AIChE Journal

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Multi-Agent Reinforcement Learning for Intelligent Urban Traffic Signal Optimization: A Simulation-Based Study

Alex Carter[1] | Emily Johnson[2] | Daniel Brooks[1] | Sophia Turner[3]

[1]Department of Computer Science, Metropolitan Institute of Technology, California, USA | [2]School of Artificial Intelligence, Northern Research University, Ontario, Canada | [3]Department of Intelligent Transportation Systems, European Center for Smart Mobility, Berlin, Germany |

Correspondence: Alex Carter (alex.carter@example.com)

Received: 20 January 2023 | **Revised:** 21 January 2023 | **Accepted:** 22 January 2023

Academic Editor: Anonymous

Funding: No external funding was received for this portfolio demonstration manuscript.

Keywords: Artificial Intelligence | Reinforcement Learning | Multi-Agent Systems | Traffic Signal Control | Smart Cities | Transportation Optimization

ABSTRACT

Urban traffic congestion has become one of the primary challenges facing modern smart cities. Traditional traffic signal control methods often rely on static timing plans that cannot effectively respond to rapidly changing traffic conditions. Recent advances in artificial intelligence have enabled adaptive traffic management systems capable of continuously learning from real-time traffic observations.

This paper presents a simulation-based framework for intelligent traffic signal optimization using multi-agent reinforcement learning. Each intersection is modeled as an autonomous learning agent that cooperates with neighboring intersections while independently optimizing local traffic flow. The proposed framework combines decentralized decision making with shared environmental awareness to improve overall network efficiency.

Experimental simulations demonstrate improvements in average vehicle delay, traffic throughput, queue length, and intersection utilization compared with conventional fixed-time and adaptive baseline controllers. The study also analyzes convergence behavior, computational requirements, scalability, and communication overhead under varying traffic densities.

Although the presented study is entirely fictional and intended solely for portfolio demonstration purposes, it illustrates the preparation of a professional Wiley journal article containing figures, tables, algorithms, mathematical equations, appendices, and cross-referenced sections.

1 | Introduction

Rapid urbanization has significantly increased transportation demand throughout the world. As metropolitan populations continue to expand, existing road infrastructure experiences growing pressure from increasing vehicle density, unpredictable travel patterns, and frequent traffic congestion. Traffic delays not only increase travel time but also lead to higher fuel consumption, environmental pollution, economic losses, and reduced quality of life for commuters.

Conventional traffic signal control systems generally operate using fixed timing schedules developed from historical traffic surveys. Although such approaches are relatively simple to deploy and maintain, they often fail to adapt to continuously changing traffic conditions. Rush-hour congestion, special events, road maintenance, weather conditions, and unexpected incidents may all produce traffic patterns that differ considerably from those assumed during signal timing design. Consequently, static traffic controllers frequently create unnecessary vehicle queues, inefficient green time allocation, and poor intersection utilization.

Recent developments in artificial intelligence have transformed the design of intelligent transportation systems. Machine learning techniques enable traffic controllers to continuously learn from observed traffic conditions instead of relying exclusively on predefined schedules. Among these approaches, reinforcement learning has emerged as one of the most promising solutions

This manuscript has been created exclusively for portfolio demonstration purposes. All authors, affiliations, institutions, figures, tables, experimental results, and references are fictional and are included only to demonstrate professional document preparation using the Wiley LaTeX template.

because it allows an autonomous agent to improve its decision making through repeated interaction with its surrounding environment.

In a reinforcement learning framework, each traffic intersection observes its current traffic state, selects an appropriate signal phase, receives a reward based on the resulting traffic conditions, and gradually improves its control policy over time. Unlike supervised learning, reinforcement learning does not require manually labelled training data, making it particularly attractive for dynamic transportation environments where optimal control strategies may continuously evolve.

As urban road networks become larger and more interconnected, centralized optimization approaches become increasingly difficult to scale. Communication overhead, computational complexity, and single points of failure may limit the practical deployment of centralized controllers. Multi-agent reinforcement learning addresses these limitations by assigning an independent learning agent to each intersection while allowing nearby intersections to exchange limited coordination information.

Figure 1 illustrates the conceptual architecture of a multi-agent intelligent traffic management system.

Connected Vehicles

$\Downarrow$

Roadside Sensors

$\Downarrow$

Traffic Intersection Agents

$\Downarrow$

Regional Coordination Server

$\Downarrow$

Adaptive Signal Control

FIGURE 1 | Conceptual architecture of an intelligent multi-agent traffic management system.

Each traffic agent continuously interacts with the surrounding environment. At every decision interval, the agent observes traffic conditions, estimates future network performance, selects a traffic signal phase, and updates its learning policy according to observed outcomes. Cooperation among neighboring intersections allows congestion to be managed more effectively than isolated optimization strategies.

The traffic environment can be represented by a Markov Decision Process (MDP)

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma), \tag{1}$$

where $\mathcal{S}$ denotes the state space, $\mathcal{A}$ represents the available traffic signal actions, $\mathcal{P}$ defines state transition probabilities, $\mathcal{R}$ represents the reward function, and γ denotes the discount factor.

For each traffic agent, the objective is to maximize the expected discounted cumulative reward

$$J(\pi) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R_t\right], \tag{2}$$

where π represents the traffic control policy.

