

ОПТИМІЗАЦІЯ ДОРОЖНЬОГО РУХУ ЗА ДОПОМОГОЮ ПІДКРІПЛЮВАЛЬНОГО НАВЧАННЯ

М. І. Вороновський, І. Ю. Юрчак

Національний університет “Львівська політехніка”,
кафедра систем автоматизованого проектування
E-mail: mykhailo.i.voronovskiy@lpnu.ua, iryna.yurchak@lpnu.ua

© Вороновський М. І., Юрчак І. Ю., 2025

Проаналізовано сучасні підходи до побудови інтелектуальних транспортних систем (ІТС) з метою оптимізації міського трафіку. Особливу увагу приділено алгоритмам безмодельного навчання з підкріпленням (Q-Learning і Deep Q-Learning), що застосовуються для керування сигналами світлофорів у динамічних умовах дорожнього руху. За результатами симуляцій у середовищі SUMO доведено, що впровадження таких алгоритмів забезпечує суттєве зменшення черг на перехрестях і підвищує пропускну здатність транспортної мережі. Зокрема, застосування Deep Q-Learning дає змогу ефективніше обробляти великі обсяги даних, зокрема інформацію від IoT-сенсорів та V2X-технологій, що сприяє реалістичній адаптації до змінних умов трафіку. Окреслено перспективи подальших досліджень щодо вдосконалення глибоких нейронних мереж і мультиагентних методів, які дають змогу покращити результати керування дорожнім рухом у міському середовищі та стати фундаментом для створення “розумних” транспортних інфраструктур.

Ключові слова: інтелектуальні транспортні системи, дорожній рух, Q-Learning, DQL, машинне навчання, оптимізація трафіку.

Вступ

Зростаюча урбанізація та стрімке зростання кількості автомобілів у містах створюють серйозні проблеми із заторами, безпекою дорожнього руху та екологічними наслідками. Традиційні системи та методи регулювання світлофорів зазвичай базуються на фіксованих або адаптивних часових інтервалах, що не завжди дає змогу врахувати реальний стан дорожнього руху. Тому інтелектуальні транспортні системи (ІТС) терміново необхідні для задоволення потреб сучасних транспортних мереж, що розвиваються [1].

Сучасні підходи до управління трафіком дедалі частіше звертаються до методів машинного навчання та підкріплювального навчання, що дають змогу створювати гнучкі та адаптивні алгоритми керування. Одним із перспективних напрямів є застосування алгоритму Q-learning для оптимізації роботи світлофорів. Прогнозування транспортних потоків є важливим компонентом інтелектуальних транспортних систем, оскільки воно дає змогу точно оцінити майбутній стан трафіку в певному регіоні протягом заданого періоду. Застосування алгоритму Q-learning дозволяє агенту, яким є система регулювання світлофорів, навчатися оптимальній стратегії керування шляхом постійної взаємодії з середовищем та коригування дій на основі отриманих винагород.

Ще одним важливим підходом є Deep Q-Learning (DQL), який поєднує переваги класичного Q-learning з можливостями глибоких нейронних мереж. DQL дає змогу моделювати складні залежності між станами та діями, забезпечуючи більш точне прогнозування майбутніх винагород навіть у високорозмірних просторах даних, що є особливо актуальним для систем регулювання дорожнього руху.

Застосування Q-learning та DQL відкриває нові можливості для створення адаптивних систем регулювання світлофорів, здатних оперативно реагувати на зміни в трафіку та оптимізувати потоки транспорту. Це сприяє підвищенню безпеки руху, зниженню заторів і покращенню загальної ефективності транспортної системи [2].

Огляд літературних джерел

У сучасному світі можна знайти низку систем, що використовують алгоритми Q-learning та Deep Q-Learning (DQL) для регулювання світлофорів. Розглянемо кілька конкретних систем, які вже запроваджені у світі та демонструють свою ефективність у реальних умовах.

