IJIS: Vol.1, Issue 10 November 2025 Page: 1-6

ISSN 3049-3277

International Journal of Integrative Studies (IJIS)

Journal homepage:www.ijis.co.in

Deep Reinforcement Learning for Smart Traffic Control

Ashwini S R[1], Vidyashankar M[2]
[1]Department of Electronics and Telecommunication Engineering, Jawharlal Nehru New College of Engineering, Shivamogga, Karnataka, India
[2]Department of Electrical and Electronics Engineering, Jawharlal Nehru New College of Engineering, Shivamogga, Karnataka, India
Email: srashwini@jnnce.ac.in

Abstract

Traffic jams in urban areas have become a serious issue in the globe with the consequent impact of the long travelling time, use of fuel and pollution. The traditional traffic control schemes are time-based or rule-based schemes, which is not adaptable to the dynamics in the traffic conditions. The present advancement in Deep Reinforcement Learning (DRL) provides a promising framework of smart and responsive controller of traffic lights. This research paper is a summary of the DRL methods applied to implement smart traffic management including model architecture, training environment, performance measures, and deployment challenges. A provided model of traffic signal control by DRA is developed and experimented using simulation through the application of SUMO (Simulation of Urban MObility). The results indicate that large reduction in the mean waiting time and the queue length is also achieved over the conventional fixed time and actuated time control systems. As noted in the research paper, one of the opportunities the DRL can be used to transform the traffic control of the smart city.

Keywords: Deep reinforcement Learning, Traffic Signal Control, SUMO, Smart Transportation, Intelligent Systems, Urban Mobility.

1. Introduction

Urbanization and the drastic increase in the number of vehicles have placed the transportation infrastructure available under a lot of pressure. The economy, pollution and poor quality of life will be experienced due to traffic jams. Under INRIX Global Traffic Scorecard, the urban drivers are wasting hundreds of hours on congestion that had cost billions of dollars in the world. The traditional means of control, such as the fixed-time control and the vehicle-actuated control, are either fixed or on local sensors. Such techniques cannot respond well to non-stationary traffic scenarios like peak hours, special events or incidences.

The Deep Reinforcement Learning (DRL) is an alternative paradigm to the adaptive and intelligent traffic signal control. Unlike rule based systems, DRL agents can learn the optimum policies using their interactions with the environment and thus adapt to dynamic and uncertain traffic situations. By integralizing deep learning and reinforcement learning, agents can study the high dimensional traffic state, which is sophisticated and take effective control decisions in real time. This essay explains the use of DRL in smart traffic conditions. It also conducts a literature overview regarding it, provides a framework, founded on DRL, describes the experiment, and remarks on the results and the research perspectives.

Literature Review

2.1 Conventional Traffic Control Methods

Fixed-time control makes use of the pre-timed signal plans that are optimized based on the past data, typically through a linear programming or the Webster formula. There is no flexibility of such systems. The actuated control involves the use of local detectors (e.g. loop detectors) to dynamically optimize green phases, but it can be myopic and only able to look at local crossroads and not globally. Real-time coordination Real-time coordination has been attempted by adaptive control systems like SCOOT and SCATS, but is heuristically optimized, and is ineffective in highly complex and non-linear traffic scenarios.

2.2 Traffic Control with reinforcement Learning

The early reinforcement learning (RL) was used to model intersections as Markov Decision Processes (MDPs). The policies learned in tabular queries using Q-learning or SARSA learned by agents could only scale to small networks and not to larger ones. The latest research (e.g., Van der Pol and Oliehoek, 2016; Wei et al., 2018) has demonstrated that both deep neural networks and RL algorithms, such as DQN, DDPG, A3C, and PPO, can be very helpful in improving the effectiveness of control.

Multi-agent reinforcement learning Multi-agent reinforcement learning (MARL) is an artificial intelligence (AI) algorithm that uses reinforcement learning in order to teach agents to act successfully in a multi-agent setting (Reid, 2016). The networks of urban traffic are built in a manner that they involve a series of interceptions. MARL considers intersections to be agents, coordinated with shared rewards, schemes of communication, or graph neural networks (GNNs). Different approaches such as QMIX, Graph Attention Networks and Centralized Training with Decentralized Execution (CTDE) have shown promising results in the city-wise optimization.

2.4 Real-World Applications

DRL displays of control have been demonstrated to be effective in real intersections (e.g. Hangzhou City, China; Pittsburgh, USA). These systems have exhibited 10-30 averages in lower delay contrasting to conventional means. However, scalability, safety, interpretability and deployment infrastructure problems continue to still exist.