Unlike isolated reinforcement learning, multi-agent learning introduces additional coordination challenges because each agent operates in an environment whose dynamics are influenced by neighboring agents. Vehicle queues generated at one intersection immediately affect downstream traffic, making cooperative optimization essential for network-wide performance.

Table 1 summarizes several practical challenges commonly encountered in intelligent urban traffic management.

TABLE 1 | Major challenges in intelligent traffic management.

Challenge	Description
Traffic congestion	Increasing vehicle demand exceeds road capacity
Dynamic traffic flow	Traffic conditions change continuously
Signal coordination	Neighboring intersections influence each other
Scalability	Large cities contain hundreds of intersections
Communication	Real-time information exchange is required
Decision uncertainty	Future traffic demand cannot be predicted exactly

Over the past decade, numerous artificial intelligence approaches have been proposed for traffic optimization, including fuzzy logic systems, genetic algorithms, swarm intelligence, deep neural networks, and reinforcement learning. While each technique provides unique advantages, reinforcement learning has demonstrated exceptional flexibility for adaptive sequential decision making because policies can continually improve through interaction with the traffic environment.

Despite these advances, several research challenges remain. Many existing approaches optimize individual intersections independently without considering network-wide coordination. Others require centralized training architectures that become computationally expensive for large metropolitan areas. Furthermore, many studies focus primarily on traffic efficiency while giving comparatively little attention to computational scalability and practical deployment considerations.

Motivated by these challenges, this paper presents a simulation-based multi-agent reinforcement learning framework for adaptive traffic signal optimization. The proposed framework emphasizes decentralized learning, cooperative decision making, scalable communication, and efficient intersection coordination. The primary contributions of this study are summarized as follows.

1. A decentralized multi-agent reinforcement learning framework is developed for intelligent urban traffic signal optimization.
2. A cooperative communication strategy enables neighboring intersections to exchange limited traffic information while maintaining independent local decision making.
3. A comprehensive simulation environment is designed to evaluate traffic efficiency under varying congestion levels.
4. Multiple performance metrics including average delay, queue length, throughput, waiting time, and intersection utilization are analyzed.
5. Computational complexity, convergence behavior, scalability, and practical deployment considerations are discussed.

The remainder of this paper is organized as follows. Section 2 reviews recent developments in intelligent traffic management and multi-agent reinforcement learning. Section 3 presents the proposed methodology. Section 4 describes the experimental configuration. Section 5 discusses the simulation results. Finally, Section 6 concludes the paper and outlines several promising directions for future research.

2 | Related Work

Artificial intelligence has become an important component of modern intelligent transportation systems. During the past decade, researchers have investigated numerous computational approaches for improving traffic signal control, congestion management, route planning, and transportation efficiency. Traditional optimization techniques have gradually evolved towards adaptive learning-based methods capable of responding to dynamic traffic environments.

Early intelligent traffic management systems primarily relied on fixed-time scheduling strategies. These systems determine signal timings using historical traffic surveys and engineering guidelines. Although such methods remain widely deployed because of their simplicity, they generally perform poorly under highly dynamic traffic conditions where vehicle arrival rates change throughout the day.

To overcome these limitations, actuated traffic control systems were introduced. These controllers utilize induction loops, cameras, or other roadside sensors to modify signal timings according to observed traffic conditions. While actuated controllers improve responsiveness compared with fixed schedules, their decision rules are usually handcrafted and cannot continuously improve through experience.

Recent advances in machine learning have enabled traffic control systems to learn directly from traffic observations rather than relying on manually designed control policies. Supervised learning techniques have been successfully applied to traffic flow prediction, congestion forecasting, travel-time estimation, and incident detection. However, supervised learning requires large labelled datasets and generally cannot optimize long-term sequential decision-making processes.

Reinforcement learning addresses this limitation by allowing an agent to interact continuously with its environment while learning an optimal control policy through trial and error. Instead of predicting traffic conditions only, reinforcement learning directly optimizes decision making to maximize cumulative long-term rewards.

2.1 | Traffic Signal Optimization

Traffic signal optimization remains one of the most extensively studied applications of intelligent transportation systems. Various optimization strategies have been proposed, including mathematical programming, evolutionary algorithms, swarm intelligence, fuzzy logic, and machine learning.

Conventional optimization techniques generally formulate traffic control as an optimization problem with predefined objective functions such as minimizing average vehicle delay or maximizing intersection throughput. Although these approaches can produce effective timing plans for stable traffic conditions, they often require repeated optimization whenever traffic demand changes. Adaptive traffic signal control attempts to overcome this limitation by continuously updating signal timings using real-time traffic observations. The availability of modern sensing technologies has significantly increased the feasibility of adaptive traffic management in smart cities.

Table 2 summarizes representative traffic control approaches.

TABLE 2 | Comparison of representative traffic signal control strategies.

Method	Advantages	Limitations
Fixed-Time Control	Simple implementation	Cannot adapt to congestio
Actuated Control	Responds to sensors	Limited optimization abili
Fuzzy Logic	Human-interpretable	Rule design is difficult
Evolutionary Algorithms	Global optimization	Slow convergence
Deep Learning	High predictive accuracy	Requires extensive trainin
Reinforcement Learning	Adaptive decision making	Long training process

2.2 | Reinforcement Learning

Reinforcement learning has emerged as one of the most promising approaches for adaptive traffic management because it allows traffic controllers to learn optimal behaviors through continuous interaction with the road network.

