

# Adaptive Traffic Light Control Using Deep Reinforcement Learning

Dr. S. Pavan<sup>1</sup>, Shaik Jahera<sup>2</sup>, J Bhanu Prathap<sup>2</sup>, D Karthik<sup>2</sup>, K Manisha<sup>2</sup>

<sup>1</sup>Associate Professor, Department of ECE

<sup>2</sup>Students, Department of ECE

Tirumala Engineering Collge, Narasaraopet, India

**Abstract**—Traffic congestion is a major challenge in urban environments, leading to increased travel time, fuel consumption, and environmental pollution. Traditional fixed-time traffic signal systems are inefficient under dynamic traffic conditions. This paper proposes an adaptive traffic signal control system using reinforcement learning. The system dynamically adjusts signal timings based on real-time traffic density and waiting time. A Deep Q-Network (DQN) model is used to learn optimal traffic signal policies. Simulation results demonstrate improved traffic flow efficiency, reduced waiting time, and minimized congestion compared to conventional methods.

**Index Terms**—Traffic Signal Control, Reinforcement Learning, Deep Q-Network, Smart Cities, Traffic Optimization

## I. INTRODUCTION

Rapid urbanization and the continuous growth in vehicle population have led to severe traffic congestion in modern cities. As road infrastructure struggles to keep pace with increasing demand, intersections become critical bottlenecks where vehicles accumulate, resulting in delays, fuel wastage, and environmental pollution. Efficient traffic signal control plays a key role in managing these challenges and improving overall transportation efficiency.

Traditional traffic signal systems operate using fixed timing schedules that are pre-configured based on historical traffic patterns. While such systems are simple to implement, they fail to adapt to real-time traffic conditions. In many cases, vehicles are forced to wait unnecessarily at signals even when other lanes have little or no traffic. This results in inefficient utilization of road capacity and increased travel time.

To address these limitations, adaptive traffic signal control systems have been introduced. These systems adjust signal timings dynamically based on current traffic conditions using data collected from sensors, cameras, or simulation environments. Adaptive systems aim to reduce congestion by allocating green signal time more efficiently to heavily loaded lanes.

In recent years, Artificial Intelligence (AI) and Machine Learning (ML) techniques have gained significant attention in traffic management. Among these, Reinforcement Learning (RL) has emerged as a powerful approach for solving dynamic decision-making problems. RL enables a system to learn optimal control strategies through interaction with the environment, without requiring explicit programming of rules.

In traffic signal control, the intersection can be modeled as an environment, and the signal controller acts as an agent.

The agent observes the current traffic state, selects an action (signal phase), and receives a reward based on the effectiveness of that action. Over time, the agent learns to maximize cumulative rewards, leading to improved traffic flow and reduced congestion.

However, traditional RL methods such as Q-Learning face challenges when dealing with large and complex traffic environments due to the exponential growth of the state space. To overcome this limitation, Deep Reinforcement Learning (DRL) combines RL with deep neural networks, enabling the system to handle high-dimensional inputs and learn more efficient policies.

Deep Q-Network (DQN) is one of the most widely used DRL techniques for traffic signal control. It approximates the Q-value function using a neural network, allowing the system to generalize across different traffic conditions. This makes DQN suitable for real-world applications where traffic patterns are highly dynamic and unpredictable.

Inspired by these advancements, this paper proposes an adaptive traffic signal control system based on reinforcement learning. The system dynamically adjusts signal timings using real-time traffic density and waiting time information. A Deep Q-Network is employed to learn optimal signal control strategies that minimize congestion and improve traffic efficiency.

The main contributions of this work are as follows:

- Design of an adaptive traffic signal control system using reinforcement learning
- Development of a state representation based on vehicle count and waiting time
- Implementation of a Deep Q-Network for optimal signal decision-making
- Performance evaluation using simulation under different traffic scenarios

The remainder of this paper is organized as follows: Section II reviews related work, Section III presents the system model, Section IV describes the proposed methodology, Section V discusses simulation results, and Section VI concludes the paper.

