

An Approach to Optimizing CO₂ Emissions in Traffic Control via Reinforcement Learning

Olexander Ryzhanskyi^{1,†}, Eduard Manziuk^{1,*}, Olexander Barmak^{1,†}, Iurii Krak^{2,3†}, Nebojsa Bacanin^{4†}

¹ Khmelnytskyi National University, 11, Instytuts'ka str., Khmelnytskyi, Ukraine

² Taras Shevchenko National University of Kyiv, 60 Volodymyrska str., Ukraine

³ Glushkov Cybernetics Institute, 40 Glushkov ave., Kyiv, Ukraine

⁴ Singidunum University, 11000 Belgrade, Serbia

Abstract

Automotive transport plays a key role in ensuring economic development but is accompanied by significant negative impacts on the environment, particularly in areas where vehicles are concentrated. This article presents an approach that uses reinforcement learning and accounts for traffic flow pressure to optimize the travel time of vehicles through road intersections with the aim of reducing CO₂ emissions. The proposed method is based on modern approaches to optimizing traffic light operations, but with an emphasis on ecological aspects. Experimental verification on the synthetic scenario SUMO GRID 4x4 demonstrates the efficiency of the developed algorithm. Comparative analysis shows that it outperforms other algorithms, such as MaxPressure and IDQN, in particular, it improves travel time and queue length by 33%, and reduces CO₂ emissions by 32-33%. The obtained results lay the foundation for further refinement and implementation of the proposed approach in real-world conditions.

Keywords

traffic signal control, reinforcement learning, reward modeling, pollutant emissions

1. Introduction

Road transport plays an important role in ensuring economic growth and social development. It is defined as a key component of the transportation system due to its objective advantages, which are reinforced by significant achievements in the transport infrastructure of the vast majority of countries. Road transport is also widely used and is a key priority in economic development. However, such circumstances are accompanied by significant pressure on the

IntelliTISIS'2024: 5th International Workshop on Intelligent Information Technologies and Systems of Information Security, March 28, 2024, Khmelnytskyi, Ukraine

* Corresponding author.

† These authors contributed equally.

✉ alex@eventcadence.com (O. Ryzhanskyi); eduard.em.km@gmail.com (E. Manziuk);

alexander.barmak@gmail.com (O. Barmak); yuri.krak@gmail.com (I. Krak); nbacanin@singidunum.ac.rs (N.

Bacanin)

ORCID: 0009-0000-4664-5195 (O. Ryzhanskyi); 0000-0002-7310-2126 (E. Manziuk); 0000-0003-0739-9678 (O. Barmak);

0000-0002-8043-0785 (I. Krak); 0000-0002-2062-924X (N. Bacanin)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

environment, especially in places where vehicles are heavily congested. Some of these places are large cities and transportation interchanges. Hence the problem of transport regulation. There is a wide range of problems for the solution of which information technologies of traffic improvement are [1]. Accordingly, the development of systems for the formation of climate-neutral and smart cities is of great importance given the current challenges associated with climate change and urban growth. Such systems are determined by the need for the following reasons:

- They help to achieve climate neutrality, which is a strategic objective under the European Green Pact [2]. Cities make a significant contribution to greenhouse gas emissions, and the implementation of systems aimed at optimizing resources and reducing environmental impact helps to solve this problem.
- Smart cities improve the quality of life for citizens. By optimizing traffic flows and reducing air and noise pollution, they contribute to the health and well-being of the population.
- Such systems help to reduce the urban ecological footprint and support sustainable development. They aim to reduce resource consumption and develop more efficient strategies for managing urban resources.
- The development of such systems promotes political coherence and citizen participation in decision-making. This is important for ensuring the effectiveness of strategies and achieving climate neutrality.
- Smart cities are being integrated into European and global strategies, contributing to the achievement of global climate goals and providing synergies with other initiatives.

To summarize, intelligent systems are a key element of digital transformation and innovation, enabling cities to use modern technologies more effectively to achieve climate neutrality and support sustainable development.

One approach to developing intelligent systems is to use reinforcement learning, which can be applied to a similar class of tasks [3]. This approach to artificial intelligence allows systems to learn from the data they receive and gain experience to make optimal decisions in real time. One of the key challenges is the efficient management of urban resources and infrastructure to ensure sustainability and efficiency. Reinforcement learning can analyze and optimize the operation of traffic lights, transportation systems, and other aspects aimed at reducing emissions and improving energy efficiency. Particularly important is the ability to train automated transport management systems, which helps to improve traffic flow and reduce traffic congestion. This has an impact on CO₂ emissions and improves air quality in cities, which in turn affects public health and overall quality of life.