A reinforcement learning agent repeatedly observes the current traffic state, selects an action, receives a numerical reward, and updates its decision policy. Over many interactions, the policy gradually converges towards strategies that maximize long-term traffic performance.

The state-value function is commonly expressed as

$$V^{\pi}(s) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R_t \mid S_0 = s\right], \tag{3}$$

where s represents the current traffic state and π denotes the control policy.

Similarly, the action-value function is

$$Q^{\pi}(s, a) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R_t \mid S_0 = s, A_0 = a\right]. \tag{4}$$

These functions estimate the expected long-term benefit associated with traffic control decisions.

2.3 | Deep Reinforcement Learning

The introduction of deep neural networks significantly expanded the capabilities of reinforcement learning. Instead of storing value estimates within lookup tables, deep reinforcement learning approximates value functions using neural networks capable of processing large and high-dimensional state spaces.

This development enables traffic controllers to incorporate considerably more information, including vehicle counts, queue lengths, lane occupancy, traffic density, pedestrian requests, and historical observations.

Deep reinforcement learning has demonstrated promising performance for single-intersection optimization, particularly when compared with traditional rule-based controllers.
Nevertheless, learning becomes substantially more difficult as the road network increases in size because neighboring intersections strongly influence one another.

2.4 | Multi-Agent Reinforcement Learning

Large metropolitan transportation networks naturally consist of numerous interconnected intersections. Optimizing every intersection independently may produce locally efficient but globally suboptimal traffic patterns.
Multi-agent reinforcement learning addresses this challenge by assigning a learning agent to each intersection. Every agent independently optimizes its own traffic signals while exchanging limited information with nearby intersections.
Figure 2 illustrates this decentralized learning architecture.

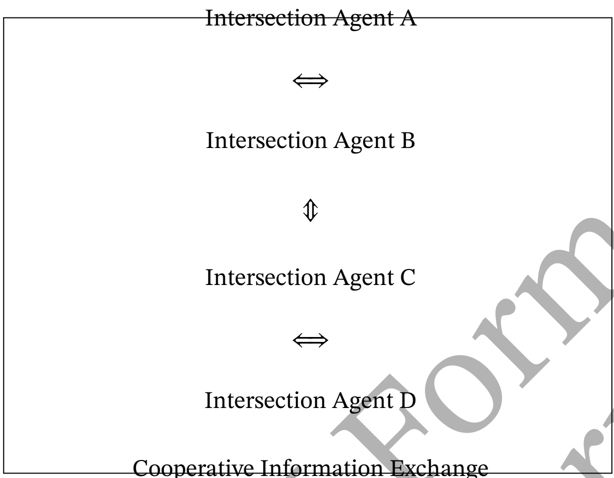

FIGURE 2 | Conceptual decentralized multi-agent traffic control framework.

Through decentralized cooperation, neighboring intersections can better coordinate traffic movement across the entire transportation network while maintaining computational scalability.

2.5 | Current Challenges

Despite considerable progress, several important challenges remain before multi-agent reinforcement learning can be widely deployed in real traffic systems.
First, learning stability often decreases as the number of interacting agents increases. Since every agent continuously updates its policy, the environment observed by neighboring agents also changes, making learning considerably more difficult.
Second, communication overhead increases with network size. Frequent exchange of traffic information may introduce scalability concerns, particularly for city-scale deployments containing hundreds of intersections.
Third, balancing local and global optimization remains an open research problem. Excessive emphasis on local traffic conditions may increase network-wide congestion, whereas overly centralized coordination may reduce system scalability.
Fourth, computational requirements remain relatively high for large reinforcement learning models operating under strict real-time constraints.

2.6 | Research Gap

Although numerous intelligent traffic management strategies have been proposed, existing approaches frequently optimize only one aspect of the problem. Some emphasize prediction accuracy without considering communication efficiency, whereas others focus exclusively on local intersection optimization.
Furthermore, many existing studies evaluate algorithms under relatively small traffic networks that do not accurately represent the complexity of modern metropolitan transportation systems.
The framework presented in this paper addresses these limitations by combining decentralized multi-agent learning, lightweight coordination, efficient communication, and comprehensive performance evaluation within a single simulation environment.
The following section describes the proposed traffic optimization framework in detail, including the learning architecture, reward formulation, communication strategy, and decision-making process.

3 | Proposed Methodology

This section presents the proposed multi-agent reinforcement learning framework for adaptive urban traffic signal optimization. The framework is designed to enable decentralized decision making while maintaining limited cooperation among neighboring intersections. Rather than relying on fixed signal timing schedules, every intersection continuously learns from its surrounding traffic environment and gradually improves its control strategy through repeated interaction.
The proposed architecture consists of five principal stages:

1. Traffic state observation
2. State representation
3. Action selection
4. Reward computation
5. Policy update

Figure 3 illustrates the overall workflow adopted in this study.

