Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/AN15.20250316
Published: February 18, 2025

Dynamic Traffic Signal Control with Reinforcement Learning: Improving Throughput and Reducing Wait Times in Urban Networks

Thomas Wiese
Assistant Professor of Analytics
School of Business
SUNY Empire State University
New York
thomas.wiese@sunyempire.edu

Abstract

This paper presents a novel approach to optimizing urban traffic flow using Reinforcement Learning (RL). Traditional traffic control systems, such as fixed-timing signals, often fail to adapt to real-time traffic conditions, leading to congestion, delays, and increased fuel consumption. By leveraging real-time data from the New York City Department of Transportation (NYC DOT) and modeling traffic signal control as an RL problem, we hypothesize that RL can outperform fixed-timing systems in reducing vehicle wait times, traffic congestion, and overall fuel consumption. Using the Simulation of Urban Mobility (SUMO) platform, we simulate traffic flows and compare the RL system to a rule-based, fixed-timing system. The RL system reduces average queue lengths by 25%, decreases vehicle wait times by 18%, improves traffic throughput by 12%, and reduces fuel consumption by 15%. These results suggest that RL offers a scalable and adaptive solution for managing urban traffic more efficiently.

Keywords
Keywords: Reinforcement Learning, Traffic Optimization, Operations Management, Process Optimization, Machine Learning Applications.

1. Introduction
Urban traffic congestion is a significant challenge faced by cities worldwide, leading to increased travel delays, fuel consumption, and environmental pollution (Papageorgiou et al., 2003). According to the United Nations (2018), by 2050, 68% of the world's population is expected to live in urban areas, intensifying the demand on infrastructure, particularly transportation networks. Congestion not only delays commuters but also results in substantial economic losses and environmental impacts. Schrank et al. (2015) estimate that traffic congestion in the United States alone costs over $160 billion annually in lost productivity and wasted fuel.

Traditional traffic management systems, such as fixed-timing signals, operate on predefined schedules and fail to adapt to real-time traffic conditions. This lack of adaptability results in inefficient traffic management, particularly during peak hours or unexpected traffic surges (Papageorgiou et al., 2003). Advanced systems like SCOOT (Split Cycle Offset Optimization Technique) and SCATS (Sydney Coordinated Adaptive Traffic System) have sought to address these shortcomings by adjusting signal timings in response to real-time traffic data (Hunt et al., 1982; Lowrie, 1992). However, these systems remain limited by rule-based logic and require extensive manual calibration, restricting their effectiveness in complex urban environments.

Machine learning techniques, specifically predictive models that forecast traffic conditions based on historical data, have emerged as potential solutions to congestion (Vlahogianni et al., 2015). However, these models are constrained by their inability to adapt to real-time conditions and unforeseen events such as accidents or sudden traffic surges. Reinforcement Learning (RL) represents a promising alternative, offering an adaptive approach that allows agents to learn from interactions with the environment and optimize their actions in real time (Sutton & Barto, 2018). RL agents can dynamically adjust traffic signals to optimize flow and minimize delays based on real-time traffic data (Wei et al., 2019).

While RL offers significant advantages, implementing it in real-world traffic systems presents several challenges. Computational complexity is a major concern, as RL algorithms often require substantial computational resources to process real-time data and make prompt decisions (Li, 2018). Data privacy and security issues also arise when collecting and transmitting real-time traffic data, necessitating robust data protection measures (Hahn et al., 2019). Additionally, infrastructure limitations, such as outdated traffic signal hardware and lack of standardized communication protocols, can hinder the deployment of RL systems in existing urban networks (Dresner & Stone, 2008). Addressing these challenges is crucial for the successful integration of RL into urban traffic management.

1.1 Objectives

This paper explores the application of RL to traffic signal optimization in New York City. By leveraging real-time data from the NYC DOT, we test the hypothesis that an RL-based traffic control system will outperform traditional fixed-timing traffic systems in reducing congestion and improving traffic flow. Additionally, we assess the environmental benefits of RL in terms of fuel consumption and emissions reduction.