## II. RELATED WORK

Traffic signal control has been widely studied over the years, with approaches ranging from traditional fixed-time methods to advanced artificial intelligence-based techniques. This

section reviews the major categories of traffic signal control systems and highlights their advantages and limitations.

#### A. Fixed-Time and Actuated Control Systems

Conventional traffic signal systems operate using fixed-time control, where signal timings are predefined based on historical traffic data. These systems are simple to implement and require minimal computational resources. However, they lack adaptability and perform poorly under dynamic traffic conditions. Vehicles often experience unnecessary delays when traffic distribution is uneven across different lanes.

To overcome this limitation, actuated traffic signal systems were introduced. These systems use sensors such as loop detectors to adjust signal timings based on vehicle presence. Although actuated systems improve performance compared to fixed-time control, they are still limited in handling complex traffic patterns and large-scale intersections.

#### B. Sensor-Based and IoT-Based Systems

With the advancement of sensing technologies, modern traffic systems utilize cameras, infrared sensors, and IoT devices to collect real-time traffic data. These systems enable better monitoring of vehicle density and traffic flow. Image processing techniques have been widely used to detect vehicle count and estimate congestion levels.

IoT-based traffic management systems allow communication between traffic signals and centralized control units, enabling coordinated signal control. While these approaches improve data collection and monitoring, they often rely on predefined rules and lack the ability to learn and adapt over time.

#### C. Machine Learning Approaches

Machine learning techniques have been applied to traffic signal control to improve adaptability. Supervised learning models can predict traffic patterns based on historical data, while unsupervised learning methods can identify traffic trends and anomalies.

However, these approaches require large datasets and may not perform well in highly dynamic environments where traffic conditions change frequently. Additionally, they typically do not provide real-time decision-making capabilities for signal control.

#### D. Reinforcement Learning-Based Methods

Reinforcement Learning (RL) has emerged as a powerful approach for traffic signal optimization. In RL-based systems, the traffic signal controller acts as an agent that learns optimal policies through interaction with the environment.

Q-Learning is one of the earliest RL techniques used for traffic control. It enables the agent to learn optimal actions based on rewards. However, Q-Learning suffers from scalability issues as the state space grows, making it unsuitable for complex traffic environments.

Deep Reinforcement Learning (DRL) extends RL by integrating neural networks to approximate value functions. Deep Q-Networks (DQN) have been successfully applied to traffic

signal control problems, allowing the system to handle high-dimensional state spaces and learn more efficient policies.

Recent studies have also explored multi-agent reinforcement learning, where multiple intersections are controlled by independent agents. These systems improve coordination across traffic networks and enhance scalability. As demonstrated in recent research, multi-agent DQN approaches significantly improve traffic throughput and reduce congestion in large urban networks.

#### E. Limitations of Existing Work

Despite the progress in intelligent traffic control systems, several challenges remain. Many existing approaches are computationally expensive and difficult to deploy in real-world environments. Additionally, most systems focus on either single intersections or simplified traffic scenarios, limiting their scalability.

There is a need for efficient, scalable, and adaptive traffic signal control systems that can operate in real-time and handle dynamic traffic conditions effectively. The proposed work addresses these challenges by implementing a reinforcement learning-based adaptive signal control system that improves traffic flow and reduces congestion.

### III. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, the traffic signal control problem is formulated as a Markov Decision Process (MDP). The intersection is modeled as an environment, and the traffic signal controller acts as an agent that learns optimal control policies.

#### A. Traffic Environment

The system considers a four-way intersection with vehicles arriving from multiple directions, namely East–West (EW) and North–South (NS). Each direction consists of incoming lanes where vehicles accumulate and wait for the green signal.

The traffic environment is dynamic in nature, where vehicle arrivals, queue lengths, and waiting times continuously change over time. The goal of the system is to optimize traffic signal timings to reduce congestion and improve flow efficiency.