CoLight

CoLight є передовою мультиагентською системою, що використовує Deep Q-Learning (DQL) для регулювання роботи світлофорів на рівні мережі перехресть. Основна ідея системи полягає в тому, що кожен світлофор розглядається як окремий агент, який приймає рішення на основі локальних даних про трафік, а також інформації, отриманої від сусідніх агентів. Використання Deep Q-Learning дає змогу агентам моделювати складні залежності між різними станами та діями у високорозмірному просторі. Глибокі нейронні мережі, що застосовуються в CoLight, забезпечують ефективне навчання навіть у випадках, коли традиційні методи підкріплювального навчання стикаються з проблемою “прокляття розмірності”.

Система інтегрує дані з кількох перехресть, що дає змогу враховувати вплив рішень одного агента на інших. Це сприяє більш узгодженому регулюванню мережі світлофорів, що призводить до покращення загального трафіку на рівні міста.

Дослідження [3] свідчить, що система CoLight значно покращує показники дорожнього руху: вона забезпечує більш рівномірний розподіл транспортного потоку, зменшує скупчення автомобілів на перехрестях та швидко реагує на зміни у трафіку, оптимізуючи режими роботи світлофорів у режимі реального часу.

IntelliLight

IntelliLight є сучасною системою керування світлофорами, розробленою на базі глибокого підкріплювального навчання (Deep Reinforcement Learning, DRL), яка дає змогу забезпечити оптимізацію транспортного потоку у режимі реального часу.

Ця система використовує дані з різноманітних сенсорів, таких як камери, детектори руху та інші джерела інформації про транспортні потоки. Ці дані обробляються для формування високорозмірного простору станів, що відображає поточну ситуацію на дорозі (кількість транспортних засобів, їх швидкість, розташування, час затримки тощо). Завдяки використанню глибоких нейронних мереж система здатна ефективно розпізнавати складні патерни в даних, завдяки чому можна точно оцінювати поточний стан дорожнього руху та прогнозувати його майбутній розвиток.

Використовуючи алгоритми глибокого підкріплювального навчання, IntelliLight навчається приймати оптимальні рішення для регулювання світлофорів. Система формулює задачу як марковський процес прийняття рішень (MDP), де кожен стан супроводжується відповідними діями (наприклад, зміна тривалості зеленого сигналу, перемикання фаз) та винагородою. Функція винагороди розробляється з метою зниження часу затримки транспортних засобів та оптимізації потоку, що сприяє більш ефективній роботі дорожньої мережі.

Система IntelliLight демонструє значні переваги порівняно з традиційними методами керування світлофорами, забезпечуючи більш плавний розподіл транспортного потоку, зменшення заторів та підвищення загальної безпеки дорожнього руху. Ці характеристики роблять її перспективним рішенням для впровадження у сучасні інтелектуальні транспортні системи [4].

PressLight

PressLight – це система, яка використовує глибоке підкріплювальне навчання для оптимізації режимів роботи світлофорів у реальному часі. Система інтегрує відеокамери та інші датчики для отримання зображень із перехрест. Отримані дані піддаються попередній обробці за допомогою згорткових нейронних мереж (CNN), що дає змогу визначити кількість транспортних засобів, їх розташування, швидкість руху та інші важливі параметри.

На основі відеопотоку система створює високорозмірне представлення поточного стану перехрестя, яке враховує як поточну завантаженість, так і прогнозовані зміни в трафіку. Це дає можливість більш точно оцінити ситуацію і забезпечити якісну адаптацію до змін у дорожньому потоці.

Використання Deep Q-Learning у системі PressLight передбачає представлення задачі керування світлофорами у вигляді марковського процесу прийняття рішень (MDP), де стан відображає актуальні умови на перехресті (наприклад, кількість автомобілів чи довжину черги), дія передбачає різні варіанти регулювання (зокрема, зміну тривалості зеленого сигналу), а функція винагороди спрямована на зменшення затримок і скорочення черг транспортних засобів. Завдяки застосуванню алгоритму Deep Q-Learning, що використовує replay метогу та таргет-мережі, система здатна послідовно вдосконалювати свою стратегію і підтримувати стабільне навчання навіть за умов високої мінливості дорожнього руху [5].

