequences. To minimise these problems, traffic needs to be managed properly, mainly at road junctions.

Traditionally, traffic signals are controlled either manually by traffic police or through traffic signal lights. These signal lights generally operate on fixed-time control systems, where predefined signal timings are set for each light and function without considering the real-time traffic density [12], [15], [16]. Actuated signal control systems use sensors to detect vehicles

and adjust signal timings accordingly. Various optimisation techniques, such as fuzzy logic and evolutionary algorithms, are also proposed to improve signal timing performance [17]. These approaches provide improvements over fixed signal systems, but they often struggle to adapt to rapidly changing traffic patterns.

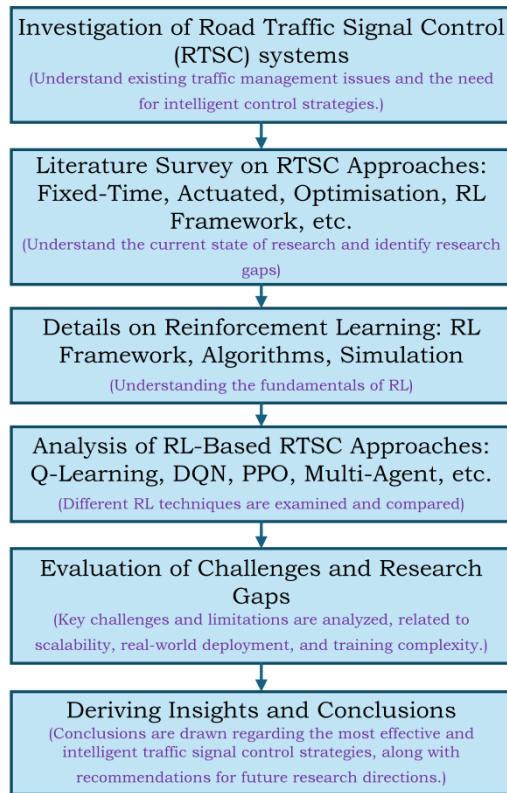
In recent years, ML techniques have been explored to improve traffic signal control. Among them, reinforcement learning (RL) has been given significant attention due to its ability to learn optimal control strategies through interaction with the traffic environment [1], [3]. RL enables an agent to observe the real-time traffic state, take actions such as changing signal phase, and receive feedback in the form of rewards based on traffic performance metrics (like average waiting times). Algorithms such as Q-Learning and Deep Reinforcement learning techniques like Deep Q-Network have shown promising results in optimising traffic signal control [2], [4], [8].

Researchers have applied RL in traffic signal control using simulation platforms to evaluate algorithm performance in different traffic scenarios [34], [35]. These methods are focused on reducing vehicle waiting times, minimising queue lengths, and improving overall traffic flow efficiency. Despite promising results, there are challenges regarding scalability, training stability, and real-world deployment.

This paper presents a survey of RL approaches used for traffic signal control [37], [38], [40]. The survey reviews classical traffic signal optimisation techniques, RL learning methods applied in traffic management, and key challenges and research trends. The objective of this study is to provide an overview of current research developments and highlight potential directions for future intelligent traffic signal control systems.

The remainder of this paper is organised as follows: Section 2 reviews traditional traffic signal control methods, including fixed-time, actuated, and optimisation-based approaches. Section 3 introduces the fundamentals of reinforcement learning and explains the key components of the RL framework and commonly used algorithms. Section 4 discusses various reinforcement learning approaches applied to traffic signal control, including classical and deep reinforcement learning methods, along with a comparison of existing studies. Section 5 highlights the major challenges and future research directions in applying reinforcement learning to real-world traffic signal systems. Finally, Section 6 concludes the paper with a summary of the findings and potential directions for future work.

Fig 1 illustrates the overall structure of the survey paper. It begins with the investigation of traffic signal control problems,



followed by a review of existing approaches. The study then focuses on reinforcement learning techniques, analysis of RL-based methods, and finally highlights research challenges and future directions.