Thus, the main contribution of the paper is the proposed approach using Reinforcement Learning to finding the optimal mode of vehicles passing through a traffic light-controlled crossroads according to the criterion of reducing CO₂ emissions.

The main contributions of the research include:

- A new approach to traffic signal control at road intersections using reinforcement learning that takes into account the environmental impact of traffic, in particular CO₂ emissions is proposed.

- The MPLightCO2 algorithm is developed, which is an extension of the existing MPLight approach with additional consideration of CO₂ emissions from vehicles queuing to enter and exit the crossroads. This makes it possible to optimize crossroads traffic modes in order to reduce environmental impact.
- It is proposed to take into account the "traffic flow pressure" metric to determine the efficiency of vehicle distribution in the crossroads network and improve throughput.
- Experimental verification and comparative analysis of the developed MPLightCO2 algorithm with other approaches, such as MaxPressure, MPLight, and IDQN, were carried out on the synthetic test scenario SUMO GRID 4x4.
- The results showed that MPLightCO2 outperforms existing approaches in terms of travel time, average queue length, and CO₂ emissions, demonstrating increased efficiency in both optimizing traffic flow and reducing its environmental impact, which allowed reducing queue length by 75-76% and reducing CO₂ emissions by 32-33%.

The article is structured as follows. In the Related Works section, we review current approaches to solving similar problems and formulate the purpose of the paper. The Methods and Materials section describes the crossroads control system, provides a formalization of its elements, presents an approach using traffic pressure, describes the DQN agent, the implementation of deep Q-learning, characterizes the SUMO GRID 4x4 synthetic test and approaches to assessing the quality of the solutions obtained. The Results and Discussion section analyzes the results of experimental testing on SUMO GRID 4x4, the quality indicators of the models, and compares them with other algorithms. The Conclusions and Future Work section summarizes the results of the study, outlines limitations and directions for further work.

2. Related Works and Basic Concepts of Approximate Dynamic Programming

A review of recent publications on the topic of the study showed that the modern reinforcement learning approach is actively used to solve such problems. Below is an overview of these publications. Modern development trends in the field of artificial intelligence are actively used to implement effective strategies for optimizing traffic flows in cities. The main goal is to reduce environmental impact through the development and application of various methods and technologies. Artificial intelligence plays a key role in this context, helping to create intelligent systems that ensure efficient traffic management. In particular, deep learning algorithms are used to develop smart traffic light control systems aimed at dynamic adaptation to changes in traffic flow in real time [4][6]. This not only minimizes stops and saves fuel, but also has a positive impact on emissions. It is important to study approaches that would meet all the requirements of AI reliability [7][8][9]. Much of the papers is aimed at optimizing traffic flows to increase the capacity of transportation routes.

Another approach is to predict and manage transportation demand, which is becoming another aspect where machine learning methods are used to accurately analyze passenger flow data and predict its changes at different times of the day [10][12]. This allows optimizing

the allocation of resources, reducing the number of empty flights and thus contributing to the reduction of CO₂ emissions.

The use of traffic monitoring and analysis systems based on data from sensors and cameras allows artificial intelligence algorithms to identify patterns and predict possible traffic congestion [13]. This opens up opportunities for taking effective measures to avoid congestion and, therefore, reduce the negative impact on the environment. The use of route optimization algorithms is also important in the context of reducing the environmental impact of transport [14] [15]. These algorithms take into account various aspects, such as minimizing the use of traffic lights and separating environmentally friendly routes.

The introduction of electric vehicles and autonomous cars is a key step in ensuring environmentally sustainable transportation. Research on the safety [16] and reliability of communication equipment is also important [17]. Artificial intelligence is used to optimize their movement and develop charging station infrastructure [18] [19]. An integrated approach to optimizing traffic flows in cities allows achieving traffic efficiency, reducing emissions and promoting sustainable urban transport. Such approaches can improve the conditions of movement of vehicles along the roads, reducing their delay, improving speed conditions, which ultimately has a positive impact on transport emissions, improving the environment. Methods for constructing neural networks are being developed and refined [20][22]. In pursuit of this goal, research is being conducted using modern reinforcement learning algorithms to optimize the performance of signal controllers in real time [23] [24]. In this approach, the state of the crossroads is determined by the parameters of vehicles (lane, speed, waiting time, queue position) and the actual signal (traffic permission). The main task of reinforcement learning, which is used in the form of an agent, is to optimize a strategy that adapts states to the signals. This approach has shown a potential reduction in vehicle delays of up to 73% compared to a fixed response time [25]. A method of multi-agent reinforcement learning known as cooperative dual Q-learning is used to solve the complexity of traffic signal synchronization in large-scale traffic control systems [26]. It uses independent dual Q-learning methods and an upper confidence bound policy to avoid overestimation problems that can occur in traditional algorithms. A new reward distribution mechanism and a local state distribution method are introduced to ensure stable and robust learning. Experiments on traffic flow scenarios show that the proposed system outperforms state-of-the-art decentralized algorithms on various traffic metrics.