Постановка задачі

Розглянути алгоритми безмодельного навчання з підкріпленням Q-Learning та Deep Q-Learning (DQL). Спроектувати та реалізувати скрипти з використанням цих алгоритмів для оптимізації управління трафіком на регульованих перехрестях. Провести симуляційні експерименти, оцінити ефективність запропонованих методів шляхом аналізу метрик, таких як пропускна здатність перехрестя та кількість автомобілів, що стоять в черзі.

Огляд алгоритмів безмодельного навчання

Q-Learning

Q-Learning є одним із найпоширеніших алгоритмів навчання з підкріпленням, який використовується для прийняття оптимальних рішень в середовищах із динамічно змінними станами. Його основна ідея полягає у формуванні таблиці Q-значень (Q-таблиці), яка містить оцінки очікуваних винагород для різних станів і дій. Алгоритм навчається методом проб і помилок, оновлюючи значення у цій таблиці відповідно до отриманих винагород за обрані дії [6].

Процес навчання у Q-Learning базується на рівнянні Беллмана:

$$Q(s, a) = Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)], \quad (1)$$

де s – поточний стан; a – виконувана дія; r – отримана винагорода; s' – наступний стан; α – швидкість навчання (learning rate); γ – коефіцієнт дисконтування (discount factor), який визначає вплив майбутніх винагород.

Алгоритм повторює такі основні кроки:

1. Вибір дії a у поточному стані s за допомогою стратегії дослідження (наприклад, ϵ -жадібною стратегії, яка поєднує випадковий вибір і використання найкращих відомих Q-значень).
2. Виконання дії та отримання винагороди r .
3. Оновлення Q-значення відповідно до рівняння Беллмана.
4. Перехід до нового стану s' і повторення процесу.

Q-Learning широко використовується у різних сферах, включаючи оптимізацію транспортних систем. У випадку регулювання світлофорів алгоритм може навчатися змінювати сигнали так, щоб мінімізувати затримки транспорту та покращувати загальну пропускну здатність перехрестя. У поєднанні з симуляційними середовищами, такими як SUMO, Q-Learning дає змогу створювати ефективні адаптивні системи керування дорожнім рухом без необхідності детального програмування правил.

DQL

Deep Q-Learning є розширенням класичного Q-Learning, що використовує глибокі нейронні мережі для апроксимації Q-функції. Основна ідея полягає у подоланні проблеми експоненційного зростання Q-таблиці, що унеможливує її використання у складних середовищах із великим числом станів. У DQL замість збереження Q-значень для кожної можливої комбінації стану та дії використовується нейронна мережа, яка навчається прогнозувати значення Q-функції на основі вхідних параметрів.

DQL використовує кілька ключових технік для стабільності навчання. По-перше, це досвід повторного відтворення, що дає змогу уникнути кореляції між послідовними прикладами шляхом збереження історії досвіду агента у буфері та випадкового вибору мініпакетів для навчання. По-друге, застосовується окрема цільова нейромережа, яка оновлюється з меншою частотою, ніж основна мережева модель. Це допомагає зменшити нестабільність у процесі навчання, оскільки цільові значення змінюються поступово.

Навчання мережі здійснюється шляхом мінімізації різниці між передбаченим та фактичним Q-значенням, використовуючи середньоквадратичну помилку (MSE) або інші функції втрат.

Deep Q-Learning знаходить широке застосування в різних задачах, зокрема в управлінні дорожнім рухом. Завдяки здатності працювати з високорозмірними просторами станів, DQL є ефективним інструментом для оптимізації складних систем, де класичний Q-Learning є непридатним [7].

Проведення симуляцій

Для моделювання та перевірки ефективності запропонованих підходів використовується симуляційне середовище SUMO, де створюються відповідні TraCI-скрипти для інтеграції навчання з підкріпленням.

На рис. 1 наведено схему дорожнього руху, розроблену в середовищі SUMO. На схемі зображено регульоване перехрестя та детектори руху, що використовуються для фіксації руху на ділянці вздовж смуг. У реальних умовах цими детекторами можуть бути камери стеження за транспортними засобами.