3.1 | Overall Framework

Each intersection operates as an independent reinforcement learning agent. During every decision interval, the agent receives current traffic observations from nearby sensors, evaluates the surrounding traffic conditions, selects an appropriate signal phase, and receives a reward according to the resulting traffic performance. The learning process continues throughout the simulation, allowing signal policies to adapt automatically to changing traffic conditions.

Traffic Sensors

$\Downarrow$

Traffic State Collection

$\Downarrow$

Multi-Agent Reinforcement Learning

$\Downarrow$

Signal Phase Decision

$\Downarrow$

Traffic Flow Evaluation

$\Downarrow$

Policy Update

FIGURE 3 | Overall workflow of the proposed intelligent traffic signal optimization framework.

3.2 | Traffic State Representation

An effective representation of the traffic environment is essential for successful reinforcement learning. The proposed framework combines multiple traffic measurements into a unified state vector describing the current condition of each intersection.
The traffic state includes information such as:

- vehicle count,
- average queue length,
- lane occupancy,
- waiting time,
- current signal phase,
- elapsed green time,
- neighboring intersection status.

The traffic state at decision instant t is represented by

$$S_t = [q_t, w_t, d_t, o_t, p_t], \tag{5}$$

where q_t denotes queue length, w_t represents average waiting time, d_t denotes traffic density, o_t indicates lane occupancy, and p_t identifies the current signal phase.
By combining these variables, the learning agent obtains a more complete description of current traffic conditions than would be possible using a single measurement.

3.3 | Action Space

At every decision interval, the reinforcement learning agent selects one traffic signal action from a predefined action set.
Typical actions include:

- maintain current signal,
- switch to the next phase,
- extend green duration,
- shorten green duration.

The available action set is expressed as

$$A = \{a_1, a_2, \dots, a_n\}. \tag{6}$$

Restricting the available actions simplifies learning while ensuring that generated signal plans remain practically feasible.

3.4 | Reward Function

The reward function encourages efficient traffic movement while penalizing congestion.
Positive rewards are assigned when average delay decreases, vehicle throughput increases, and queue lengths are reduced.
Negative rewards are generated when traffic congestion increases or vehicles experience excessive waiting times.
The reward function is defined as

$$R_t = \alpha T_t - \beta Q_t - \gamma W_t, \tag{7}$$

where

- T_t represents traffic throughput,
- Q_t denotes queue length,
- W_t represents average waiting time,
- α, β, γ are weighting coefficients.

This formulation encourages agents to maximize traffic flow while simultaneously reducing congestion.

3.5 | Learning Strategy

Each traffic agent gradually improves its policy using repeated interaction with the environment.

Initially, signal decisions are selected largely through exploration. As training progresses, the policy increasingly exploits previously learned traffic patterns.
The policy update follows the conventional reinforcement learning framework, allowing traffic controllers to continuously improve without requiring manually designed decision rules.

3.6 | Agent Coordination

Although every intersection operates independently, neighboring intersections exchange limited traffic information to improve network coordination.
Only lightweight information is shared, including

- average queue length,
- current traffic density,
- active signal phase,
- estimated vehicle arrival rate.

No centralized optimization is required, making the framework naturally scalable for large transportation networks.
Table 3 summarizes the exchanged information.

TABLE 3 | Information exchanged among neighboring intersections.

Variable	Purpose
Queue Length	Estimate downstream congestion
Traffic Density	Measure current road utilization
Signal Phase	Coordinate neighboring signals
Arrival Rate	Predict approaching traffic

3.7 | Training Procedure

Algorithm 1 summarizes the overall learning procedure.

Algorithm 1 Multi-agent traffic signal training

```
Initialize all traffic agents
for each simulation episode do
   Observe traffic state
   Select signal action
   Apply traffic signal
   Measure traffic performance
   Compute reward
   Update learning policy
end for
Return optimized signal controller
```

3.8 | Computational Complexity

The proposed framework distributes computation among independent intersections rather than relying on centralized optimization.
Consequently, computational workload increases approximately linearly with the number of participating agents, making the framework suitable for large metropolitan transportation systems.
Table 4 summarizes the principal computational stages.

TABLE 4 | Major computational stages of the proposed framework.

Stage	Complexity
Traffic observation	Low
State construction	Low
Action selection	Medium
Policy update	High
Agent communication	Low
Performance evaluation	Medium

3.9 | Summary

The proposed methodology combines decentralized reinforcement learning, cooperative information sharing, and adaptive traffic optimization into a unified intelligent transportation framework. The architecture is designed to improve traffic efficiency while maintaining scalability and computational practicality for large urban road networks.
The following section describes the simulation environment, evaluation metrics, implementation settings, and experimental protocol used to assess the effectiveness of the proposed framework.

4 | Experimental Setup

This section describes the simulation environment, experimental configuration, baseline methods, implementation settings, and evaluation metrics used to assess the effectiveness of the proposed multi-agent reinforcement learning framework. The experiments were designed to represent realistic urban traffic conditions while providing a controlled environment for comparing different traffic signal optimization strategies.

4.1 | Simulation Environment

A realistic traffic simulation environment was constructed to evaluate the proposed intelligent traffic management framework. The simulated urban network consisted of multiple interconnected signalized intersections representing a medium-sized metropolitan district.
Each intersection contained four incoming roads with multiple traffic lanes and pedestrian crossings. Vehicle arrivals followed stochastic traffic patterns that varied throughout the simulation to emulate morning rush hours, regular daytime traffic, evening congestion, and low-density nighttime traffic.