Fig. 1. Survey Structure

II. LITERATURE REVIEW

Traffic signal control has traditionally been based on predefined time-based strategies. Early systems were designed using traffic engineering principles and considering the area's factors where the traffic signal is located. These methods aim to manage vehicle flow at intersections efficiently. Mostly, traffic signal systems are either fixed-time, actuated, or use some optimisation-based methods. Although they are effective in some scenarios, they face limitations under dynamic traffic conditions and unpredictable traffic conditions as they don't consider real-time data or manage the phases efficiently.

A. Fixed-Time Traffic Signal Control

Fixed-timer traffic signal control is one of the earliest and most widely used methods for managing traffic at intersections. In this, signal phases and their durations are predetermined based on historical traffic data and traffic engineering analysis [12], [15], [16]. Each signal phase operates for a fixed duration, and the cycle repeats continuously regardless of the current traffic demand. The main advantage of fixed-time systems is their simplicity and ease of implementation. As signal timings are predefined, the system does not require real-time traffic detection infrastructure. This makes fixed-time control suitable for intersections with mostly stable and predictable traffic patterns. However, fixed-time control systems are not capable of adapting to real-time traffic fluctuations. During periods of low

traffic demand, vehicles may experience unnecessary waiting times due to fixed signal duration. Similarly, during peak traffic periods, fixed timing plans may lead to long queues and increased congestion. As a result, fixed-time control methods often struggle to maintain optimal performance under dynamic traffic conditions.

B. Actuated Traffic Signal Control

Actuated traffic signal systems were introduced to improve upon the limitations of fixed-time control. These systems use vehicle detection devices like inductive loop detectors, cameras, or radar sensors to detect the presence of vehicles at intersections [16]. Based on the detected traffic demand, the signal controller can extend or terminate signal phases dynamically. In actuated systems, green signal duration can be extended if vehicles are detected approaching the intersection, reducing unnecessary waiting times. This allows the traffic signal system to respond more effectively to variations in traffic demand compared to fixed-time control. Regardless of these improvements, actuated traffic signal systems still rely on predefined rules and threshold parameters to determine signal changes. As a result, their adaptability is limited when dealing with complex traffic patterns or large-scale urban traffic networks. Additionally, the installation and maintenance of detection infrastructure may increase system costs.

C. Adaptive Traffic Signal Control Systems

Adaptive traffic signal control systems provide a more advanced approach to traffic management. These systems continuously monitor traffic conditions and adjust signal timings in real-time to improve traffic flow. Adaptive control systems often use centralised traffic management platforms that coordinate multiple intersections across the road network. Examples of such systems include widely implemented adaptive traffic signal control solutions that adjust signal phases based on real-time data collected from sensors and detectors [13], [14]. By dynamically adjusting signal timings, adaptive systems aim to minimise congestion and improve overall network performance.

Although adaptive traffic control systems provide significant improvements compared to traditional fixed-time or actuated methods, they still face several challenges. These systems often require extensive sensor infrastructure and sophisticated traffic management centres, which can increase implementation and maintenance costs. Furthermore, designing effective adaptive control strategies for complex traffic environments remains a challenging task.

D. Optimisation-Based Traffic Signal Control

In addition to rule-based control systems, researchers have explored various optimisation techniques to improve traffic signal performance. These approaches try to determine optimal signal timing plans by applying mathematical optimisation methods. Techniques such as fuzzy logic control, evolutionary algorithms, and differential evolution have been used to optimise traffic signal timings and improve intersection performance [17]. Optimisation-based approaches try to minimise key traffic metrics such as vehicle delay, queue length, and travel time. These methods often rely on mathematical models that describe traffic flow and intersection behaviour. By optimising these models, traffic signal timing plans can be improved.

However, optimisation-based methods also have limitations. Many optimisation techniques require accurate modelling of traffic behaviour, which can be difficult in real-world scenarios. Additionally, these methods may struggle to adapt quickly to rapidly changing traffic conditions. As a result, researchers have increasingly explored machine learning-based approaches for traffic signal control.