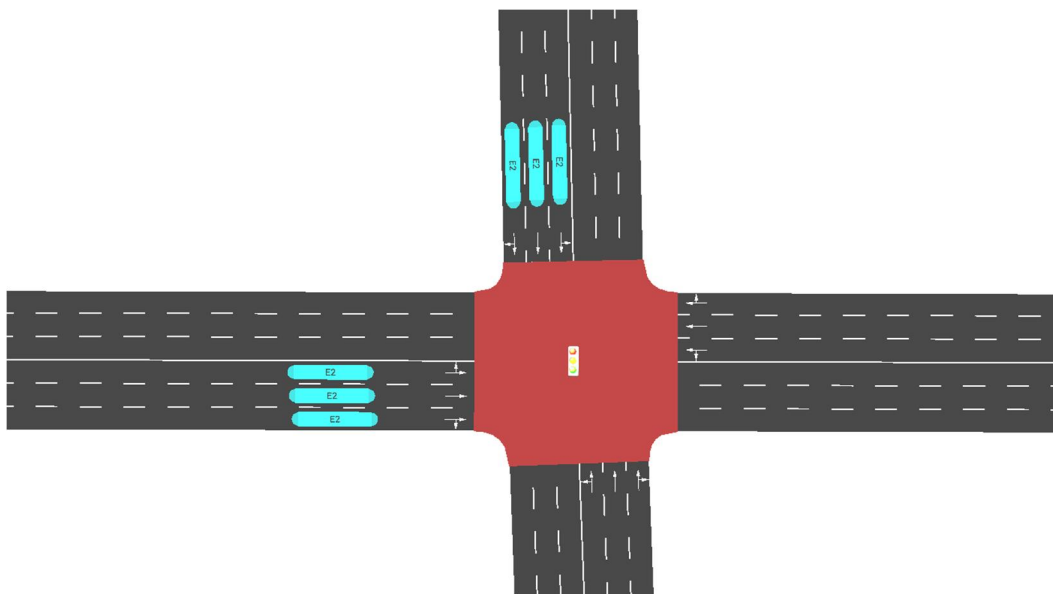


Рис. 1. Схема дорожнього руху

Для проведення симуляцій налаштовано транспортні потоки (рис. 2), що мають різну щільність появи, випадково генеруються на смугах одного напрямку і мають встановлену максимальну швидкість на рівні 50 км/год.

