

DQN-Based Coordinated Traffic Flow Control for Adjacent Intersections

Anushka Kalbande

Department of Computer Science and Engineering
Anjuman College of Engineering and Technology Nagpur,
India
anushkakalbande28@gmail.com

Dr. Mohammad Anwarul Siddique

Department of Computer Science and
Engineering Anjuman College of
Engineering and Technology Nagpur,
India
anwar.ideacomputsoft@gmail.com

Abstract—Urban traffic congestion remains a major challenge in intelligent transportation systems, particularly in urban corridors containing closely spaced signalized intersections. Conventional fixed-time control schemes operate with predetermined phase durations and do not adapt effectively to dynamic traffic patterns, queue spillback, or interactions between neighboring intersections. Reinforcement Learning (RL) has emerged as a promising approach for adaptive signal control because it enables a controller to learn from direct interaction with the traffic environment. However, a large part of the existing work focuses on isolated intersections, while real-world urban traffic networks require coordinated decision-making across connected nodes. This paper presents a research-based framework for coordinated traffic flow control at adjacent intersections using a Deep Q-Network (DQN) in a SUMO-TraCI simulation environment. The study formulates the corridor control problem as a Markov Decision Process in which a coordinated state representation captures queue conditions and active signal phases across neighboring intersections. A DQN agent is designed to learn adaptive signal switching policies that account for local congestion as well as downstream traffic conditions. The paper emphasizes methodology, system design, and evaluation strategy rather than reporting fixed numerical claims. The proposed framework is intended to support comparative study against conventional fixed-time control and independent local decision strategies using metrics such as delay, queue length, throughput, and travel-time behavior. The work contributes a structured, simulation-driven, and academically grounded basis for multi-node traffic signal research that can be extended toward larger corridors, multi-agent control, and sustainability-aware traffic management.

Index Terms—Reinforcement Learning, Deep Q-Network, Traffic Signal Control, Adjacent Intersections, Coordinated Control, SUMO, TraCI, Intelligent Transportation Systems

I. INTRODUCTION

Traffic congestion has become one of the most persistent operational and economic challenges in contemporary urban transportation systems. Rapid urban growth, mixed land use, increasing travel demand, and rising private vehicle ownership have placed severe pressure on existing road infrastructure. Signalized intersections often emerge as the most critical bottlenecks in this setting because they regulate competing traffic movements and influence the progression of vehicles across entire corridors. When signals are not timed appropriately, congestion builds rapidly, queues extend upstream, and travel times become unstable. This not only reduces road

efficiency but also increases driver frustration, fuel wastage, environmental burden, and uncertainty in daily commuting.

Conventional traffic signal controllers in many urban areas still rely on fixed-time plans. These plans are typically designed using historical averages, assumed demand patterns, or periodically updated field studies. Although such approaches can be acceptable under stable and predictable conditions, they respond poorly to real-time traffic fluctuations. Morning and evening peaks, incidents, uneven directional flow, road works, and random disturbances often create conditions far from the assumptions under which fixed-time controllers are designed. Under these circumstances, green time may be allocated to empty or lightly loaded approaches while congested approaches continue to accumulate queues. As a result, static control logic often underutilizes intersection capacity and propagates congestion to neighboring locations.

The limitations of traditional signal control become more pronounced when intersections are closely spaced. In arterial roads or urban corridors, the outflow from one intersection becomes the inflow to the next. This introduces a strong interdependence between adjacent intersections. If the upstream intersection releases a vehicle platoon without considering the downstream queue condition, the downstream intersection may quickly saturate, causing spillback into the link between the nodes. Once spillback occurs, it can block upstream discharge, reduce effective green utilization, and deteriorate corridor-wide performance. Therefore, optimizing one intersection in isolation does not necessarily improve overall traffic efficiency. Instead, effective corridor management requires coordination between neighboring signals.

This challenge has motivated the development of adaptive and intelligent traffic control methods. Sensor-driven adaptive systems use live traffic information to adjust phase duration and phase order. While such systems are more responsive than fixed-time schemes, they are often based on handcrafted logic, thresholds, or rule-based heuristics. Their performance depends heavily on



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

calibration, and they may struggle to generalize across changing or previously unseen traffic conditions. In contrast, Reinforcement Learning offers a data-driven framework in which a controller learns an effective policy

by observing the environment, taking actions, and receiving reward feedback. This learning paradigm is particularly attractive for traffic control because the problem naturally involves sequential decision-making under uncertainty.