Hypotheses

Null Hypothesis (H_0): There is no statistically significant difference between the performance of an RL-based traffic signal control system and a traditional fixed-timing system in terms of reducing vehicle wait times, improving traffic throughput, and minimizing fuel consumption.

Alternative Hypothesis (H_1): The RL-based traffic signal control system significantly outperforms a traditional fixed-timing system by reducing vehicle wait times, improving traffic throughput, and minimizing fuel consumption.

Research Questions

Can RL-based traffic signal control significantly reduce vehicle wait times compared to traditional rule-based systems in a highly congested urban environment like New York City? What are the effects of RL-based signal control on key performance metrics such as traffic throughput and fuel consumption?

2. Literature Review

Urban traffic management has been a focal point of research for decades, with various systems developed to address congestion, environmental impact, and economic efficiency. Traditional fixed-timing traffic signals, while simple, are increasingly inadequate for managing the complex and dynamic conditions in modern urban environments (Papageorgiou et al., 2003). Their limitations are particularly evident during peak traffic hours when their inability to adapt results in long vehicle queues and increased wait times.

Adaptive traffic control systems (ATCS) like SCOOT and SCATS were introduced to overcome these limitations by leveraging real-time data from traffic sensors to adjust signal timings dynamically (Hunt et al., 1982; Lowrie, 1992). SCOOT adjusts the offset between green lights at different intersections based on real-time data, optimizing traffic flow. SCATS operates similarly, using sensor data to optimize signal timings at individual intersections. While these systems improve upon fixed-timing signals, they rely on rule-based algorithms that require manual calibration, limiting their responsiveness to sudden and unpredictable changes in traffic patterns (Lowrie, 1992).

Machine learning approaches, particularly supervised learning models such as decision trees, support vector machines, and neural networks, have been explored as potential solutions. These models can predict future traffic patterns based on historical data, allowing traffic managers to anticipate congestion and adjust signals accordingly (Vlahogianni et al., 2015). However, their reliance on historical data limits adaptability, especially when faced with unexpected events like accidents or sudden traffic surges.

Reinforcement Learning represents a more flexible approach to traffic management. Unlike supervised learning models, RL agents learn by interacting with their environment, continuously improving their decision-making policies based on real-time feedback (Sutton & Barto, 2018). This makes RL well-suited to the dynamic and unpredictable nature of urban traffic systems. Wei et al. (2019) applied RL to traffic signal control and demonstrated that RL-based systems could reduce average vehicle delays by up to 20% compared to traditional fixed-timing systems.

Deep Reinforcement Learning (DRL) has further expanded RL's capabilities in traffic management. DRL combines RL with deep neural networks, enabling agents to handle high-dimensional state spaces and learn complex policies (Mnih et al., 2015). Gao et al. (2020) proposed a DRL-based traffic signal control method that improved traffic flow in simulated urban networks by approximately 15%. Similarly, Mousavi et al. (2017) utilized deep policy-gradient methods for traffic light control, showing improved adaptability to changing traffic conditions with delay reductions of around 10%.
Multi-Agent Reinforcement Learning (MARL) extends RL to environments with multiple interacting agents, which is particularly relevant for traffic systems where multiple intersections must coordinate to optimize overall network performance. Chu et al. (2019) applied MARL to large-scale traffic signal control, demonstrating enhanced scalability and efficiency over single-agent systems, achieving delay reductions of up to 25% in complex networks.

Despite these advancements, challenges remain in deploying RL-based systems in real-world settings. Computational requirements for training and inference can be substantial, and ensuring the stability and convergence of learning algorithms in dynamic environments is an ongoing research topic (Oroojlooyjadid & Hajinezhad, 2019). Moreover, integrating RL systems with existing traffic infrastructure requires addressing compatibility and standardization issues (Hahn et al., 2019).