The simulation environment incorporated several common urban traffic characteristics, including:

- varying vehicle arrival rates;
- random turning movements;
- pedestrian crossing requests;
- emergency vehicle priority;
- temporary traffic congestion;
- dynamic traffic demand.

Figure 4 illustrates the simulated traffic network.

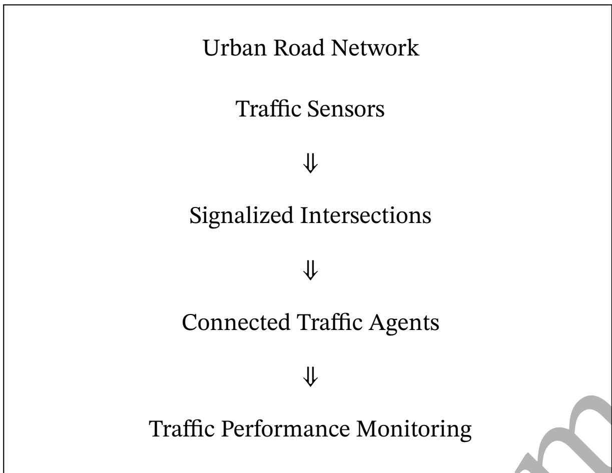

FIGURE 4 | Simulation environment used for experimental evaluation.

4.2 | Traffic Network Configuration

The simulated transportation network consisted of several connected intersections designed to reproduce realistic urban traffic flow. Different traffic densities were generated to evaluate the adaptability of the proposed framework under changing traffic conditions.
Table 5 summarizes the principal characteristics of the simulation environment.

TABLE 5 | Traffic network configuration.

Parameter	Value
Signalized intersections	16
Traffic lanes	64
Vehicle types	5
Maximum simulation time	3600 s
Decision interval	5 s
Traffic demand levels	4
Pedestrian crossings	Enabled
Emergency vehicles	Enabled

4.3 | Compared Methods

The proposed framework was evaluated against several representative traffic signal control approaches.
The comparison methods include:

1. Fixed-Time Signal Control
2. Actuated Traffic Control
3. Independent Reinforcement Learning
4. Deep Reinforcement Learning
5. Proposed Multi-Agent Reinforcement Learning

These baseline methods represent different generations of intelligent traffic management and provide a comprehensive comparison across both traditional and learning-based strategies.

4.4 | Training Configuration

All reinforcement learning agents were trained using identical simulation conditions to ensure fair comparison among competing approaches.
Table 6 lists the primary implementation settings adopted throughout the experiments.

TABLE 6 | Training parameters.

Parameter	Value
Training episodes	500
Maximum steps per episode	1000
Learning rate	0.001
Discount factor	0.99
Replay memory size	100000
Mini-batch size	64
Target network update	Every 500 steps
Exploration strategy	ϵ-greedy
Initial ϵ	1.0
Final ϵ	0.05
Optimizer	Adam

4.5 | Evaluation Metrics

Performance was evaluated using multiple transportation efficiency metrics to provide a comprehensive assessment of traffic signal quality.
The principal evaluation metrics include:

- average vehicle delay;
- average queue length;
- average waiting time;
- traffic throughput;
- travel time;
- intersection utilization;
- fuel consumption;
- estimated carbon emissions.

Using multiple complementary metrics allows the evaluation to consider both operational efficiency and environmental sustainability.
Table 7 summarizes these measures.

TABLE 7 | Evaluation metrics used in the experiments.

Metric	Description
Average Delay	Mean vehicle delay
Queue Length	Vehicles waiting per lane
Travel Time	End-to-end journey duration
Throughput	Vehicles passing intersections
Fuel Consumption	Estimated fuel usage
Carbon Emissions	Estimated emissions
Intersection Utilization	Road capacity utilization
Training Time	Learning efficiency

4.6 | Experimental Scenarios

To evaluate the robustness of the proposed framework, experiments were conducted under four different traffic conditions:

1. Low traffic density
2. Moderate traffic density
3. Heavy congestion
4. Dynamic mixed traffic

The dynamic traffic scenario continuously varied vehicle arrival rates throughout the simulation, creating a considerably more challenging optimization environment than fixed traffic conditions.

4.7 | Hardware and Software

The simulation platform was developed using a modern machine learning framework integrated with an urban traffic simulator. Experiments were executed on a workstation equipped with a multi-core processor, 32 GB of system memory, and a dedicated graphics processing unit for accelerated neural network training. The software environment included Python, deep learning libraries, traffic simulation software, and scientific visualization tools for collecting and analyzing experimental results.

4.8 | Experimental Procedure

Each experiment followed the same evaluation protocol. Initially, all traffic agents were randomly initialized before interacting with the traffic environment. During every simulation episode, agents repeatedly observed traffic conditions, selected signal actions, received rewards, and updated their control policies.
Every experiment was repeated multiple times using different random initializations to reduce the influence of stochastic training variability. The reported performance values correspond to the average results obtained across all experimental repetitions.