Current research is overwhelmingly focused on the use of intelligent systems to optimize traffic flows and reduce congestion. Strategies should also actively seek to improve the environmental performance of transportation. Research focuses on traffic optimization rather than on the full range of environmental aspects and practical measures to reduce the environmental impact of transport. To achieve environmentally sustainable transport, it is important to consider not only traffic efficiency, but also improvements in air quality and overall environmental sustainability.

Thus, the aim of the study is to develop and test the effectiveness of an approach that uses reinforcement learning and traffic pressure to optimize vehicle travel times through road intersections. Particular emphasis is placed on traffic signal control to reduce CO₂ emissions. The research includes validation of the proposed approach on a synthetic grid4x4 test and further analysis of the results.

3. Methods and Materials

To systematize the traffic flow at a crossroads, we define typical scenarios, which we will call "states". At a signalized crossroads, there are incoming and outgoing roads, each of which may include one or more lanes. For each crossroads, we define a set of states ST , where each specific state $st \in ST$ is associated with a specific direction of traffic.

The states are considered as conflicting if they cannot be activated simultaneously due to traffic crossroads. At each stage, the signaling controller is responsible for establishing a specific combination of non-conflicting states in order to optimize the long-term objective function. For reinforcement learning-based controllers, the signalized crossroads environment can be modeled using the following description.

The state space is formed by the mapping of incoming traffic and active states. It is particularly important to consider the differences in research approaches, where some take into account high-resolution traffic detection technologies such as real-time observations of vehicle counts, waiting times, and average speeds, while others are limited to less informative data such as visibility of queue lengths or waiting times for the first vehicle. In terms of the envisioned sensing radius, some take a broad approach that covers all entrance roads, but a more realistic approach is to use a fixed sensing radius r_s . This may depend on the technological capability of the detection to provide reliable results, and take into account local features such as terrain, visibility limitations, or the presence of obstacles such as buildings or trees. In the context of alarm management, at each time step, the controller determines a set of non-conflicting states that are allowed to move, which is indicated by the green light. If there is a difference between the selected states and the active states, a mandatory yellow state is automatically entered for a fixed period of time. The assignment of yellow states is a constraint on the sequence of environmental control, not part of the action space.












| Road number at the crossroads | Lane 2 | Lane 1 | | |
|-------------------------------|---|---|---|---|
| 1 |  |  |  | |
| 2 | |  |  | |
| 3 |  |  |  | |
| 4 | |  |  |  |

Figure 1: Examples of typical states when crossing a crossroads.

The transition function is determined by the development of traffic after the signal is activated. This dynamics can be modeled according to a specific traffic model in the simulated environment or taken from real traffic progress data as part of a real-world implementation. The reward function typically uses the reduction in queue length as the sum of the respective scores of all incoming lanes and is expressed as an integral reward. This is effective in reducing congestion, but does not always normalize the benefits of signal optimization over the travel time of a particular route. Therefore, other reward functions are used, such as total delays, crossroads delays, crossroads waiting time, traffic volume, and others.

However, for proper systematization of traffic at the crossroads, it is necessary to take into account the action space. At each time step, the controller selects a set of non-conflicting states that receive permission to move, which is indicated by turning on the green light. If the selected states differ from the current active states, a mandatory yellow state is automatically entered for a predefined period of time. The assignment of yellow states is a constraint imposed on the sequence of environmental control, not part of the action space.

The consistency of the defined action space and the choice of optimal states determines the efficiency and safety of traffic at the crossroads. Even taking into account the different approaches to traffic detection, it is important to consider that parameters such as the prescribed sensing radius can affect the accuracy and reliability of the data obtained. Let's formalize the elements of the control system at the crossroads. To do this, let's define the main parameters taking into account the states.

Let us denote the set of states as $St = \{st_1, st_2, \dots, st_n\}$, s_i is the specific state associated with the direction of traffic; $n \in N$ - the total number of states at the crossroads, determined by the number of specific directions of traffic that are selected for modeling or need to be taken into account.

The number of states can be determined by a ratio that takes into account the number of possible options for each direction of traffic on each road. If there are Rd input roads, each of which has dr_i possible directions of movement. Then the total number of states N is calculated by formula:

$$N = \prod_{i=1}^{Rd} dr_i. \quad (1)$$

This formula presents the product of the number of possible directions of movement on each input road. Thus, the total number of unique states in the state space is determined N^* .