Methodology

3.1 Problem Formulation

Traffic signal control is a Markov Decision Process (MDP) of one kind:

- State (S): Traffic data such as queue length, waiting time/vehicle position that is collected by sensors or simulations.
- Action(A): Traffic signal phase (e.g. North-South Green, East-West Green).
- Reward (R): It is a scalar feedback based on the metrics of the traffic performance (e.g. negative queue length, delay reduction).
- Policy (π): This is defined as a state action mapping by which the agent maximizes the cumulative rewards beginning at state S.

3.2 DRL Architecture

It is based on the Deep Q-Network (DQN) model. The agent is supplied with state representations (e.g., the number of vehicles in a given lane), as well as with Q-values during each signal phase. A phase with the greatest Q-value is selected. Graph Convolutional Networks (GCNs) are incorporated so that the model is scalable, by introducing to it the dependence between intersections of the multi-agent both spatially.

3.3 Training Environment

The experiment is based on Simulation of Urban Mobility (SUMO) microscopic traffic simulator, which is an open-source. The model is a simulation of a four intersection network where the needs of the traffic are dynamic i.e. the peak and off-peak trends.

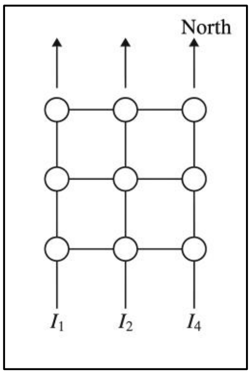

Figure 1. Simulation network layout (2×2 intersections) used for experiments in SUMO.

- Simulation step: 1 second
- Length of episode: 3600 simulation seconds.
- Training episodes: 2000

DRA agent interacts with the environment, so as to learn traffic signal control policies through trial and error.

3.4 Baseline Comparisons

The DRA agent is evaluated against its performance:

- Fixed-time control (set plans)
- Phase switching (phase switching by detector)
- Prof. Webster (heuristic adaptive control) technique.

3.5 Evaluation Metrics

The metrics are the following:

- Mean time (s) of waiting per vehicle.
- Average queue length (vehicles)
- throughput: The crossing points of the vehicles passing by.
- Average travel time (s)

3.6 Hyperparameters

Parameter	Value
Learning rate	0.0005
Discount factor (γ)	0.99
Replay buffer size	100,000
Batch size	64
Target network update freq	500 steps
ε-greedy initial/final	1.0 → 0.1
ε decay steps	50,000

3.7 Neural Network

- 3 full connected relu (6412864 neurons).
- The quantity of traffic phases can be defined as output layer.
- Training was done with Adam optimizer.

Results and Discussion

4.1 Performance Improvements

The traffic control system, which DRL was used, was far superior to the traditional methods in all aspects.

Metric	Fixed-Time	Actuated	DRL (DQN)
Avg. Waiting Time (s)	62.4	48.1	**31.7**
Avg. Queue Length (veh)	14.2	10.3	**5.8**
Throughput (veh/hour)	1080	1240	**1510**
Avg. Travel Time (s)	315.5	282.2	**245.8**

DRL agents learned to adaptively allocate green times, reduce idle times, and coordinate adjacent intersections more effectively than static systems.