In an RL-based traffic signal framework, the controller functions as an agent interacting with a traffic environment. The state may include detector readings, queue lengths, occupancy, or active phase information. The action may correspond to keeping the current signal phase, extending green, or switching to another phase. The reward reflects traffic-control objectives such as reducing queues, minimizing waiting time, or improving throughput. Through repeated interaction, the agent gradually improves its policy to maximize long-term reward. This formulation allows control strategies to emerge from experience rather than from manually designed timing plans.

Deep Reinforcement Learning extends this idea by using neural networks to approximate action-value functions or policies in problems with larger state spaces. Among these methods, the Deep Q-Network has become a widely adopted baseline for traffic signal control because it combines Q-learning with neural-network approximation, replay memory, and target-network stabilization. DQN is well suited to traffic environments in which the controller must process multiple detector inputs, traffic states, and signal configurations. It also offers a practical balance between implementation complexity and learning capability, making it attractive for research-based simulation studies.

Despite the growing use of RL in traffic signal optimization, a large portion of existing work still concentrates on single intersections. Such studies are valuable because they isolate local decision dynamics and make controlled experimentation easier. However, real-world traffic systems are not composed of isolated nodes. Vehicles move through connected corridors and networks in which traffic states are strongly coupled across intersections. A controller that performs well locally may still generate poor corridor performance if it ignores downstream congestion or upstream release patterns. This gap between isolated optimization and network-aware coordination motivates the present study.

The focus of this paper is therefore on coordinated traffic flow control for adjacent intersections using DQN. The work is intentionally framed as research-based and simulation-based rather than result-claim oriented. Instead of presenting hard numerical improvements without implementation evidence, the paper provides a detailed methodological framework that can support experimental study, comparative evaluation, and future extension. The proposed system models two connected intersections within a SUMO simulation and uses TraCI to interact with a DQN controller in real time. The controller receives a coordinated state representation that includes traffic information from both intersections and selects actions with corridor-level awareness.

The main contributions of this paper are summarized below:

- It formulates adjacent-intersection traffic signal coordination as a research-oriented Deep Reinforcement Learning problem in a reproducible simulation environment.

- It proposes a coordinated state representation that captures traffic conditions across neighboring intersections instead of treating each node independently.
- It defines a DQN-based control architecture suitable for adaptive multi-node traffic signal operation using SUMO and TraCI.
- It establishes an evaluation framework for comparing co-ordinated control against fixed-time and independent local strategies using congestion- and mobility-related metrics.
- It provides a scalable basis for future work in corridor control, multi-agent reinforcement learning, reward engineering, and sustainability aware traffic optimization.

The rest of the paper is organized as follows. Section II reviews the evolution of traffic signal control methods and related work on reinforcement learning for coordinated intersections. Section III presents the problem statement and research objectives. Section IV explains the proposed methodology and control formulation. Section V describes the simulation environment, controller design, and evaluation strategy. Section VI presents a structured discussion framework for the expected analysis and qualitative interpretation of outcomes. Section VII concludes the paper and identifies promising future directions.

II. LITERATURE REVIEW

A. Evolution of Traffic Signal Control

Traffic signal control has a long history rooted in analytical timing and operational heuristics. Early approaches treated the problem as one of deterministic optimization based on estimated arrival rates, saturation flows, and cycle-length calculations. Classical fixed-time design methods aimed to minimize average delay under assumed steady-state traffic demand and were appropriate in settings where traffic variation was limited. Over time, these methods were enhanced through coordination offsets, progression bands, and corridor-level timing plans intended to improve arterial movement.

Although analytical and fixed-time methods remain important from an operational perspective, they are inherently limited by their dependence on prior assumptions. Urban traffic patterns rarely remain stationary for long durations. Variability caused by school hours, office peaks, weather changes, incidents, and route choice behavior creates conditions that challenge rigid timing plans. To address these issues, actuated and adaptive control methods were introduced. Actuated systems rely on detectors to extend or terminate green phases based on real-time demand, while network-level adaptive systems update splits, cycles, and offsets according to broader traffic patterns. These methods can improve responsiveness, but they still depend on rule-based logic and extensive calibration.