#### B. MDP Formulation

The traffic signal control problem is defined as a Markov Decision Process (MDP) represented by the tuple:

$$(S, A, R, P, \gamma) \quad (1)$$

where:

- $S$  represents the set of states
- $A$  represents the set of actions
- $R$  is the reward function
- $P$  is the state transition probability
- $\gamma$  is the discount factor

At each time step  $t$ , the agent observes the current state  $S_t$ , selects an action  $A_t$ , and receives a reward  $R_t$ .

### C. State Representation

The state represents the current traffic condition at the intersection. It is defined using key traffic parameters:

$$S_t = \{N_{EW}, N_{NS}, W_{EW}, W_{NS}\} \quad (2)$$

where:

- $N_{EW}$  = Number of vehicles in East–West direction
- $N_{NS}$  = Number of vehicles in North–South direction
- $W_{EW}$  = Waiting time in East–West direction
- $W_{NS}$  = Waiting time in North–South direction

This state representation captures both congestion level and delay, enabling the agent to make informed decisions.

### D. Action Space

The action space defines the possible traffic signal control decisions. In this system, the agent can choose between:

- $A_1$ : Green signal for East–West direction
- $A_2$ : Green signal for North–South direction

At each time step, the agent selects one action based on the observed state.

### E. Reward Function

The reward function guides the learning process by encouraging actions that reduce congestion. It is defined as:

$$R_t = -(\alpha W_t + \beta Q_t) \quad (3)$$

where:

- $W_t$  = Average waiting time
- $Q_t$  = Queue length
- $\alpha, \beta$  = weighting factors

A lower waiting time and queue length result in a higher reward, encouraging efficient traffic signal control.

### F. Objective Function

The objective of the system is to maximize the cumulative reward over time:

$$\max \sum_{t=0}^T \gamma^t R_t \quad (4)$$

This ensures that the system learns a policy that minimizes congestion and improves traffic flow.

### G. Problem Statement

The problem can be summarized as designing an intelligent traffic signal controller that dynamically adjusts signal timings based on real-time traffic conditions. The controller should minimize vehicle waiting time, reduce queue length, and improve overall traffic efficiency.

## IV. PROPOSED METHODOLOGY

This section presents the proposed adaptive traffic signal control system based on reinforcement learning. The system dynamically adjusts signal timings by learning optimal control policies through interaction with the traffic environment.

### A. Overall Framework

The proposed system follows a reinforcement learning framework where the traffic signal controller acts as an agent interacting with the traffic environment. The workflow consists of state observation, action selection, reward evaluation, and policy update.

At each time step, the agent observes the current traffic state, selects an action corresponding to a signal phase, and receives a reward based on traffic performance. This process is repeated continuously, allowing the system to learn optimal traffic signal strategies over time.

### B. Traffic Simulation Environment

A traffic simulation environment is used to model real-world traffic conditions. The system considers a four-way intersection with vehicles arriving from East–West and North–South directions.

Traffic parameters such as vehicle count, queue length, and waiting time are continuously updated. The simulation environment provides real-time feedback to the learning agent, enabling safe and efficient training without affecting real traffic systems.

### C. State Representation

The state representation is a crucial component of the reinforcement learning model. It captures the current traffic conditions at the intersection.

The state is defined as:

$$S_t = \{N_{EW}, N_{NS}, W_{EW}, W_{NS}\} \quad (5)$$

where:

- $N_{EW}$  and  $N_{NS}$  represent vehicle counts
- $W_{EW}$  and  $W_{NS}$  represent waiting times

This representation provides sufficient information for the agent to identify congestion levels and make effective decisions.

### D. Action Selection Strategy

The action space consists of possible signal phases:

- Green signal for East–West direction
- Green signal for North–South direction

An epsilon-greedy strategy is used to balance exploration and exploitation:

- With probability  $\epsilon$ , a random action is selected (exploration)
- With probability  $1 - \epsilon$ , the best action is selected (exploitation)

This ensures that the agent explores different strategies while gradually converging to optimal policies.

### E. Reward Function Design

The reward function is designed to minimize traffic congestion and waiting time:

$$R_t = -(W_t + \lambda Q_t) \quad (6)$$

where:

- $W_t$  is the average waiting time
- $Q_t$  is the queue length
- $\lambda$  is a weighting factor

A lower waiting time results in a higher reward, guiding the agent toward efficient traffic signal control.