E. Section summary

Traditional traffic signal control methods have played an important role in managing traffic flow at urban intersections. Approaches such as fixed-time control, actuated control, and optimisation-based techniques have been widely used in transportation systems. However, these methods mostly rely on predefined rules or static models and may not perform well under highly dynamic traffic conditions. Consequently, recent research has focused on the use of intelligent and adaptive techniques such as reinforcement learning (RL) to develop more flexible and efficient traffic signal control systems.

Fig 2 presents the evolution of traffic signal control systems, starting from manual traffic control to fixed-time and actuated systems, followed by adaptive and optimisation-based approaches. Recent advancements focus on reinforcement learning and multi-agent systems, which enable intelligent and scalable traffic management in modern urban environments.

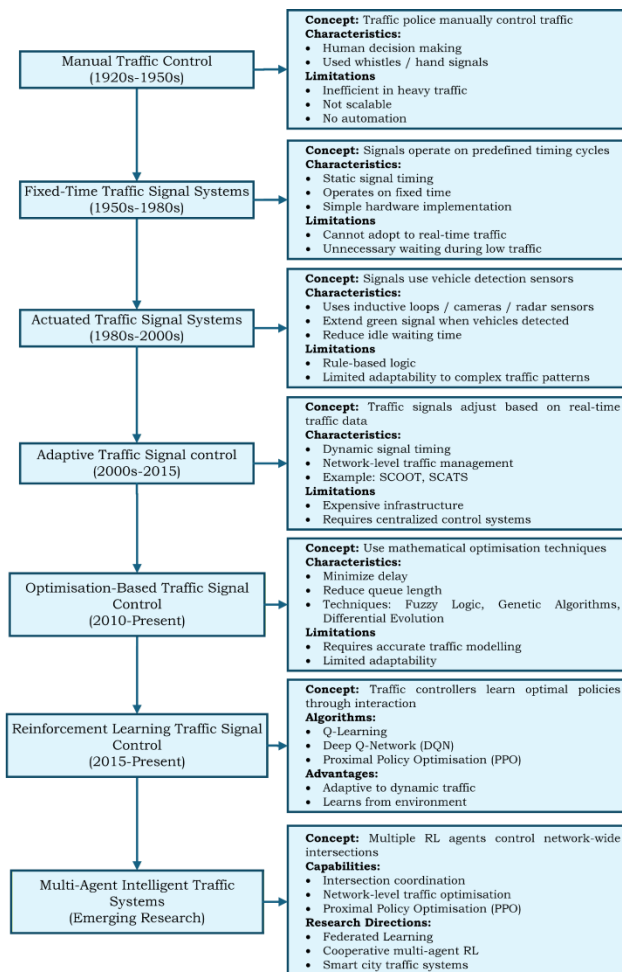


Fig. 2. Evolution of Traffic Signal Control

III. REINFORCEMENT LEARNING FUNDAMENTALS

Reinforcement learning (RL) is a branch of machine learning (ML) in which an agent learns to make decisions by interacting with its environment [1], [3]. Unlike supervised learning or unsupervised learning, which rely on datasets, RL focuses on learning optimal actions through trial-and-error interactions. The agent observes the current state of the environment, performs an action, and receives feedback in the form of rewards or penalties. Over time, the agent learns a policy that maximises the cumulative reward.

Reinforcement learning has been widely applied in various domains such as robotics, autonomous driving, and game playing [36], [39]. In the context of traffic signal control, reinforcement learning allows traffic signal controllers to learn optimal signal timing policies by observing traffic conditions and adjusting signal phases accordingly. By continuously interacting with the traffic environment, RL-based systems can adapt to dynamic traffic patterns and improve traffic flow efficiency.

A. Reinforcement Learning Framework

A reinforcement learning system typically consists of several key components, including the agent, environment, state, action, and reward. The agent represents the decision-making entity that interacts with the environment. In traffic signal control applications, the agent is usually the traffic signal controller responsible for selecting signal phases. Fig 3 shows the working of the reinforcement learning framework.