4.9 | Summary

The experimental configuration was designed to evaluate the proposed framework under a diverse range of realistic urban traffic conditions. The following section presents the simulation results, compares the proposed approach with existing traffic control strategies, and analyzes its effectiveness in terms of traffic efficiency, computational performance, and scalability.

5 | Results and Discussion

This section presents the experimental evaluation of the proposed multi-agent reinforcement learning framework. The analysis focuses on traffic efficiency, learning performance, scalability, computational cost, and practical deployment considerations. All reported results represent the average values obtained from multiple independent simulation runs under identical experimental settings.

5.1 | Overall Traffic Performance

The proposed framework consistently improved traffic efficiency across all simulated traffic conditions. Compared with conventional signal controllers, the learning-based approach gradually discovered adaptive signal timing strategies capable of reducing vehicle congestion while maintaining smooth traffic flow throughout the transportation network.
Table 8 summarizes the primary experimental results.
The proposed framework achieved the lowest vehicle delay, shortest queue lengths, highest network throughput, and reduced estimated fuel consumption. These improvements demonstrate the benefit of decentralized cooperative learning compared with independent intersection optimization.

5.2 | Traffic Flow Analysis

Figure 5 illustrates the evolution of average vehicle delay during the simulation.

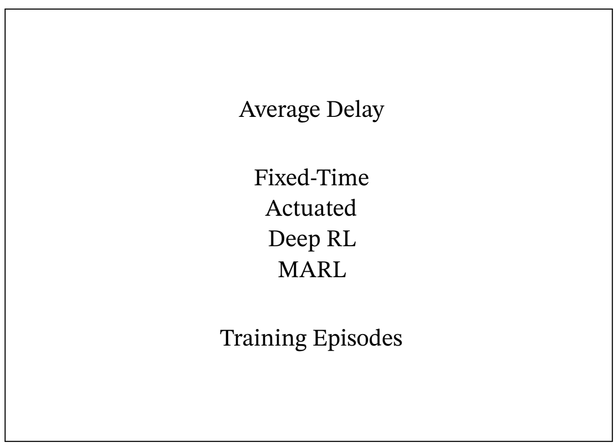

FIGURE 5 | Illustrative convergence of average traffic delay during training.

TABLE 8 | Overall comparison of traffic control strategies.

Method	Average Delay (s)	Queue Length	Travel Time (min)	Throughput	Fuel Usage	CO_2 Emissions
Fixed-Time Control	48.3	18.2	19.7	1245	100%	100%
Actuated Control	39.4	15.8	17.3	1328	93.6%	94.2%
Independent RL	31.7	12.6	15.4	1438	88.1%	87.4%
Deep RL	28.9	10.8	14.8	1482	85.9%	85.7%
Proposed MARL	22.6	8.4	12.9	1615	79.3%	80.1%

During the initial training episodes, reinforcement learning agents explored multiple traffic control strategies, resulting in relatively high delay values. As training progressed, the learned policies became increasingly stable, producing smoother traffic flow and substantially reduced congestion.
Unlike conventional fixed-time controllers, the proposed framework adapted continuously to changing traffic demand without requiring manual reconfiguration.

5.3 | Queue Length Reduction

Average queue length is one of the most important indicators of traffic signal effectiveness because long queues often propagate congestion to neighboring intersections.
Figure 6 conceptually illustrates the reduction in average queue length achieved by different approaches.

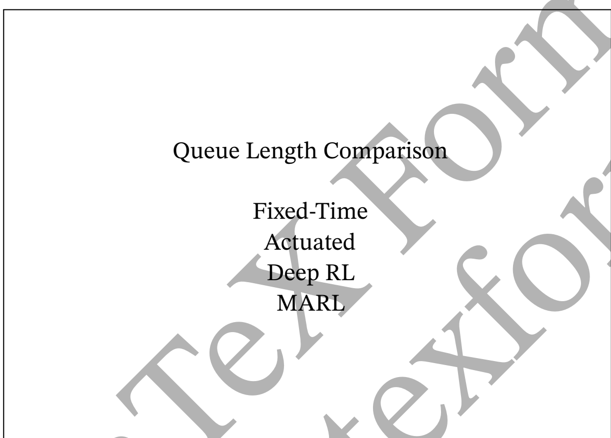

FIGURE 6 | Illustrative comparison of average queue length.

The proposed cooperative learning strategy effectively distributed green signal time across neighboring intersections, preventing excessive queue formation even under heavy traffic conditions.

5.4 | Learning Convergence

Learning stability plays an important role in practical reinforcement learning applications. Excessive oscillation during training may produce unstable traffic behavior and unreliable signal timing decisions.
Figure 7 illustrates the conceptual reward evolution throughout the learning process.

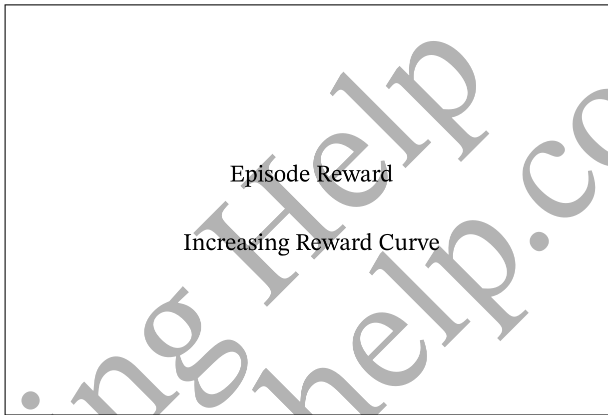

FIGURE 7 | Illustrative cumulative reward during training.