### F. Deep Q-Network (DQN)

To handle large state spaces, a Deep Q-Network (DQN) is used. The DQN approximates the Q-value function using a neural network:

$$Q(s, a; \theta) \quad (7)$$

where  $\theta$  represents network parameters.

The network takes the traffic state as input and outputs Q-values for each action. The optimal action is selected as:

$$a^* = \arg \max_a Q(s, a) \quad (8)$$

### G. Experience Replay

To improve learning stability, experience replay is used. The agent stores past experiences in a replay buffer:

$$(s_t, a_t, r_t, s_{t+1}) \quad (9)$$

Mini-batches of experiences are randomly sampled during training, reducing correlation between consecutive samples and improving convergence.

### H. Target Network

A separate target network is used to stabilize training. The target network parameters are periodically updated from the main network.

The loss function used for training is:

$$L = \mathbb{E}[(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta))^2] \quad (10)$$

where  $\theta^-$  represents target network parameters.

### I. Training Process

The training process involves multiple episodes of interaction with the environment:

- 1) Initialize Q-network and target network
- 2) Observe initial state
- 3) Select action using epsilon-greedy policy
- 4) Execute action in environment
- 5) Receive reward and next state
- 6) Store experience in replay buffer
- 7) Sample mini-batch and update network
- 8) Update target network periodically

Over time, the model learns to optimize traffic signal timings and reduce congestion.

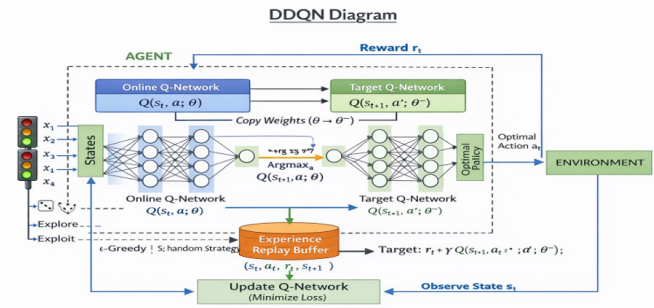


Fig. 1. DDQN Architecture

### J. System Workflow

The overall workflow of the proposed system can be summarized as follows:

- 1) Traffic data is collected from the environment
- 2) State representation is generated
- 3) Reinforcement learning agent selects action
- 4) Traffic signal phase is updated
- 5) New traffic conditions are observed
- 6) Reward is calculated
- 7) Model parameters are updated

This continuous feedback loop enables adaptive and intelligent traffic signal control.

## V. RESULTS AND DISCUSSION

This section evaluates the performance of the proposed reinforcement learning-based traffic signal control system under different traffic scenarios. The system is compared with a conventional fixed-time traffic signal system to demonstrate its effectiveness.

### A. Experimental Setup

The proposed model is implemented using Python and evaluated in a simulated traffic environment. The simulation considers a four-way intersection with vehicles arriving from East–West and North–South directions.

The reinforcement learning agent is trained over multiple episodes, where each episode represents a complete simulation cycle. During training, the agent interacts with the environment and learns optimal signal control policies based on reward feedback.

### B. Performance Metrics

The system performance is evaluated using the following metrics:

- **Average Waiting Time:** Measures the average delay experienced by vehicles at the intersection.
- **Queue Length:** Indicates the number of vehicles waiting at the signal.
- **Traffic Throughput:** Represents the number of vehicles successfully passing through the intersection.
- **Signal Efficiency:** Evaluates how effectively the signal timing adapts to traffic conditions.

### C. Training Performance

During the initial training phase, the agent explores different actions, resulting in fluctuating performance. As training progresses, the agent gradually learns optimal policies, leading to improved traffic flow and reduced congestion.

The convergence of the model indicates that the Deep Q-Network successfully learns to balance traffic load across different directions.



Fig. 2. Average waiting time over episodes

### D. Comparison with Fixed-Time System

The proposed adaptive system is compared with a traditional fixed-time traffic signal system.