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

Modern traffic management increasingly requires systems that can learn complex interactions rather than merely react to threshold conditions.

B. Reinforcement Learning for Traffic Signal Control

Reinforcement Learning has gained attention as a suitable paradigm for adaptive traffic signal control because it directly models the interaction between a decision-making agent and a dynamic environment. In RL, the traffic signal controller receives observations of traffic conditions, selects actions that affect signal timing, and updates its decision policy based on observed rewards. This sequential framework is well aligned with the operational nature of traffic control, where each signal decision affects future queues, arrivals, and delays.

Early traffic RL research used tabular methods such as Q-learning. These studies demonstrated that an agent can learn to improve intersection performance without relying on pre-computed timing plans. State definitions in these works often included queue counts, waiting vehicles, or occupancy values, while actions typically involved phase switching decisions. Although promising, tabular approaches become difficult to scale when the number of detectors, lanes, or intersections increases. The size of the state-action table grows rapidly, and the learned policies may become unstable or impractical in realistic scenarios.

Deep Reinforcement Learning addressed this limitation by replacing lookup tables with neural-network approximators. Deep Q-Networks, in particular, became influential because they enabled learning in higher-dimensional spaces while stabilizing training through replay memory and target networks. In traffic signal control, DQN-based methods have been studied for single intersections, arterial corridors, and grid networks. These methods generally aim to reduce congestion-related measures such as waiting time, queue length, or cumulative delay. Their advantage lies in the ability to learn non-linear control behavior from raw or semi-processed traffic observations.

Many simulation-based studies employ the SUMO platform together with TraCI for RL experimentation. SUMO allows researchers to create customizable road networks, define route demand, simulate microscopic vehicle behavior, and access simulation states programmatically. TraCI enables real-time interaction between the simulation and an external RL controller, making it possible to observe detector outputs, update signals, and compute rewards at each control step. This combination has become a standard environment for prototyping traffic signal RL algorithms because it balances realism, control, and reproducibility.

C. Single-Intersection and Multi-Intersection Studies

A substantial amount of RL-based traffic signal research focuses on isolated intersections. This is understandable because single-node settings simplify the experimental design and allow clearer analysis of controller learning. Many such studies have reported that adaptive learning approaches can outperform fixed-time control in terms of queue dissipation, waiting time, or traffic progression. These studies also high-light the importance of state representation, reward definition, and training duration.

However, corridor traffic behavior cannot be understood solely from isolated nodes. Adjacent intersections influence each other through platoon movement, signal progression, residual queues, and spillback. When traffic released from one node reaches the next node during an unfavorable phase or when downstream links become saturated, upstream operations are affected immediately. This interdependency means that lo-cal optimality does not guarantee network-level efficiency. As a result, the shift from single-intersection to multi-intersection control is not merely a scale increase; it is a structural change in the control problem.

Research on multi-intersection RL has taken several forms. Some studies assign an independent agent to each node, where each controller learns locally and interacts indirectly through the shared traffic environment. Other studies use coordinated or centralized frameworks in which a joint state representation is used to generate corridor-level actions. Multi-agent RL has also been explored for larger networks, but such approaches introduce communication, credit-assignment, and scalability challenges. For a research-oriented M.Tech study, a two-node adjacent-intersection corridor provides a suitable balance: it is rich enough to capture coordination effects, yet small enough to permit clear modeling and explanation.

D. Deep Q-Networks in Coordinated Control

DQN is often selected for traffic control research because it is conceptually simpler than many newer DRL methods while still powerful enough for meaningful adaptive control. The use of experience replay helps decorrelate training samples, and the target network reduces instability caused by rapidly changing Q-value targets. In traffic applications, these features make DQN a strong baseline for testing whether learned control policies can respond to variable demand and cross-node interactions.