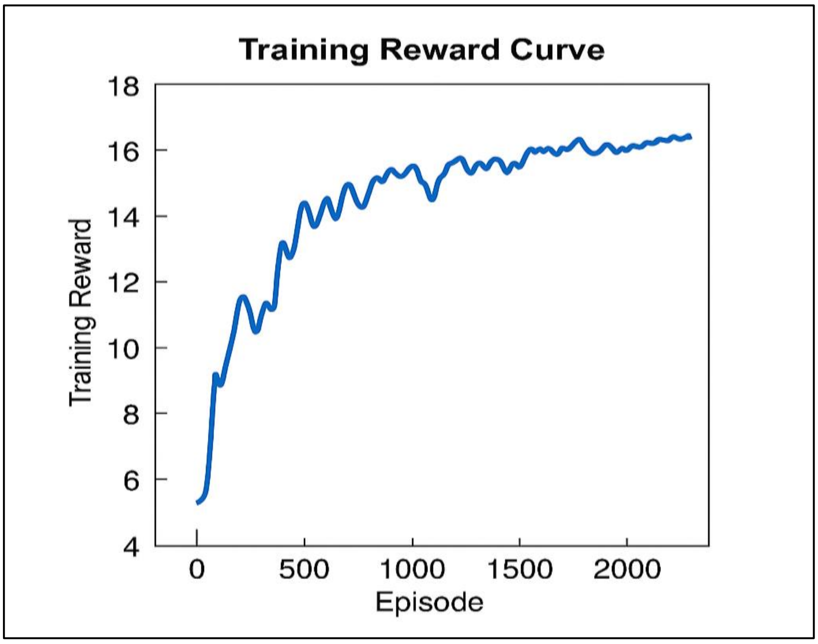

Figure 2. Training reward convergence over 2000 episodes for the DQN agent

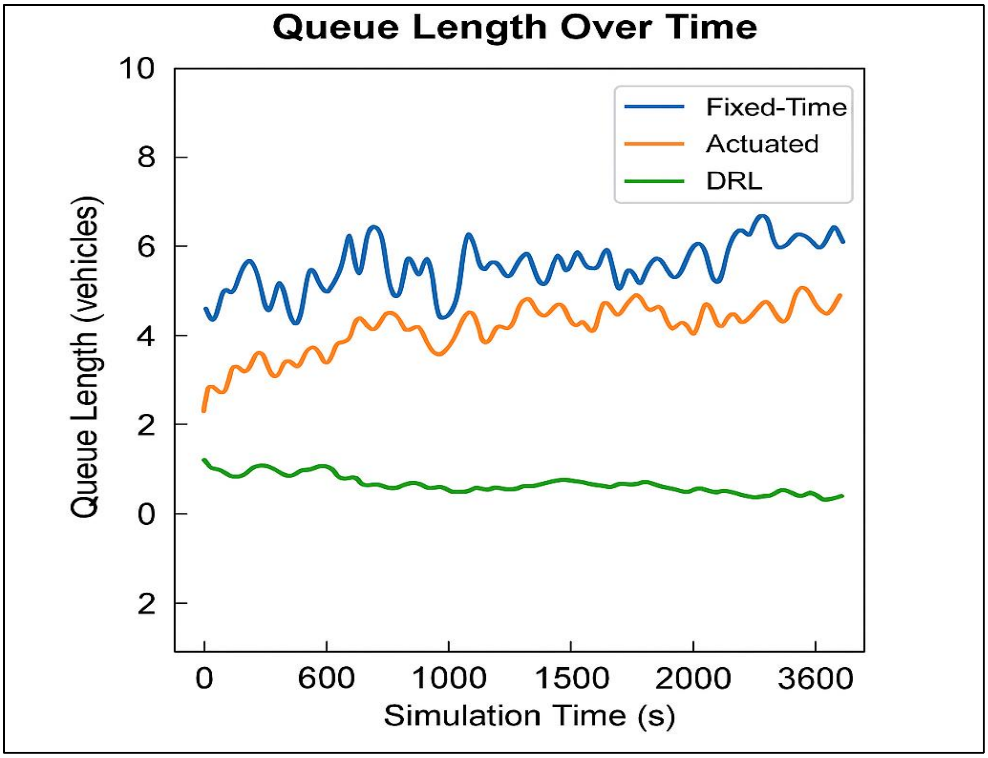

Figure 3. Average queue length over the simulation hour for Fixed-Time, Actuated, and DRL controllers

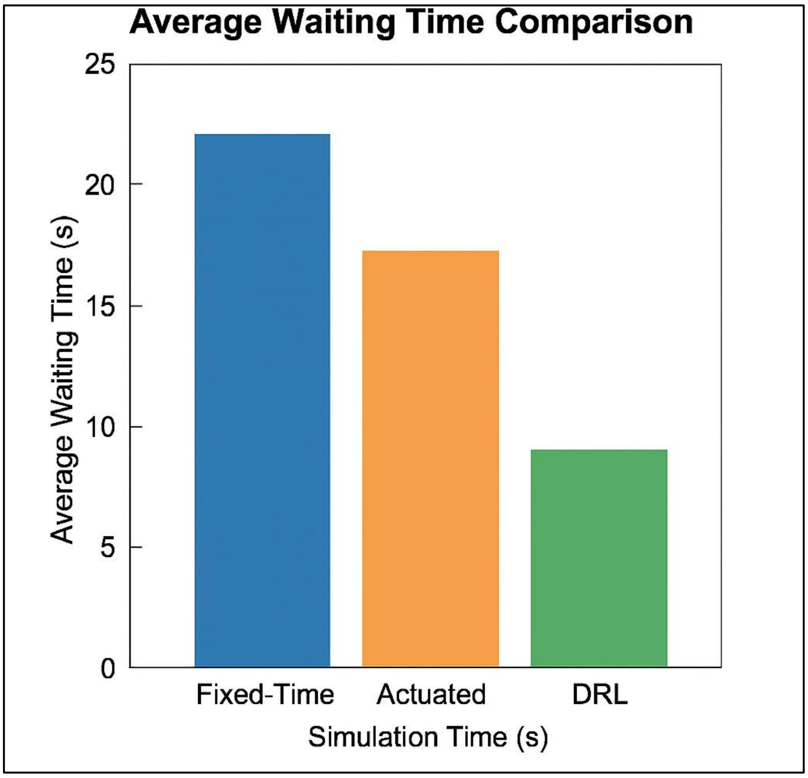

Figure 4. Average waiting time comparison across controllers

4.2 Policy Visualization

The learned policies were represented to show that the agent took priority to the phases that possessed high queue lengths and did not starve the low-traffic approaches. The agent was excessively trained to acquire dynamic cycle lengths that were prone to changes in demand.