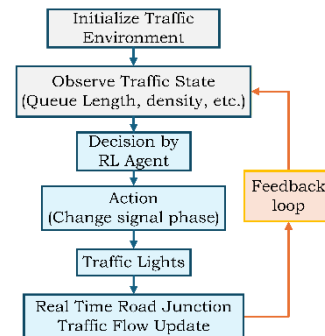


Fig. 3. Reinforcement Learning Framework for traffic control

The environment represents the system with which the agent interacts. In traffic signal control studies, the environment often consists of the traffic intersection and the surrounding road network. Traffic conditions such as vehicle queues, waiting times, and traffic flow patterns form the state of the environment that the agent observes. The state represents the current condition of the environment. In traffic signal control, the state may include variables such as queue length at each lane, vehicle waiting time, traffic density, or the number of vehicles approaching the intersection. Based on the observed state, the agent selects an action that influences the system. Actions correspond to the decisions made by the agent. In traffic signal control applications, actions typically involve selecting a signal phase or adjusting the duration of green lights. Once an action is performed, the environment transitions to a new state, and the agent receives a reward signal. The reward represents feedback provided to the agent after performing an action. In traffic signal control problems, reward functions are often designed to minimise traffic congestion by reducing vehicle waiting times,

queue lengths, or delays. By maximising cumulative rewards over time, the agent gradually learns optimal traffic signal control strategies.

B. Reinforcement Learning Algorithms

Several reinforcement learning algorithms have been proposed to solve decision-making problems. One of the earliest and most widely used algorithms is Q-Learning, which learns an action-value function that estimates the expected reward for each state-action pair [2]. Q-learning allows agents to learn optimal policies without requiring a model of the environment. With the advancement of deep learning techniques, researchers introduced deep reinforcement learning algorithms that combine neural networks with reinforcement learning. One popular method is the Deep Q-Network, which uses a neural network to approximate the Q-value function for large and complex state spaces [4], [8]. This approach enables reinforcement learning to be applied to more complex environments such as traffic networks. Another widely used algorithm is Proximal Policy Optimisation, which belongs to the family of policy gradient methods. PPO improves training stability by restricting large policy updates during learning [6]. Due to its stability and performance, PPO has been applied in several traffic signal control studies.

C. Reinforcement Learning in Traffic Simulation

Evaluating reinforcement learning algorithms in real-world traffic environments can be difficult due to safety concerns and infrastructure limitations. Therefore, researchers commonly use traffic simulation platforms to test and evaluate RL-based traffic signal control strategies. One of the most widely used simulation tools is SUMO, which provides a microscopic traffic simulation environment for modelling vehicle movement and traffic signal systems [34]. Traffic simulation platforms allow researchers to create realistic road networks, generate traffic flows, and evaluate different traffic control strategies under various conditions. Reinforcement learning agents can interact with the simulation environment, observe traffic states, and adjust signal timings while learning optimal policies through repeated simulations.

D. Section Summary

Reinforcement Learning provides a powerful framework for solving decision-making problems in dynamic environments. By allowing agents to learn from interactions with the environment, reinforcement learning can adapt to changing traffic conditions and optimise signal control strategies. Various RL algorithms such as Q-learning, deep Q-networks, and policy gradient methods have been applied to traffic signal control problems. These approaches have demonstrated promising results in simulation studies, motivating further research in intelligent traffic signal systems.

IV. REINFORCEMENT LEARNING APPROACHES FOR TRAFFIC SIGNAL CONTROL

Reinforcement Learning (RL) has emerged as a promising approach for adaptive traffic signal control due to its ability to learn optimal control strategies through interaction with dynamic traffic environments. Unlike traditional rule-based or fixed-time signal control methods, RL-based systems allow traffic controllers to adapt their policies based on real-time traffic conditions. Unlike traditional rule-based or fixed-time signal control methods, RL-based systems allow traffic controllers to adapt their policies based on real-time traffic

conditions. Various reinforcement learning techniques have been explored in the literature, ranging from classical Q-learning methods to more advanced deep reinforcement learning approaches. These methods differ in terms of state representation, reward functions, and learning architectures [37], [38], [39].