For coordinated adjacent-intersection control, DQN offers an attractive framework because the joint state can include traffic measurements from both nodes, and the action can be defined in terms of coordinated switching behavior. This allows one controller to learn how upstream and downstream states should influence one another. For example, the controller can learn not to discharge an upstream queue aggressively when downstream storage is constrained. Similarly, it can learn timing patterns that support corridor progression rather than independent stop-start behavior.

Nevertheless, DQN-based coordinated control also faces design challenges. The state space grows when multiple inter-sections are included. The action space becomes more complex because coordinated decisions involve multiple signals simul-taneously. Reward design must also



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

be chosen carefully so that the controller does not optimize one node at the expense of another. These issues make coordinated DQN control a meaningful research problem even before numerical results are considered.

TABLE I
COMPARISON OF RELATED WORKS ON TRAFFIC SIGNAL CONTROL

Author / Year	Control Type	Network Size	Method	Reward Focus	Main Limitation	Relevance to This Work
Webster	Fixed-time	Single intersection	Analytical timing	Delay minimization	Assumes stable demand and cannot adapt to traffic fluctuations	Provides classical baseline for signal timing
Papageorgiou et al.	Adaptive / coordinated	Network-level	Survey of control strategies	Operational efficiency	Many methods require calibration and are not learning-based	Motivates the need for adaptive and intelligent control
Watkins and Dayan	General RL foundation	Generic MDP setting	Q-learning	Reward-driven learning	Not traffic-specific; tabular scaling limitations	Provides theoretical basis for RL control
Mnih et al.	General DRL	Generic sequential decision tasks	DQN	Value maximization	Not designed for traffic directly	Supports use of deep value approximation
Traffic RL studies on single intersections	Adaptive	Single intersection	Q-learning / DQN	Queue or delay reduction	Limited to isolated nodes and local optimization	Highlights the gap between local and coordinated control
Coordinated arterial control studies	Coordinated	Two or more intersections	RL / max-pressure / hybrid methods	Queue progression and throughput	Often complex or not directly comparable across settings	Closely related to adjacent-intersection coordination
This paper	Coordinated adaptive control	Two adjacent intersections	DQN in SUMO	Corridor-level queue, delay, and throughput	Research-based simulation study; experimental results to be evaluated separately	Focuses on adjacent-intersection coordination using a clean RL framework

E. Role of State, Action, and Reward Design

The literature consistently shows that the success of RL-based traffic control depends not only on the learning algorithm but also on the design of the state, action, and reward spaces. A weak state representation may hide critical traffic conditions from the controller. An overly large or poorly defined action space may slow learning or generate unstable signal behavior. A narrow reward function may improve one metric while degrading another.

In multi-node traffic control, these design choices become even more important. The state must reflect both local demand and neighboring conditions. The action space must allow coordinated yet practical signal transitions. The reward must capture corridor-level objectives such as queue stability, progression, delay reduction, and spillback prevention. If any of these components is designed poorly, the learned policy may not reflect meaningful coordination even if the algorithm converges.

F. Research Gap

The review of prior work suggests several gaps that motivate the present study. First, a substantial portion of the existing literature remains centered on single intersections, limiting its applicability to corridor-level traffic management. Second, many multi-intersection studies either assume more complex network settings than necessary for initial research or adopt architectures that are difficult to interpret in a compact academic study. Third, fixed numerical improvement claims often vary widely across studies because network topology, traffic demand, training duration, and detector placement are not standardized. This makes it difficult to isolate the effect of coordination itself.

Therefore, there is value in developing a research-based, simulation-driven paper that emphasizes methodological clarity over unsupported result claims. A DQN-based coordinated control framework for two adjacent intersections provides an effective middle ground between oversimplified single-node studies and highly complex network-scale systems. It allows investigation of state coordination, action design, and evaluation planning in a setting that remains academically meaningful and practically manageable.

III. PROBLEM STATEMENT AND RESEARCH OBJECTIVES

A. Problem Statement

Closely spaced urban intersections exhibit strong operational interdependence because traffic departing from one node influences queue formation, arrival patterns, and service efficiency at the next node. Fixed-time controllers and independently optimized signal strategies do not explicitly account for these interactions, which can lead to queue spillback, unstable progression, and corridor inefficiency. Existing traffic RL studies often focus on isolated intersections or treat coordination only implicitly. There is a need for a structured research framework that explicitly models adjacent-intersection coordination using Deep Reinforcement Learning and evaluates it in a reproducible simulation setting.