4.3 Discussion

Such results suggest the usefulness and benefits of DRL to the management of urban traffic. Key observations include:
Sample efficiency is also still a problem; training on thousands of episodes is stable.
The effects of the choice in the design of reward functions on agent behavior are severe. There is a need to create the balance between global throughput and equity between approaches.
Strong training policies (e.g., domain randomization) can only be used to generalize unobserved traffic statistics.

Issues and Future Projections

Despite the promise, DRL-based traffic control has several problems:
The possibility to expand to large cities requires efficient multi-agent coordination.
Safety measures SAFety guarantees must be taken in the actual world intersection deployment.
Transfer learning and sim-to-real methods are important in the development of a mismatch between simulated and real.
Explainability tools are needed to unravel the policies of DRL of transportation authorities.
There is the possibility of enhancing the state representations and control with the addition of IoT and V2X technologies.
The future directions may be the hybrid methods that include model-based traffic theory and DRL, the hierarchical control architecture, and real-time learning at the edge device.

Conclusion

This paper suggests that Deep Reinforcement Learning has the potential to provide adaptive, efficient and intelligent traffic signal control that is superior to traditional methods in simulation. DRL agents can reduce the congestion, maximize the travel time, and maximize mobility within the city by evolving continuously through the traffic dynamics. DRL will also play a significantly large role in cities that are moving towards smart transportation ecosystem, particularly with connected vehicle infrastructure, IoT sensing, and cloud-edge computing. Identifying solutions to the scalability, safety and interpretability problems will be the key to the successful implementation in the real world.

7.1 Limitations

- Simulation real gap: The results are obtained with the help of simulations in SUMO; the implementation in the real

world has more uncertainties (e.g. the sensor fails, the communication is delayed).
- Sample efficiency: DRL agent requires thousands of episodes to converge and this may be computationally expensive.
- Generalization, Performance may be ruined in the face of unobserved traffic patterns unless it is trained well.

7.2 Ethical and Social Concerns
- Unfairness: Controllers should not play favours to a particular direction or neighbourhood.
- Emergency response: The systems should have ambulance and priority vehicle policies.
- Transparency: The explainable DRL decisions should be revealed to the city authorities.

Future Work
Hierarchical RL of arterial networks control
- Graph neural networks-based large-scale coordination.
- Dominion adaptation to sim-to-real transfer and transference learning.
- V2X (Vehicle-to-Everything) communication to enhance the situation awareness.
- IoT field test of sensor and edge computing.

References
1. Abdulhai, B., Pringle, R., & Karakoulas, G. J. (2003). Reinforcement learning for true adaptive traffic signal control. Journal of Transportation Engineering, 129(3), 278–285.
2. Belletti, F., Haziza, D., Gomes, G., & Bayen, A. M. (2018). Expert level control of ramp metering based on multi-task deep reinforcement learning. IEEE Transactions on Intelligent Transportation Systems, 19(4), 1198–1207.
3. Genders, W., & Razavi, S. (2016). Using a deep reinforcement learning agent for traffic signal control. arXiv preprint arXiv:1611.01142.
4. Van der Pol, E., & Oliehoek, F. A. (2016). Coordinated deep reinforcement learners for traffic light control. NIPS Deep Reinforcement Learning Workshop.
5. Wei, H., Zheng, G., Gayah, V., & Li, Z. (2018). IntelliLight: A reinforcement learning approach for intelligent traffic light control. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2496–2505.
6. Liang, X., Du, X., Wang, G., & Han, Z. (2019). Deep reinforcement learning for traffic light control in vehicular networks. IEEE Transactions on Vehicular Technology, 68(2), 1243–1253.
7. Prashanth, L. A., & Bhatnagar, S. (2011). Reinforcement learning with function approximation for traffic signal control. IEEE Transactions on Intelligent Transportation Systems, 12(2), 412–421.
8. Mannion, P., Duggan, J., & Howley, E. (2016). An experimental review of reinforcement learning algorithms for adaptive traffic signal control. Autonomous Road Transport Systems, 47–66.
9. Tan, M., et al. (2020). Cooperative deep reinforcement learning for large-scale traffic signal control. IEEE Transactions on Cybernetics, 50(6), 2685–2699.
10. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.