3. Methodology

This study employs real-time traffic data provided by the New York City Department of Transportation (NYC DOT), collected through Automated Traffic Recorders (ATR) located at major intersections. The dataset includes variables such as time, location, vehicle count, and sensor ID, with traffic volumes reported at hourly intervals. To simulate real-time traffic conditions in the Simulation of Urban Mobility (SUMO) platform, the hourly data were interpolated to minute-level intervals.

The interpolation involved distributing the hourly traffic volumes uniformly across each hour, recognizing that this simplification may not capture short-term fluctuations but offers a reasonable approximation for the simulation (Little & Rubin, 2019). To address potential inaccuracies, stochastic variations were introduced by adding random noise drawn from a normal distribution with a mean of zero and a standard deviation equal to 10% of the hourly volume.

Additional feature engineering was performed to refine the dataset. Time-of-day indicators were created to account for daily traffic patterns, distinguishing between peak and off-peak hours. Rolling averages over 15-minute intervals were computed to capture short-term trends in traffic flow. Traffic volumes were normalized using z-score normalization to ensure consistent inputs for the reinforcement learning (RL) agent and to prevent any single feature from dominating the learning process.

The core of the RL framework involves modeling each traffic signal as an RL agent. The agent's objective is to minimize vehicle wait times and reduce overall congestion by dynamically adjusting signal timings based on real-time traffic conditions. The agent learns through trial and error, receiving feedback in the form of rewards based on its actions.
The state space S is defined as:

$$S = \{ Q\_i, P\_i, T\_i, D \}$$

where Qi represents the queue length for each lane i, Pi denotes the current signal phase for each lane (e.g., red, yellow, green), Ti is the elapsed time since the last signal change for lane i, and D is a discrete time-of-day indicator (e.g., peak, off-peak).

The action space A consists of the possible actions the agent can take at each time step:

$$A = \{ C, E\_g, S\_g \}$$

where C represents changing the signal phase (e.g., from green to yellow), Eg is extending the green light duration by Δt seconds, and Sg is shortening the green light duration by Δt seconds.

The reward function R(s,a) is designed to encourage actions that minimize vehicle wait times and reduce queue lengths:

$$R(s,a) = -[\alpha \times \Sigma Q\_i + \beta \times \Sigma W\_i]$$

In this equation, Qi is the queue length at lane i, Wi is the total wait time of vehicles in lane i, and α and β are weighting factors set to 1 to balance queue length and wait time. The summations Σ are over all lanes i.

The RL agent was trained using a Deep Q-Network (DQN) proposed by Mnih et al. (2015). Key components of the training algorithm include experience replay, which stores past experiences (s,a,r,s′) to break correlations in the observation sequence, and a target network with delayed parameters to stabilize learning. An epsilon-greedy policy was employed to balance exploration and exploitation, with ϵ decreasing from 1 to 0.1 over 10,000 steps.

The hyperparameters used in the training are as follows:

Learning rate: 0.001

Discount factor γ: 0.95

Batch size: 64

Replay memory size: 100,000 experiences

Epsilon decay: ϵ decreases from 1 to 0.1 over 10,000 steps

The simulation was conducted using the SUMO platform (Krajzewicz et al., 2012). The simulated network included ten interconnected intersections in Midtown Manhattan, selected for their high traffic volumes and complexity. The road network was modeled based on actual map data, and traffic signal configurations matched those used by the NYC DOT.

The simulation parameters included a duration of 24 hours, covering both peak and off-peak periods, and a time step of one second per simulation step. Approximately 100,000 vehicles were simulated over the 24-hour period. The simulations were run on a workstation equipped with an Intel Core i9 processor and 32 GB of RAM.

Calibration was performed using historical traffic data to adjust vehicle acceleration and deceleration rates and route choice models, ensuring that the simulated traffic patterns closely matched real-world conditions. Validation was conducted by comparing simulated traffic volumes and speeds with actual measurements from the NYC DOT, achieving satisfactory accuracy.

Two systems were compared in this study: the RL-based system, where traffic signals are controlled by the trained RL agent dynamically adjusting in real-time, and the fixed-timing system, with traditional traffic signals operating on predefined cycle lengths and not adapting to real-time conditions.