A. Early Reinforcement Learning Approaches

Early research on reinforcement learning for traffic signal control primarily focused on tabular learning algorithms such as Q-learning [2], [18]. In these approaches, the traffic signal controller learns optimal actions by updating a Q-table that maps traffic states to signal control actions. Typical state representations included traffic parameters such as vehicle queue lengths, traffic density, and waiting times at intersections. Actions generally involved selecting signal phases or adjusting the duration of green lights for specific directions. The reward function in these systems was often designed to minimise congestion-related metrics, including average waiting time, queue length, and vehicle delays. By continuously interacting with the traffic environment, the RL agent gradually learned signal policies that improved traffic flow efficiency. However, these early methods faced several limitations. The tabular representation of Q-values made it difficult to scale the system to large state spaces, particularly in complex traffic networks. As traffic environments became more dynamic and multi-dimensional, traditional RL methods struggled to handle the increased complexity.

B. Deep Reinforcement Learning Methods

To address the limitations of traditional RL approaches, researchers began incorporating deep learning techniques into reinforcement learning frameworks. Deep Reinforcement Learning (DRL) uses neural networks to approximate value functions or policies, enabling the system to handle high-dimensional traffic states. In DRL-based traffic signal control systems, the state of the traffic environment may include detailed information such as lane occupancy, vehicle counts, queue lengths, and vehicle waiting times. These inputs are processed by neural networks that estimate the expected reward for different signal control actions. One commonly used algorithm in this domain is the Deep Q-Network (DQN), which extends traditional Q-learning by replacing the Q-table with a neural network [4], [8], [9]. The network learns to predict Q-values for different actions based on observed traffic states. Several studies have demonstrated that DRL-based signal controllers can significantly improve traffic efficiency by reducing delays and improving traffic throughput [20], [21], [22]. These systems are often evaluated using traffic simulation platforms such as SUMO, which provide realistic traffic environments for training and testing reinforcement learning models. Despite these improvements, DRL-based systems also introduce challenges such as increased training complexity and the need for large amounts of simulation data.

Table I presents a comparative analysis of various reinforcement learning-based traffic signal control approaches. It can be observed that deep reinforcement learning methods outperform traditional approaches in handling complex traffic environments. Furthermore, multi-agent reinforcement learning techniques demonstrate improved scalability and coordination in large traffic networks, although they introduce additional challenges such as communication overhead and training complexity.

| Study | Method | Type | Environment | State Representation | Reward Function | Performance Metric | Key Contribution | Limitation |
|-------------------------|-----------------|--------------|-----------------|-------------------------------|------------------------------|-----------------------------|------------------------------------|-----------------------------------|
| Li et al. (2016) [20] | Deep RL | Single-Agent | Simulation | Queue length, traffic flow | Minimize delay | Average delay, queue length | Adaptive traffic signal control | Limited to simple intersections |
| Wei et al. (2018) [21] | DQN | Single-Agent | SUMO | Vehicle density, waiting time | Reduce waiting time | Waiting time, throughput | IntelliLight system | High training complexity |
| Van der Pol (2016) [22] | Deep RL | Single-Agent | Simulation | Traffic state representation | Maximise cumulative reward | Traffic efficiency | Policy learning for signal control | Limited scalability |
| Mannion (2016) [23] | Multi-Agent RL | Multi-Agent | Traffic Network | Multi-intersection states | Minimise network delay | Network throughput | Coordinated intersections | Complex training and coordination |
| Chu et al. (2020) [24] | Multi-Agent DRL | Multi-Agent | SUMO | Traffic density, queue length | Minimise travel time | Travel time, delay | Large-scale traffic control | Communication overhead |
| Wei et al. (2019) [25] | MARL (CoLight) | Multi-Agent | SUMO | Graph-based traffic states | Maximise coordination reward | Network performance | Network-level cooperation | High computational cost |