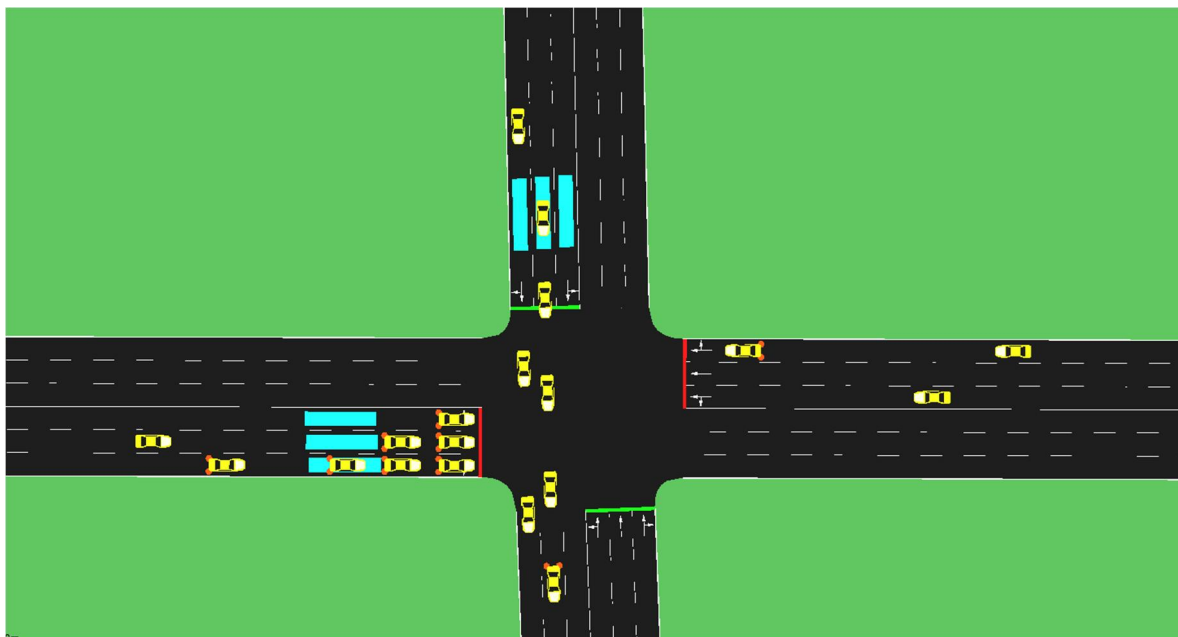


Рис. 2. Рух транспортних потоків

Використовуючи однакові налаштування, було проведено три випробування, кожне з яких мало обмеження в 10 000 кроків. На рис. 3 показано схему руху транспортних потоків, а в табл. 1 наведено стани світлофорів, що використовуються на перехресті.

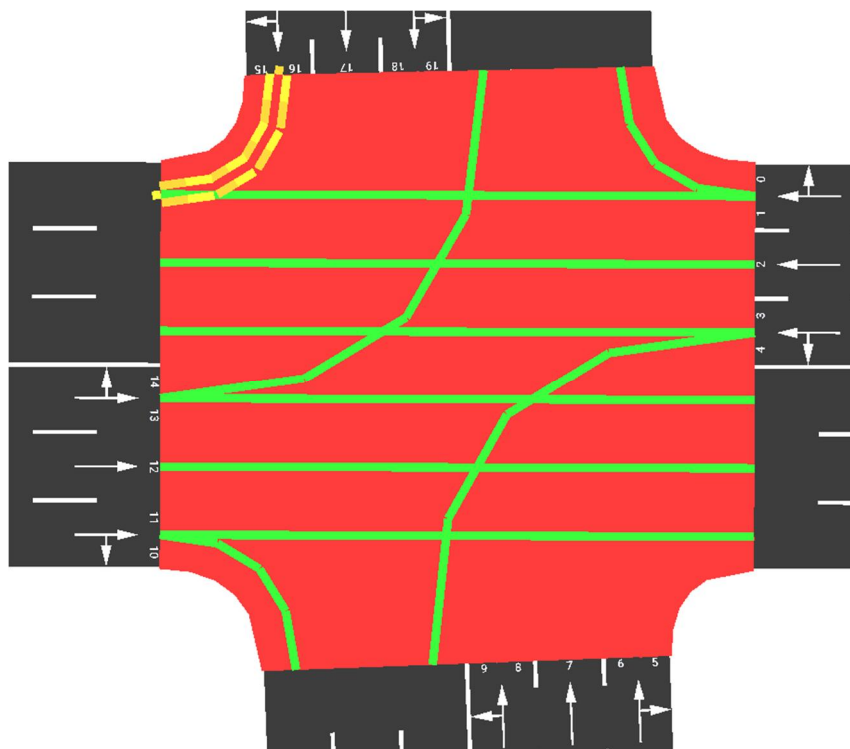


Рис. 3. Стани регульованого перехрестя

Таблиця 1

Стани регульованого перехрестя

Стан	Час, с	Сигнали
0	42	GGGGgrrrrGGGGgrrrr
1	3	Yuuuugrrrruuuugrrrr
2	42	rrrrGGGGgrrrrGGGGg
3	3	rrrruuuugrrrruuuu

Для збору статистики в усіх трьох симуляціях та реалізації Q-Learning і DQL алгоритмів було написано скрипти на мові програмування Python, що використовують бібліотеки підключення до SUMO, а також бібліотеку TensorFlow, що дає змогу легко створювати моделі машинного навчання.

Результати дослідження

Після проведення симуляцій в середовищі SUMO було отримано інформацію про те, як змінювалися транспортні черги на кожному кроці симуляції, а також загальну кількість транспортних засобів, що проїхали перехрестя.

На рис. 4 зображено зміну автомобільних черг, що утворювалися перед перехрестям.