B. Research Objectives

The objectives of this work are as follows:

- 1) To model two adjacent signalized intersections as a coordinated traffic control problem in a SUMO-based simulation environment.
- 2) To formulate corridor traffic signal coordination as a Markov Decision Process suitable for DQN learning.

- 3) To design a joint state representation that captures traffic conditions and active phases across neighboring intersections.
- 4) To define a practical coordinated action space for corridor-aware traffic signal switching.
- 5) To establish a reward formulation that encourages queue reduction, delay minimization, and progression across adjacent intersections.
- 6) To provide an evaluation framework for comparing coordinated DQN control with conventional fixed-time and independent local control strategies.

IV. PROPOSED METHODOLOGY

A. Overall Framework

The proposed framework integrates a microscopic traffic simulator, a real-time control interface, and a Deep Reinforcement Learning agent. SUMO simulates the motion of vehicles

through a two-intersection corridor, while TraCI enables the external controller to access detector readings and modify signal phases during runtime. The DQN agent receives traffic

observations, selects actions according to the current policy,

and updates its internal value function based on reward feed-back.

The methodology is intentionally structured so that the study can remain research-based without making unsupported claims about performance magnitude. Its primary purpose is to formalize the coordinated control problem, define a reproducible architecture, and establish how comparative evaluation should be performed.

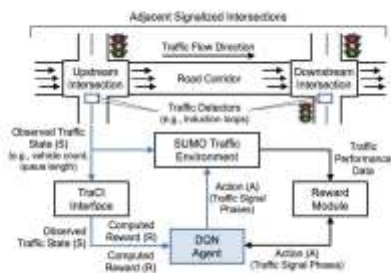


Fig. 1. Overall architecture of the proposed DQN-based coordinated traffic flow control framework for adjacent intersections.

B. Corridor Representation

The traffic network considered in this study consists of two adjacent intersections placed along a common corridor. Vehicles traverse the upstream node and then continue toward the downstream node. The link between the intersections serves as a key coordination element because its occupancy and queue buildup influence whether released vehicles can progress smoothly or become blocked. Such a setup reflects many practical urban road segments where adjacent signals are separated by short distances.

Each intersection contains detector-equipped approaches. Lane-area detectors or equivalent sensing elements can be positioned near the stop line and, if needed, at multiple upstream distances to observe queue growth. These detector values serve as the basis of the state representation. Additional state information may include current signal phase, elapsed phase time, or simple occupancy indicators.

C. MDP Formulation

The coordinated control problem is modeled as a Markov Decision Process (S, A, P, R, γ) . At each decision interval, the controller observes the system state, selects a coordinated action, receives a reward, and transitions to a new state according to the traffic dynamics generated by SUMO.

The state vector at time t can be expressed as:

$$S_t = [Q^t, Q^t, O^t, O^t, \phi^t, \phi^t]$$

where Q^t and Q^t denote queue-related observations for the two intersections, O^t and O^t denote occupancy or traffic load indicators, and ϕ^t, ϕ^t represent the current active signal

phases. Depending on the desired model granularity, each queue variable may itself consist of multiple detector values from different approaches.

This joint state representation is important because it allows the controller to capture interdependence between nodes rather than treating them independently. A downstream congested condition should influence upstream signal release decisions, and this relationship can only be learned if both nodes are visible to the agent.

D. Action Design

The action space must allow meaningful coordination while remaining compact enough for tractable learning. A practical coordinated action set can include the following possibilities:

- Maintain the current phase configuration at both intersections.
- Switch the upstream intersection while keeping the down-stream unchanged.
- Switch the downstream intersection while keeping the upstream unchanged.
- Switch both intersections according to a coordinated transition rule.

This action design is intentionally simple. It enables corridor-aware control without creating an excessively large combinational action space. It also reflects practical signal operation, where abrupt or highly granular action alternatives may be undesirable.