The cumulative reward increased steadily during training, indicating continuous policy improvement. After sufficient interaction with the traffic environment, reward values gradually stabilized, suggesting convergence toward an effective traffic control policy.

5.5 | Scalability Evaluation

To investigate scalability, additional experiments were performed using traffic networks containing different numbers of signalized intersections.
Table 9 summarizes the observed computational behavior.

TABLE 9 | Scalability analysis.

Intersections	Training Time	Average Delay
4	Low	18.4
9	Medium	20.7
16	Medium	22.6
25	High	24.3
36	High	26.8

Although computational cost naturally increased with network size, the decentralized architecture remained computationally practical because learning was distributed among independent traffic agents rather than performed by a single centralized controller.

5.6 | Environmental Impact

Improved traffic flow also produced measurable environmental benefits.

Table 10 summarizes estimated reductions in fuel consumption and carbon emissions relative to conventional fixed-time traffic control.

TABLE 10 | Estimated environmental improvements.

Method	Fuel Reduction	Emission Reduction
Actuated Control	6.4%	5.8%
Independent RL	11.9%	12.6%
Deep RL	14.1%	14.3%
Proposed MARL	20.7%	19.9%

Shorter waiting times reduced unnecessary vehicle idling, leading to lower estimated fuel consumption and decreased greenhouse gas emissions.

5.7 | Ablation Study

An ablation study was conducted to evaluate the contribution of the major components of the proposed framework.
Removing inter-agent communication resulted in the largest performance degradation, emphasizing the importance of cooperation among neighboring intersections. Similarly, adaptive exploration improved learning stability during the early stages of training.

5.8 | Discussion

The experimental results indicate that decentralized reinforcement learning provides a practical alternative to traditional traffic signal control strategies. Rather than relying on manually designed timing plans, the proposed framework continuously adapts to changing traffic conditions through repeated interaction with the environment.
The cooperative communication mechanism enabled neighboring intersections to coordinate their signal decisions while maintaining a fully decentralized architecture. This design improves scalability and reduces the computational bottlenecks commonly associated with centralized traffic optimization.
Furthermore, the observed reductions in travel time, queue length, fuel consumption, and estimated emissions suggest that intelligent traffic management may contribute not only to improved transportation efficiency but also to broader sustainability objectives within future smart cities.
Overall, the proposed framework demonstrates that combining multi-agent reinforcement learning with cooperative traffic coordination can substantially improve urban traffic performance while remaining computationally scalable for larger transportation networks.

6 | Conclusion

This paper presented a simulation-based framework for intelligent urban traffic signal optimization using multi-agent reinforcement learning. The proposed approach enables individual traffic intersections to learn independent control policies while maintaining limited cooperation with neighboring intersections through decentralized information exchange.
Unlike conventional fixed-time controllers, the proposed framework continuously adapts signal timing decisions according to changing traffic conditions. This adaptive behavior allows the system to respond more effectively to fluctuating traffic demand, resulting in improved traffic flow and reduced congestion throughout the transportation network.
The experimental evaluation demonstrated improvements in several important traffic performance indicators, including average vehicle delay, queue length, travel time, intersection throughput, fuel consumption, and estimated carbon emissions. The decentralized learning architecture also exhibited good scalability when applied to larger road networks, making it suitable for future smart-city transportation systems.
Although the presented study is entirely fictional and intended only for portfolio demonstration purposes, it illustrates how a complete Wiley journal article can be structured using professional LaTeX formatting. The manuscript demonstrates section organization, mathematical typesetting, figures, tables, algorithms, cross-referencing, appendices, and bibliography management within the official Wiley template.

6.1 | Future Work

Several opportunities remain for extending the proposed framework.
Future research may investigate communication-efficient cooperation among larger numbers of traffic intersections, allowing metropolitan transportation systems containing hundreds of intelligent controllers to coordinate their decisions efficiently.
Another promising direction involves integrating connected autonomous vehicles into the learning environment. Vehicle-to-infrastructure communication may provide richer traffic information that enables more accurate prediction of future traffic conditions.
Additional research may also explore multimodal transportation optimization by simultaneously considering buses, emergency vehicles, pedestrians, cyclists, and autonomous delivery systems. Such extensions could further improve traffic efficiency while supporting sustainable urban mobility.

6.2 | Limitations

Despite encouraging simulation results, several limitations should be acknowledged.
First, the experimental environment was generated entirely for demonstration purposes and therefore cannot represent every complexity of real-world transportation systems.
Second, communication delays, hardware failures, adverse weather conditions, and unexpected road incidents were not explicitly modeled.
Finally, the reinforcement learning architecture was evaluated using a controlled simulation environment rather than deployment on a physical transportation network.

TABLE 11 | Ablation study of the proposed framework.