The effectiveness of the RL and fixed-timing systems was evaluated using several performance metrics, including average queue length (mean number of vehicles waiting at intersections), total vehicle wait time (cumulative time spent by vehicles waiting at traffic signals), traffic throughput (total number of vehicles processed through each intersection during the simulation), and fuel consumption (estimated based on vehicle speed profiles and idle times, using standard fuel consumption models as per Barth & Boriboonsomsin, 2008).

A two-sample Z-test was performed for each performance metric to compare the RL and fixed-timing systems. The Z-test was chosen due to the large sample sizes and known variances from the simulation data. Assumptions of normality were verified using the Central Limit Theorem. A significance level of α=0.05α = 0.05α=0.05 was used. Effect sizes were calculated using Cohen's d to quantify the magnitude of the differences observed.

4. Data Collection

The data supporting this study's findings are available on data.gov. These data were derived from the following resources available in the public domain: https://catalog.data.gov/dataset/?organization=city-of-new-york&q=Traffic+Counts&tags=atr].

5. Results and Discussion

The results demonstrate that an RL-based traffic signal control system can improve urban traffic management compared to traditional fixed-timing systems. The RL system's ability to dynamically adjust signal timings in response to real-time traffic data allowed it to outperform the fixed-timing system, with statistically significant improvements across key performance metrics.

The observed improvements align with the literature's findings. Wei et al. (2019) achieved a 20% reduction in average vehicle delays using RL, while Gao et al. (2020) reported improvements of around 15% in traffic flow. This study's 25% reduction in average queue length and 18% decrease in vehicle wait times validate the effectiveness of RL in complex urban traffic networks.

While previous studies have demonstrated the potential of RL in traffic signal control, most have been conducted in simplified simulation environments or focused on small-scale networks. This study distinguishes itself by applying RL to a complex urban network using real-world data from NYC DOT. By simulating 10 interconnected intersections in a densely populated area of New York City, we assess the RL system's performance in a challenging setting, providing insights into its applicability in real-world urban environments.

Additionally, this study incorporates environmental metrics by evaluating fuel consumption, offering a broader perspective on the benefits of RL-based traffic control. The 15% reduction in fuel consumption offers economic benefits through cost savings and contributes to environmental sustainability by lowering greenhouse gas emissions.

Despite the positive results, the study has limitations. Its reliance on simulated data and environments means that real-world complexities may not be fully captured. Interpolating hourly traffic data to minute-level intervals introduces assumptions that could affect accuracy. Moreover, we compared the RL system only with a fixed-timing system and did not include advanced adaptive systems like SCOOT or SCATS in our experimental comparison. Future studies should consider including such systems to provide a more comprehensive evaluation.

Future work should involve field trials or real-world deployments. Such trials would validate the RL system in dynamic and uncontrolled environments, where additional factors—such as driver behavior and unpredictable incidents—cannot be fully replicated in simulations. Finally, as the comparisons here focus on two systems, ANOVA or other multi-group statistical methods could be employed in the future if additional adaptive traffic control systems (e.g., SCOOT or SCATS) are included. Such multi-system comparisons would further validate whether the observed RL benefits hold across different traffic management strategies, enhancing robustness and applicability in real-world scenarios.

5.1 Numerical Results

The RL-based system demonstrated improvements over the fixed-timing system across all performance metrics.
Average Queue Length
The RL system reduced the average queue length at intersections to 26 vehicles (SD = 6), compared to 35 vehicles (SD = 7) for the fixed-timing system, representing a 25% reduction. The Z-score for this comparison was 2.20, with a p-value of 0.028. Cohen's d was calculated to be 1.36, indicating a moderate effect size.
Total Vehicle Wait Time

The RL system recorded an average wait time of 3,700 vehicle-seconds (SD = 600), while the fixed-timing system had 4,500 vehicle-seconds (SD = 800), a reduction of 18%. The Z-score was 2.15, with a p-value of 0.032. The effect size (Cohen's d) was 1.11.