E. Reward Formulation

The reward function should reflect corridor-level traffic efficiency rather than the performance of a single node alone. A generic reward can be defined as:



International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

$$R_t = -\alpha Q_t - \beta D_t + \delta T_t$$

where Q_t denotes aggregate queue conditions, D_t denotes delay-related penalty, and T_t denotes throughput-related reward. The weighting coefficients α , β , and δ control the influence of each term.

A corridor-specific spillback penalty can be introduced when queue overflow on the interconnecting link is of interest:

$$R_t = -\alpha Q_t - \beta D_t + \delta T_t - \lambda S_t$$

where S_t represents spillback severity or downstream block-age. This reward encourages the controller to consider whether upstream discharge is likely to worsen downstream congestion. The reward function in the present paper is kept intention-ally general because the emphasis is on research design and methodological clarity. The exact coefficient values can be tuned during implementation and sensitivity analysis.

F. Deep Q-Network Design

The DQN controller approximates the action-value function $Q(s, a)$ using a feed-forward neural network. The input layer receives the coordinated state vector. One or more hidden layers learn latent features related to corridor traffic behavior, and the output layer produces Q-values for the defined coordinated actions.

The training process uses an experience replay buffer to store transitions (s_t, a_t, r_t, s_{t+1}) . Random minibatch sampling from this buffer reduces temporal correlation in training data. A target network is maintained to provide stable Q-value targets during updates. Exploration is managed using an ϵ -greedy policy so that the agent can balance trial-and-error learning with exploitation of the currently learned strategy.

The choice of DQN is justified by its suitability for moderate-scale, discrete-action control problems such as adjacent-intersection coordination. While more advanced algorithms exist, DQN remains an interpretable and widely accepted baseline for traffic signal learning studies.

G. Training Workflow

The training workflow consists of repeated simulation episodes. At the start of each episode, the traffic network is initialized and route demand is loaded. At every control step, the following operations are performed:

- 1) Read detector data and phase information from both intersections through TraCI.
- 2) Construct the joint state vector.
- 3) Select a coordinated action using the DQN policy.
- 4) Apply the action to the signal controllers in SUMO.
- 5) Advance the simulation by the chosen control interval.
- 6) Compute reward from the updated traffic condition.
- 7) Store the transition in replay memory and update the Q-network.

This process continues until the episode ends. After training, the learned policy can be evaluated under unseen test episodes for comparative analysis.

V. EXPERIMENTAL SETUP AND EVALUATION STRATEGY

A. Simulation Environment

The evaluation environment is based on SUMO due to its flexibility in modeling microscopic traffic behavior and its support for programmatic signal control through TraCI. The corridor is composed of two adjacent intersections connected by a road segment through which vehicle platoons propagate. Traffic demand may be defined using route files that represent balanced or asymmetric flow conditions, depending on the intended experiment.

A fixed episode duration can be selected to capture both transient and stabilized traffic behavior. The control interval should be long enough to prevent unrealistic oscillatory switching and short enough to allow meaningful adaptation to queue changes.

B. Baseline Strategies

To make the study methodologically meaningful, the coordinated DQN controller should be compared with at least two baseline strategies:

- **Fixed-Time Control:** both intersections operate with pre-determined signal phases and durations.
- **Independent Local Control:** each intersection follows a local decision strategy without explicit coordination.

This comparison makes it possible to isolate the contribution of coordination as distinct from adaptation alone.

C. Evaluation Metrics

The proposed framework can be evaluated using widely used traffic-performance metrics:

- Total delay across vehicles in the corridor.
- Average queue length at the intersections.
- Maximum queue length as an indicator of severe congestion.
- Throughput representing vehicles discharged successfully.
- Average travel time through the corridor.
- Tail-end travel metrics such as upper-percentile travel time. These metrics provide a balanced view of both average and worst-case traffic behavior. If desired, future work may also include environmental indicators such as fuel use or emissions.

D. Analysis Plan

Because this paper is research-based and does not seek to claim unverified numerical outcomes, the analysis plan focuses on how results should be interpreted once experiments are conducted. The comparison should examine whether coordinated DQN control:

- improves corridor progression,
- stabilizes downstream queues,
- reduces the likelihood of spillback,
- balances service between adjacent intersections,
- adapts more effectively to changing demand than static or local strategies.