In the fixed-time system, signal durations remain constant regardless of traffic conditions. This often leads to inefficient traffic flow, especially under uneven traffic distribution.

In contrast, the proposed system dynamically adjusts signal timings based on real-time traffic density. This results in better allocation of green signal time to congested lanes.

The results show that the adaptive system significantly reduces average waiting time and queue length compared to the fixed-time approach.

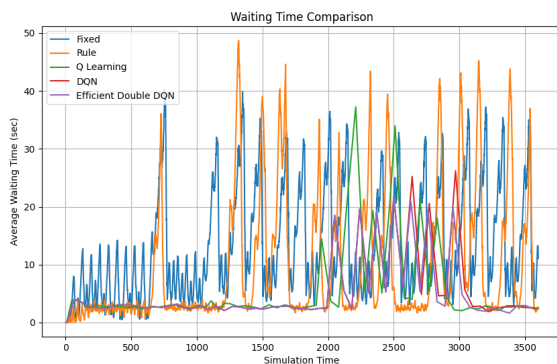


Fig. 3. Waiting comparison

### E. Performance Under Different Traffic Conditions

The system is evaluated under three different traffic scenarios:

- Low traffic density
- Moderate traffic density
- High traffic density

In low traffic conditions, both systems perform similarly, as congestion is minimal. However, in moderate and high traffic conditions, the proposed system demonstrates superior performance by efficiently managing traffic flow.

The adaptive system prioritizes heavily loaded lanes, reducing congestion and preventing queue buildup.

### F. Discussion

The results indicate that reinforcement learning provides an effective solution for traffic signal optimization. The DQN-based model successfully learns traffic patterns and adapts signal timings accordingly.

One of the key advantages of the proposed system is its ability to respond to dynamic traffic conditions without requiring predefined rules. This makes it suitable for real-world deployment in smart city environments.

However, the performance of the system depends on the quality of the reward function and training process. Proper tuning of parameters is necessary to achieve optimal results.

Overall, the proposed system demonstrates significant improvements in traffic efficiency, reduced delays, and better utilization of road infrastructure.

## VI. ADVANTAGES

- Adaptive signal control
- Reduced congestion
- Improved efficiency
- Scalable system

## VII. LIMITATIONS

- Requires training data
- Computational complexity
- Simulation vs real-world gap

## VIII. FUTURE SCOPE

- Multi-intersection control
- Integration with IoT sensors
- Real-time deployment

## IX. CONCLUSION

This paper presents a reinforcement learning-based adaptive traffic signal control system. The proposed approach improves traffic efficiency and reduces congestion. The system demonstrates strong potential for smart city applications.

## ACKNOWLEDGMENT

The authors thank the faculty and institution for their support.

REFERENCES

- [1] E. Ei Mon, H. Ochiai and C. Aswakul, "Application of Traffic Light Control in Oversaturated Urban Network Using Multi-Agent Deep Reinforcement Learning," *IEEE Access*, vol. 12, pp. 82384–82395, 2024, doi: 10.1109/ACCESS.2024.3397495.
- [2] H. Huang, Z. Hu, Z. Lu and X. Wen, "Network-Scale Traffic Signal Control via Multiagent Reinforcement Learning With Deep Spatiotemporal Attentive Network," *IEEE Transactions on Cybernetics*, vol. 53, no. 1, pp. 262–274, Jan. 2023, doi: 10.1109/TCYB.2021.3087228.
- [3] M. A. Hassan, M. Elhadeif and M. U. G. Khan, "Collaborative Traffic Signal Automation Using Deep Q-Learning," *IEEE Access*, vol. 11, pp. 136015–136032, 2023, doi: 10.1109/ACCESS.2023.3331317.
- [4] J. Wei, Y. Chen, Y. Lin and C. Xu, "Deep Reinforcement Learning for the Co-Optimization of Vehicular Flow Direction Design and Signal Control Policy for a Road Network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 14430–14442, Sept. 2024, doi: 10.1109/TITS.2021.3124567.