Traffic Throughput

The RL system processed 950 vehicles (SD = 110), whereas the fixed-timing system processed 850 vehicles (SD = 120), an improvement of 12%. The Z-score for traffic throughput was 1.96, with a p-value of 0.050. Cohen's d was calculated as 0.85.

Fuel Consumption

The RL system consumed 110 liters of fuel (SD = 15), while the fixed-timing system consumed 130 liters (SD = 20), representing a 15% reduction. The Z-score was 2.00, with a p-value of 0.046. The effect size (Cohen's d) was 1.11.

To further strengthen the statistical evaluation, 95% confidence intervals were calculated for all mean differences. They were derived from repeated simulation runs by applying standard formulas for differences between two independent groups. In all cases, the confidence intervals for the difference between the RL-based system and the fixed-timing system did not cross zero, reaffirming that the RL-based approach consistently outperformed the baseline. Below are the 95% confidence intervals for each metric:

Average Queue Length Difference (RL − Fixed-Timing): 95% CI = [−12,−6][-12, -6][−12,−6]
This interval indicates the RL system effectively reduces queue lengths by 6 to 12 vehicles on average.

Total Vehicle Wait Time Difference (RL − Fixed-Timing): 95% CI = [−1,100,−500][-1{,}100, -500][−1,100,−500] vehicle-seconds

This suggests the RL system reduces total vehicle wait time by 500 to 1,100 seconds.

Traffic Throughput Difference (RL − Fixed-Timing): 95% CI = [40,160][40, 160][40,160] vehicles
This interval shows the RL system processes 40 to 160 more vehicles than the fixed-timing approach.

Fuel Consumption Difference (RL − Fixed-Timing): 95% CI = [−27,−13][-27, -13][−27,−13] liters
The RL system uses 13 to 27 fewer liters of fuel compared to fixed-timing operation.

6. Conclusion

This study demonstrates that a Reinforcement Learning-based traffic signal control system can improve urban traffic flow compared to traditional fixed-timing systems. The RL system reduced average queue lengths by 25%, decreased vehicle wait times by 18%, improved traffic throughput by 12%, and reduced fuel consumption by 15%. These findings suggest that RL-based systems have the potential to provide scalable, adaptive solutions for managing traffic congestion in urban environments.

By leveraging real-time data and advanced machine learning algorithms, RL offers a promising approach to urban traffic management. The implementation of RL systems could lead to economic benefits, environmental sustainability, and improved quality of life for urban residents. As cities continue to grow, adopting innovative technologies like RL will be essential in addressing the complex challenges of urban transportation.

References

Aslani, M., Mesgari, M. S., and Wiering, M., Adaptive traffic signal control with actor-critic methods in a real-world traffic network with different traffic disruption events, *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 732–752, 2017. Available: https://doi.org/10.1016/j.trc.2017.09.020.

Barth, M., and Boriboonsomsin, K., Real-world carbon dioxide impacts of traffic congestion, *Transportation Research Record*, vol. 2058, no. 1, pp. 163–171, 2008. Available: https://doi.org/10.3141/2058-20.

Chu, T., Wang, J., Codecà, L., and Li, Z., Multi-agent deep reinforcement learning for large-scale traffic signal control, *IEEE Transactions on Intelligent Transportation Systems*, no. 3, pp. 1086–1095, March 2020. Available: https://doi.org/10.1109/TITS.2019.2901791.

Dresner, K., and Stone, P., A multiagent approach to autonomous intersection management, *Journal of Artificial Intelligence Research*, vol. 31, pp. 591–656, 2008. Available: https://doi.org/10.1613/jair.2502.

Gao, J., Shen, Y., Liu, J., Ito, M., and Shiratori, N., Adaptive traffic signal control: Deep reinforcement learning algorithm with experience replay and target network, *arXiv preprint*, arXiv:1705.02755, May 2017. Available: https://doi.org/10.48550/arXiv.1705.02755.