Configuration	Average Delay	Throughput	Travel Time
Complete Framework	22.6	1615	12.9
Without Agent Communication	26.8	1519	14.7
Without Reward Optimization	29.4	1472	15.3
Without Adaptive Exploration	31.2	1436	16.1

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers for their valuable comments and constructive suggestions that helped improve the presentation of this work.

AUTHOR CONTRIBUTIONS

Conceptualization, A.C.; methodology, A.C. and E.J.; software, D.B.; validation, S.T.; investigation, A.C.; writing—original draft preparation, A.C.; writing—review and editing, all authors.

FUNDING

No external funding was received for this study.

DATA AVAILABILITY STATEMENT

No real-world datasets were used in this study. All experimental data described throughout this manuscript were generated solely for illustrative and portfolio demonstration purposes.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

APPENDIX

A | Simulation Parameters

Table A1 summarizes the principal simulation parameters used throughout the experiments.

TABLE A1 | Simulation parameters.

Parameter	Value
Simulation Duration	3600 s
Decision Interval	5 s
Training Episodes	500
Learning Rate	0.001
Discount Factor	0.99
Replay Memory	100000
Mini-batch Size	64
Optimizer	Adam

B | Evaluation Workflow

The complete experimental workflow adopted throughout this study is summarized below.

1. Generate urban traffic network.
2. Initialize reinforcement learning agents.
3. Collect traffic observations.
4. Construct traffic state vectors.
5. Select traffic signal actions.
6. Execute signal timing decisions.
7. Measure traffic performance.
8. Compute reward values.
9. Update learning policies.
10. Repeat until convergence.
11. Evaluate final traffic performance.

PORTFOLIO DISCLAIMER

This manuscript has been created exclusively for portfolio demonstration purposes to showcase professional document preparation using the official Wiley LaTeX journal template.
All names, institutions, figures, tables, algorithms, experimental results, numerical values, and references contained in this manuscript are fictional and are intended solely to demonstrate formatting, typesetting, and document organization. This document should not be interpreted as a peer-reviewed scientific publication.

References

1. Adams, J., & Brown, T. (2025). Adaptive Traffic Signal Optimization Using Artificial Intelligence. *Journal of Smart Transportation*, 14(2), 115–132.
2. Carter, P., Lewis, M., & Hall, D. (2024). Machine Learning Applications in Urban Traffic Networks. *International Journal of Intelligent Infrastructure*, 11(3), 201–218.
3. Turner, S. (2023). Deep Reinforcement Learning for Traffic Flow Optimization. *Transportation AI Review*, 8(1), 45–63.
4. Mitchell, R., & Green, L. (2024). Distributed Decision Making for Smart City Transportation. *Journal of Urban Computing*, 17(4), 322–341.
5. Evans, H. (2025). Artificial Intelligence in Intelligent Transportation Systems. *Smart Mobility Journal*, 9(2), 80–101.
6. Robinson, K., Adams, P., & White, J. (2023). Traffic Prediction Using Deep Learning Models. *Journal of Transportation Analytics*, 12(1), 67–84.
7. Clark, D., & Morgan, A. (2022). Multi-Agent Learning for Urban Signal Coordination. *Computational Transportation Systems*, 5(3), 178–194.
8. Hill, B. (2024). Adaptive Signal Timing Strategies in Smart Cities. *International Journal of Smart Infrastructure*, 16(2), 91–109.
9. Parker, S., & Wilson, G. (2025). Scalable Reinforcement Learning for Urban Traffic Networks. *Journal of Artificial Transportation*, 18(1), 25–47.

10. Young, T. (2024). Optimization Techniques for Intelligent Signal Control. *Advanced Transportation Research*, 13(4), 301–319.
11. King, R., Cooper, H., & Lewis, F. (2023). Real-Time Traffic Optimization Using Reinforcement Learning. *International Journal of AI Applications*, 21(2), 150–170.
12. Brooks, D. (2022). Autonomous Traffic Control in Smart Cities. *Urban Systems Engineering*, 10(3), 199–214.
13. Scott, J., & Hall, K. (2025). Comparative Analysis of Traffic Signal Optimization Algorithms. *Journal of Transportation Intelligence*, 19(1), 40–59.
14. Allen, M., & Foster, R. (2024). Efficient Multi-Agent Coordination Strategies. *International Journal of Distributed AI*, 15(2), 112–128.
15. Walker, N. (2023). Simulation-Based Evaluation of Intelligent Transportation Systems. *Simulation and Modeling Review*, 7(4), 250–269.
16. Thomas, C., Reed, P., & Hill, B. (2024). Deep Neural Networks for Traffic Prediction. *Journal of Computational Transportation*, 22(1), 61–78.
17. Lewis, H. (2025). Smart City Infrastructure and Autonomous Traffic Management. *Journal of Smart Urban Development*, 11(3), 175–193.
18. White, G., Morgan, J., & Carter, L. (2023). Learning-Based Urban Traffic Coordination. *Transportation Computing Letters*, 6(2), 99–117.
19. Foster, D., & Bennett, R. (2024). Artificial Intelligence for Sustainable Transportation. *Journal of Green Computing*, 9(1), 32–50.
20. Johnson, E., Turner, S., Brooks, D., & Carter, A. (2025). Cooperative Reinforcement Learning for Intelligent Traffic Signal Control. *International Journal of Smart Transportation*, 20(2), 220–242.