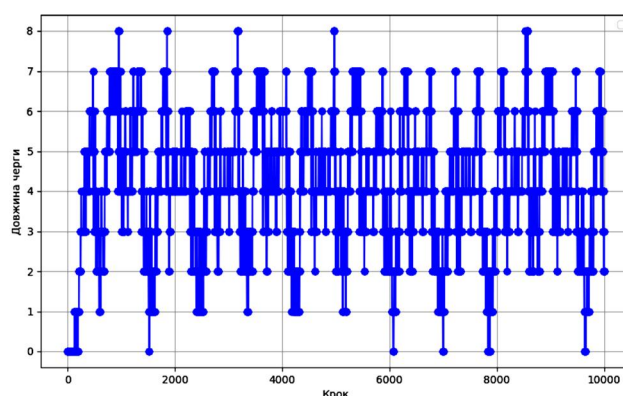


Рис. 4. Зміна довжини черги у стандартних умовах

На рис. 5 і 6 показано, як змінювалися автомобільні черги в симуляціях, де використовувалися алгоритми безмодельного навчання.

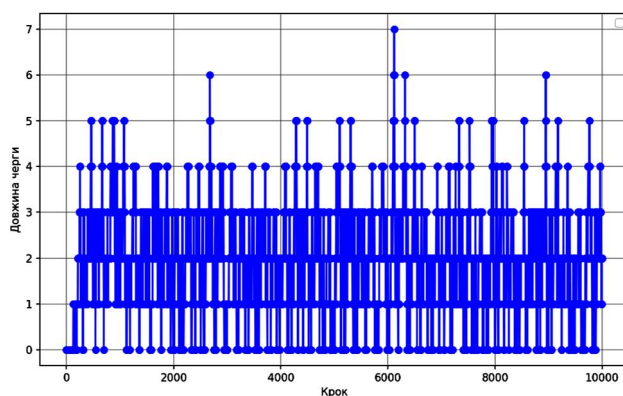


Рис. 5. Зміна довжини черги під час використання Q-Learning

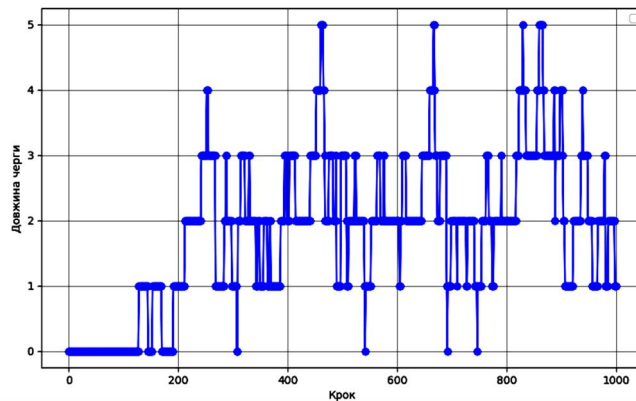


Рис. 6. Зміна довжини черги під час використання DQL

У табл. 2 наведені дані пропускної здатності перехрестя для всіх симуляцій.

Таблиця 2

Сумарна пропускна здатність після 10 000 кроків

Алгоритм	Транспортних засобів
Без алгоритму	914
Q-Learning	933
DQL	948

Висновки

Проведене дослідження підтверджує, що застосування Q-Learning та Deep Q-Learning у регулюванні світлофорів впливає на зменшення черг і збільшення пропускної здатності транспортної мережі.

У симуляціях пропускна спроможність перехрестя без інтелектуального керування становила 914 транспортних засобів, тоді як Q-Learning забезпечив збільшення до 933 одиниць, а Deep Q-Learning – до 948. Черги на світлофорах знизилися із застосуванням алгоритмів машинного навчання. Q-Learning показує скорочення черг на приблизно 25 %, тоді як симуляція з використанням DQL ще більше покращує цей показник.

Класичний Q-Learning ефективний у середовищах із порівняно невеликим простором станів, однак з розширенням дорожньої мережі й кількості вхідних даних його переваги помітно знижуються. Натомість підхід на основі глибокого підкріплювального навчання дає змогу враховувати різноманітні показники на під'їзних дорогах, оптимізувати режими керування світлофорами в умовах постійних змін трафіку та досягати кращих результатів за складних дорожніх сценаріїв.