Hahn, D., Munir, A., and Behzadan, V., Security and privacy issues in intelligent transportation systems: Classification and challenges, *IEEE Intelligent Transportation Systems Magazine*, no. xx, pp. 1–1, 2019. Available: https://doi.org/10.1109/MITS.2019.2898973.

Hunt, P. B., Robertson, D. I., Bretherton, R. D., and Royle, M. C., The SCOOT on-line traffic signal optimisation technique, *Traffic Engineering & Control*, vol. 23, no. 4, pp. 190–192, 1982.

Krajzewicz, D., Erdmann, J., Behrisch, M., and Bieker, L., 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.

Li, Y., Deep reinforcement learning: An overview, *arXiv preprint*, arXiv:1810.06339, 2018.

Little, R. J., and Rubin, D. B., *Statistical Analysis with Missing Data*, 3rd Edition, Wiley, 2019.

Lowrie, P., SCATS, Sydney Co-Ordinated Adaptive Traffic System: A Traffic Responsive Method of Controlling Urban Traffic, *Roads and Traffic Authority NSW,* Darlinghurst, 1990.

Mnih, V., et al., Human-level control through deep reinforcement learning, *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

Mousavi, S. S., Schukat, M., and Howley, E., Traffic light control using deep policy-gradient and value-function-based reinforcement learning, *IET Intelligent Transport Systems*, September 20, 2017. Available: https://doi.org/10.1049/iet-its.2017.0153.

Oroojlooyjadid, A., and Hajinezhad, D., A review of cooperative multi-agent deep reinforcement learning, *arXiv preprint*, arXiv:1908.03963, 2019.

Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., and Wang, Y., Review of road traffic control strategies, *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2043–2067, 2003.

Schrank, D., Eisele, B., Lomax, T., and Bak, J., *2015 Urban Mobility Scorecard*, Texas A&M Transportation Institute, 2015.

Sutton, R. S., and Barto, A. G., *Reinforcement Learning: An Introduction*, 2nd Edition, MIT Press, 2018.

United Nations, *World Urbanization Prospects: The 2018 Revision*, Department of Economic and Social Affairs, Population Division, 2018.

Vlahogianni, E. I., Karlaftis, M. G., and Golias, J. C., Short-term traffic forecasting: Where we are and where we're going, *Transportation Research Part C: Emerging Technologies*, vol. 43, 2014. Available: https://doi.org/10.1016/j.trc.2014.01.005.

Wei, H., Zheng, G., Yao, H., and Li, Z., IntelliLight: A reinforcement learning approach for intelligent traffic light control, *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18)*, London, United Kingdom, August 19–23, 2018. Available: https://doi.org/10.1145/3219819.3220096.

Biography

Dr. Thomas Wiese is an Assistant Professor of Analytics and the Coordinator of the MBA program at SUNY Empire State University. He teaches courses in data analysis, artificial intelligence, machine learning, python programming, operations management, and statistics, among others. He has designed and launched graduate programs, including the Master of Science in Applied Analytics, Healthcare Analytics, Marketing Analytics, and the MBA in Business Analytics, equipping students with advanced analytical and leadership skills. Before his academic career, Dr. Wiese spent over a decade in industry, holding leadership positions in automation, analytics, and engineering within the energy, transportation, and healthcare sectors. His work focused on improving operational efficiency, developing data-driven solutions, and leading large-scale projects. Dr. Wiese holds certifications as a Project Management Professional (PMP) and a Certified Maintenance and Reliability Professional (CMRP), further underscoring his expertise in managing complex systems and ensuring operational reliability. His publication record includes significant contributions to applying analytics and artificial intelligence in critical infrastructure and healthcare, emphasizing innovative methodologies and practical implementation strategies. A U.S. Navy veteran, Dr. Wiese served as a nuclear reactor operator aboard submarines, where he honed his technical expertise and leadership abilities in high-stakes environments. He holds a Ph.D. in Information Science, an MBA, and a bachelor's degree in Nuclear Engineering. Dr. Wiese is dedicated to advancing the field of analytics through impactful research, rigorous teaching, and the development of future leaders in analytics and operations.