TABLE I. COMPARISON OF REINFORCEMENT LEARNING APPROACHES

C. Multi-Agent Reinforcement Learning Approaches

Recent research has extended reinforcement learning methods to multi-intersection traffic networks using Multi-Agent Reinforcement Learning (MARL). In these systems, each traffic intersection is controlled by an independent RL agent, allowing multiple intersections to learn and operate simultaneously [19], [24]. Multi-agent approaches enable coordination among neighbouring intersections, which is essential for managing traffic flow across large urban networks. By sharing information or learning cooperative strategies, agents can optimise traffic conditions beyond individual intersections [25]. These systems often focus on improving network-level performance metrics such as overall travel time, network throughput, and congestion levels. Collaborative learning among agents can lead to more efficient traffic management compared to isolated single-intersection control [26], [27]. However, multi-agent reinforcement learning systems also face challenges related to communication overhead, coordination complexity, and training stability. Ensuring consistent learning behaviour across multiple agents remains an active area of research in intelligent transportation systems.

V. CHALLENGES AND FUTURE SCOPE

Although reinforcement learning has shown promising results in traffic signal control, several challenges still limit its widespread deployment in real-world traffic systems. Current research has primarily focused on simulation-based evaluations, and many practical issues remain unresolved. This section discusses key challenges and potential future research directions for reinforcement learning-based traffic signal control.

A. Scalability in Large Traffic Networks

One major challenge in applying reinforcement learning to traffic signal control is scalability. Many studies focus on controlling a single intersection or a small number of intersections in simulated environments. However, real-world urban traffic networks consist of numerous interconnected intersections with highly dynamic traffic patterns.

As the size of the traffic network increases, the state space and action space also grow significantly. This makes it difficult for reinforcement learning models to learn efficient policies within a reasonable time. Future research may focus on developing scalable multi-agent reinforcement learning frameworks that can efficiently manage large-scale traffic networks [24], [25].

B. Real-World Deployment Challenges

Most reinforcement learning-based traffic signal control systems are evaluated using simulation environments. While simulation platforms provide controlled environments for testing algorithms, they may not accurately represent real-world traffic conditions. Real-world deployment introduces additional challenges such as sensor noise, unpredictable traffic behaviour, and infrastructure limitations. Integrating reinforcement learning models with existing traffic management systems also requires careful consideration of safety and reliability [37]. Future research should focus on bridging the gap between simulation-based studies and real-world traffic deployments.

C. Training Stability and Data Requirements

Deep reinforcement learning models typically require large amounts of training data and extensive computational resources. Training these models can be time-consuming, especially when

complex traffic environments are involved. Additionally, reinforcement learning algorithms may experience instability during training due to issues such as reward design, exploration strategies, and convergence difficulties. Designing efficient training methods and improving algorithm stability remains an important research direction [39].

D. Coordination in Multi-Agent Systems

In multi-intersection traffic networks, multiple RL agents must coordinate their decisions to optimise traffic flow across the network. Achieving effective coordination between agents is challenging due to communication constraints and the dynamic nature of traffic environments. Poor coordination among agents may lead to suboptimal traffic signal policies or increased congestion in certain areas. Future research may explore improved communication protocols, cooperative learning strategies, and decentralised control architectures for multi-agent traffic signal systems. Despite these challenges, reinforcement learning continues to be an active area of research in intelligent transportation systems. Advances in deep learning, multi-agent coordination, and real-time traffic sensing technologies are expected to further improve the effectiveness of RL-based traffic signal control systems in the future [26], [27].