Впровадження таких алгоритмів у міську інфраструктуру має потенціал не лише розвантажити дороги, а й покращити загальну безпеку, оскільки вони створюють передумови для більш оперативного реагування на різні транспортні ситуації. На базі цих підходів можна розбудовувати глобальні системи, що охоплюють пріоритезацію громадського та екстреного транспорту, виявлення інцидентів у реальному часі та комплексну оптимізацію трафіку загалом. Подальші дослідження можуть передбачати поєднання кількох типів алгоритмів підкріплювального навчання та доопрацювання глибоких нейронних мереж, що дозволить системам навчатися ще швидше й приймати ефективніші рішення у високорозмірних просторах станів. Отже, упровадження сучасних засобів машинного навчання у сфері керування дорожнім рухом сприяє підвищенню пропускної здатності, зниженню заторів та інтеграції міських транспортних систем у концепцію “розумного” міста майбутнього.

Список літератури

1. Majstorović Ž., Tišljarić L., Ivanjko E., Carić T. Urban Traffic Signal Control under Mixed Traffic Flows: Literature Review. *Applied Sciences*. 2023. 13(7). 4484. Doi: <https://doi.org/10.3390/app13074484>.
2. Ouyang Chen, Zhan Zhenfei, Lv Fengyao. A Comparative Study of Traffic Signal Control Based on Reinforcement Learning Algorithms. *World Electric Vehicle Journal*. 2024. 15. 246. 10.3390/wevj15060246.
3. Jin J. et al. CoLight: Learning Network-Level Cooperation for Traffic Signal Control. *Proceedings of the AAAI Conference on Artificial Intelligence*. Hua Wei, Nan Xu, Huichu Zhang, Guanjie Zheng, Xinshi Zang, Chacha Chen, Weinan Zhang, Yanmin Zhu, Kai Xu, and Zhenhui Li. 2019. CoLight: Learning Network-level Cooperation for Traffic Signal Control. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM '19)*. Association for Computing Machinery, New York, NY, USA, 2019. 1913–1922. Doi: <https://doi.org/10.1145/3357384.3357902>.
4. Wei Hua, Zheng Guanjie, Yao Huaxiu, Li Zhenhui. IntelliLight: A Reinforcement Learning Approach for Intelligent Traffic Light Control. 2018. 2496- 2505. Doi: 10.1145/3219819.3220096.
5. Rasheed Faizan, Yau Kok-Lim, Md. Noor Rafidah, Wu Celimuge, Low Yeh Ching. Deep Reinforcement Learning for Traffic Signal Control. *IEEE Access*. 8. 2020. 208016- 208044. Doi: 10.1109/ACCESS.2020.3034141.
6. Abid Ali Awan. An Introduction to Q-Learning: A Tutorial For Beginners. URL: <https://www.datacamp.com/tutorial/introduction-q-learning-beginner-tutorial>.
7. Samina Amin. Deep Q-Learning (DQN). URL: <https://medium.com/@samina.amin/deep-q-learning-dqn-71c109586bae>.

OPTIMIZING ROAD TRAFFIC THROUGH REINFORCEMENT LEARNING

M. I. Voronovskyi, I. Y. Yurchak

Lviv Polytechnic National University,
Department of Computer Design Systems

© Voronovskyi M. I., Yurchak I. Y., 2025

In the article, modern approaches to the development of Intelligent Transportation Systems (ITS) aimed at optimizing urban traffic are analyzed. Special attention is paid to model-free reinforcement learning algorithms (Q-Learning and Deep Q-Learning) used for controlling traffic lights in dynamic road traffic conditions. Simulation results in the SUMO environment have proven that implementing such algorithms significantly reduces intersection queues and increases the capacity of the transportation network. In particular, the use of Deep Q-Learning enables more efficient processing of large volumes of data, including information from IoT sensors and V2X technologies, fostering realistic adaptation to changing traffic conditions. The prospects for further research involve enhancing deep neural networks and multi-agent methods, which will improve traffic management outcomes in urban environments and lay the foundation for creating “smart” transportation infrastructures.

Keywords: Intelligent Transportation Systems; road traffic; Q-Learning; DQL; machine learning; traffic optimization.