The final interpretation should emphasize policy behavior, coordination logic, and qualitative traffic



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

dynamics in addition to raw metric values.

VI. DISCUSSION FRAMEWORK

The expected discussion of experimental outcomes should go beyond simply identifying which method performs better. In coordinated traffic control, the key question is how the learned policy uses cross-node information to manage corridor flow. If the DQN controller successfully accounts for down-stream occupancy before releasing upstream traffic, then the policy can be interpreted as learning a form of progression-aware coordination. Such behavior would indicate that the joint state representation is meaningful and that the action design is expressive enough for corridor control.

Another important aspect of the discussion is the relationship between local and global objectives. A strategy that minimizes queue at one intersection alone may still create unstable flow if the neighboring node becomes overloaded. Therefore, even when the coordinated controller does not minimize every local metric independently, it may still produce better corridor-level stability. This trade-off is central to multi-node traffic control and should be analyzed carefully.

The discussion should also address training stability and model practicality. DQN-based control is computationally more demanding than fixed-time logic, but it offers the potential for adaptability and policy improvement without manual retuning. The paper should therefore consider whether the increased modeling complexity is justified by the qualitative advantages of coordination, especially in research and smart-city contexts.

Finally, the study should recognize that a two-node corridor is not the full complexity of an urban network. However, this limitation is also a strength because it allows the coordination problem to be studied clearly and transparently. A well-constructed adjacent-intersection study can provide a strong conceptual basis for later extension toward larger networks, multi-agent systems, or transfer learning across road segments.

VII. CONCLUSION AND FUTURE SCOPE

This paper has presented a research-based framework for DQN-driven coordinated traffic flow control at adjacent intersections. The study addresses a meaningful gap between isolated single-intersection optimization and large-scale network control by focusing on two connected nodes whose operations directly influence one another. Using SUMO and TraCI, the framework supports simulation-based experimentation in a controlled and reproducible setting. The formulation includes a coordinated state representation, a practical joint action space, and a corridor-oriented reward structure suitable for Deep Reinforcement Learning.

The primary contribution of this work lies in its methodological clarity. Rather than making unsupported numerical claims, the paper establishes how coordinated control for adjacent intersections can be modeled, trained, and evaluated academically. This makes it suitable as a research-oriented M.Tech contribution and a foundation for future publication-oriented experimentation.

Several extensions are possible in future work. The two-node framework can be expanded to three or more

connected intersections, allowing study of larger arterial corridors. A multi-agent variant could assign one learning agent to each node while enabling coordination through information sharing. Additional research can explore alternative reward designs, fairness-aware control, emission-sensitive objectives, and transfer learning between network layouts. Real-world deployment studies can follow later, but the current framework already provides a strong simulation-based basis for advancing coordinated traffic signal research.

REFERENCES

- [1] F. V. Webster, "Traffic signal settings," Road Research Technical Paper 39, Road Research Laboratory, UK, 1958.
- [2] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2043–2067, 2003.
- [3] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3–4, pp. 279–292, 1992.
- [4] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [5] L. A. Prashanth and S. Bhatnagar, "Reinforcement learning with average cost for adaptive traffic signal control," in *Proceedings of the 18th IFAC World Congress*, 2011, pp. 8855–8860.
- [6] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, "Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers," *Transportation Research Record*, vol. 2324, pp. 1–10, 2013.
- [7] G. Zheng, Y. Xiong, X. Zhu, D. Zhao, H. Zhao, and F.-Y. Wang, "Learning phase competition for traffic signal control," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 1243–1250.
- [8] H. Wei, G. Zheng, H. Yao, and Z. Li, "PressLight: Learning max pressure control to coordinate traffic signals in arterial network," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, pp. 1290–1298.
- [9] E. Wei, S. Xu, M. Zhang, D. Zhao, and Z. Li, "CoLight: Learning network-level cooperation for traffic signal control," in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management Workshops*, 2019.
- [10] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO—Simulation of Urban Mobility," *International Journal On Advances in Systems and Measurements*, vol. 5, no. 3–4, pp. 128–138, 2012.
- [11] Sample reference style adapted from the user-provided conference paper format on academic two-column manuscript preparation.