VI. CONCLUSION

Traffic congestion remains a significant challenge in modern urban transportation systems, making efficient traffic signal control an important research area. Traditional traffic signal control methods, such as fixed-time and actuated systems, often struggle to adapt to dynamic traffic conditions, motivating the development of intelligent control strategies. This survey reviewed various reinforcement learning approaches applied to traffic signal control. Early studies primarily used classical reinforcement learning methods such as Q-learning to optimise signal timing based on traffic parameters, including queue length and vehicle waiting time. While these approaches improved performance compared to traditional control strategies, they faced limitations in large and complex traffic environments. Recent advancements in deep reinforcement learning have significantly enhanced the capabilities of intelligent traffic signal control systems by enabling models to process high-dimensional traffic data and learn effective control policies. In addition, multi-agent reinforcement learning approaches have been proposed to coordinate multiple intersections and improve network-wide traffic efficiency. Despite promising results, challenges such as scalability, training complexity, and real-world deployment remain. Overall, reinforcement learning presents a promising direction for developing adaptive and intelligent traffic signal control systems for future smart cities.

ACKNOWLEDGEMENT

The author would like to express sincere gratitude to the project supervisor for their continuous guidance, valuable suggestions, and encouragement during the preparation of this survey paper. The author also thanks Dr Bapuji Salunkhe Institute of Engineering and Technology for providing the necessary resources and academic support. Special thanks are extended to all the researchers whose valuable publications and studies contributed to the development of this survey on reinforcement learning approaches for traffic signal control.

REFERENCES

- [1] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, MIT Press, 1998.
- [2] C. J. C. H. Watkins and P. Dayan, "Q-learning," Machine Learning, vol. 8, no. 3, pp. 279–292, 1992.
- [3] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," J. Artif. Intell. Res., vol. 4, pp. 237–285, 1996
- [4] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, no. 7540, pp. 529–533, 2015.
- [5] V. Mnih et al., "Playing Atari with deep reinforcement learning," arXiv preprint arXiv:1312.5602, 2013.
- [6] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," arXiv preprint arXiv:1707.06347, 2017.
- [7] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in Proc. ICML, pp. 1889–1897, 2015.
- [8] V. Mnih et al., "Asynchronous methods for deep reinforcement learning," in Proc. ICML, pp. 1928–1937, 2016.
- [9] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in Proc. AAAI, pp. 2094–2100, 2016.
- [10] Z. Wang et al., "Dueling network architectures for deep reinforcement learning," in Proc. ICML, pp. 1995–2003, 2016.
- [11] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," arXiv preprint arXiv:1509.02971, 2015.
- [12] Traffic Signal Timing Manual, U.S. Department of Transportation, Federal Highway Administration, 2008.
- [13] P. B. Hunt, D. I. Robertson, R. D. Bretherton, and M. C. Royle, "The SCOOT on-line traffic signal optimisation technique," Traffic Engineering & Control, vol. 23, no. 4, pp. 190–192, 1982.
- [14] G. C. Lowrie, "SCATS: Sydney co-ordinated adaptive traffic system," Roads and Traffic Authority of NSW, 1990.
- [15] P. Koonce et al., "Traffic signal timing manual," U.S. Dept. of Transportation, FHWA, Tech. Report, 2008.
- [16] N. H. Gartner, C. J. Messer, and A. Rathi, Traffic Flow Theory: A State of the Art Report, Transportation Research Board, 2001.
- [17] A. Trabia, M. S. Kaseko, and M. Ande, "Traffic signal optimization using fuzzy logic," Transportation Research Part C, vol. 7, pp. 353–367, 1999.
- [18] B. Abdulhai, R. Pringle, and G. J. Karakoulas, "Reinforcement learning for true adaptive traffic signal control," J. Transp. Eng., vol. 129, no. 3, pp. 278–285, 2003.
- [19] M. A. Wiering, "Multi-agent reinforcement learning for traffic light control," in Proc. ICML, pp. 1151–1158, 2000.
- [20] L. Li, Y. Lv, and F.-Y. Wang, "Traffic signal timing via deep reinforcement learning," IEEE/CAA J. Autom. Sin., vol. 3, no. 3, pp. 247–254, 2016.
- [21] H. Wei, G. Zheng, H. Yao, and Z. Li, "IntelliLight: A reinforcement learning approach for intelligent traffic light control," in Proc. ACM SIGKDD, pp. 2496–2507, 2018.
- [22] E. Van der Pol and F. A. Oliehoek, "Coordinated deep reinforcement learners for traffic light control," in Proc. NIPS Workshop, 2016.
- [23] P. Mannion, J. Duggan, and E. Howley, "An experimental review of reinforcement learning algorithms for adaptive traffic signal control," in Autonomous Road Transport Support Systems, Springer, 2016.
- [24] T. Chu, J. Wang, L. Codecà, and Z. Li, "Multi-agent deep reinforcement learning for large-scale traffic signal control," IEEE Trans. Intell. Transp. Syst., vol. 21, no. 3, pp. 1086–1095, 2020.
- [25] H. Wei et al., "CoLight: Learning network-level cooperation for traffic signal control," in Proc. ACM CIKM, pp. 1913–1922, 2019.
- [26] G. Zheng et al., "Learning phase competition for traffic signal control," in Proc. ACM CIKM, pp. 1963–1972, 2019.
- [27] M. Aslani, M. S. Mesgari, and M. Wiering, "Adaptive traffic signal control with actor-critic methods in a real-world traffic network," Transportation Research Part C, vol. 85, pp. 663–684, 2017.
- [28] I. Arel, C. Liu, T. Urbanik, and A. G. Kohls, "Reinforcement learning-based multi-agent system for network traffic signal control," IET Intell. Transp. Syst., vol. 4, no. 2, pp. 128–135, 2010.

- [29] J. Gao, Y. Shen, J. Liu, M. Ito, and N. Shiratori, "Adaptive traffic signal control: Deep reinforcement learning algorithm with experience replay and target network," arXiv preprint arXiv:1705.02755, 2017.
- [30] D. Garg, M. Chli, and G. Vogiatzis, "Deep reinforcement learning for autonomous traffic light control," in Proc. IEEE ITSC, 2018.
- [31] H. Ge, Y. Song, C. Wu, J. Ren, and G. Tan, "Cooperative deep Q-learning with Q-value transfer for multi-intersection signal control," IEEE Access, vol. 7, pp. 40797–40809, 2019.
- [32] W. Liu et al., "Distributed cooperative reinforcement learning-based traffic signal control integrating V2X networks," IEEE Trans. Veh. Technol., vol. 66, no. 10, pp. 8667–8681, 2017.
- [33] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO – Simulation of Urban MObility," Int. J. Adv. Syst. Meas., vol. 5, no. 3&4, pp. 128–138, 2012.
- [34] G. Brockman et al., "OpenAI Gym," arXiv preprint arXiv:1606.01540, 2016.
- [35] H. Wei, G. Zheng, V. Gayah, and Z. Li, "A survey on traffic signal control methods," arXiv preprint arXiv:1904.08117, 2019.
- [36] L. Yau, J. Qadir, H. L. Khoo, M. H. Ling, and P. Komisarczuk, "A survey on reinforcement learning models and algorithms for traffic signal control," ACM Comput. Surv., vol. 50, no. 3, pp. 1–38, 2017.
- [37] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "A brief survey of deep reinforcement learning," IEEE Signal Process. Mag., vol. 34, no. 6, pp. 26–38, 2017.
- [38] [A. Noaen et al., "Reinforcement learning in urban network traffic signal control: A systematic literature review," Expert Syst. Appl., vol. 199, p. 116830, 2022.
- [39] D. C. Gazis and R. B. Potts, "The oversaturated intersection," in Proc. 2nd Int. Symp. Theory Road Traffic Flow, 1963.
- [40] M. Abdoos, N. Mozayani, and A. L. C. Bazzan, "Traffic light control in non-stationary environments based on multi-agent Q-learning," in Proc. IEEE ITSC, pp. 1580–1585, 2011.