Let us define the action domain $Act = \{act_1, act_2, \dots, act_m\}$, where act_j is the specific action of the controller at each time step; $m \in M$ is the total number of possible controller actions. The transition function is defined as follows $Trs: St \times Act \rightarrow St$, where $Trs(st_i, act_j)$ represents the new state to which the system will move by performing an action act_j in state st_i .

For the reward function, we establish a ratio $Rwd: St \times Act \rightarrow Rwd$ that determines the amount of reward for choosing a specific action in a specific state.

A restriction is set if the selected states differ from the active states, and a mandatory yellow state is automatically entered for a certain period of time.

Yellow state restrictions can be represented in the form of the following ratios.

We denote sets:

St_{actv} is the set of active states;

St_{slcd} is the set of selected states;

St_{ylw} is the set of yellow states to be entered as mandatory for a certain period of time.

Then the constraint can be expressed by the following formula:

$$St_{ylw} = (St_{slcd} \setminus St_{activ}) \cup (St_{activ} \cap St_{slcd}). \quad (2)$$

This formula defines the set of yellow states as the union of those selected states that are not yet active ($St_{slcd} \setminus St_{activ}$) and the crossroads of selected and active states ($St_{activ} \cap St_{slcd}$).

A mandatory yellow state can be introduced with an additional parameter, for example Tm_{ylw} , which defines the duration of the mandatory yellow state. In this way, it is possible can define the time interval during which the yellow state will be active after the selected states have changed. For example:

$$St_{ylw}(t) = \begin{cases} St_{slcd}(t) \setminus St_{activ}(t) & , t \leq Tm_{ylw} \\ \emptyset & , t > Tm_{ylw}, \end{cases} \quad (3)$$

where t is the time;

$St_{slcd}(t)$ and $St_{activ}(t)$ represent the sets of selected and active states at a given time t .

Accordingly, the crossroads control system will look like this:

$$\begin{cases} St_{t+1} = Trs(St_t, Act_t); \\ Rwd(St_t, Act_t); \& Act_t = argmax_{\forall act \in Act} Rwd(St_t, act). \end{cases} \quad (4)$$

The dependencies can be adapted and extended to fit the specific details of the crossroads management system and to take into account various conditions and constraints.

The required system parameters can be represented as follows:

$$St_{t+1} = Trw(St_t, Act_t) + \alpha \cdot (Rwd(St_t, Act_t) - \beta \cdot |Sct_t \cap Act_t|) \quad (5)$$

where: St_{t+1} is the new state of the system at the next time;

$Trw(St_t, Act_t)$ – transition function, which determines how the system moves from a state St_t to a new state under the influence of the selected actions Act_t ;

α is the coefficient that takes into account the impact of the reward on the selected actions;

$Rwd(St_t, Act_t)$ – a reward function that determines how effective the choice of a specific set of actions is in each state;

β is the coefficient that takes into account the limitations of yellow states and the number of conflict states;

$|Sct_t \cap Act_t|$ is the number of conflict states in the current state and selected actions.

This formula takes into account the dynamics of the system, the impact of the selected actions on the state, the reward for these actions, and the limitations on the number of conflict states.

It is also possible to take into account the effect of pressure on the movement of vehicles, and possible changes in the system over time t . Taking these parameters into account, the description of the system state will take the following form:

$$St_{t+1} = Trw(St_t, Act_t) + \alpha \cdot [Rwd(St_t, Act_t) - \beta \cdot \delta_t \cdot (|Sct_t \cap Act_t|) - \gamma \cdot Prs(St_t, Act_t)], \quad (6)$$

where δ_t is the dynamic coefficient that takes into account changes in the system over time;
 γ is the coefficient that takes into account the effect of pressure on the movement of transport;

$Prs(St_t, Act_t)$ is the function of pressure on the movement of vehicles, which may include factors such as traffic density, speed, and other factors.

Let's take a closer look at the traffic pressure function. One possible approach is to take into account traffic density and vehicle speed. The pressure function $Prs(St_t, Act_t)$ can be expressed as follows:

$$Prs(St_t, Act_t) = \eta \cdot Dns_t(St_t, Act_t) \cdot Vls_t(St_t, Act_t), \quad (7)$$

where: η is the coefficient that determines the weight of the impact of traffic density and vehicle speed on pressure;

$Dns_t(St_t, Act_t)$ is the traffic density, which can be measured by the number of vehicles in a certain state and with selected actions;

$Vls_t(St_t, Act_t)$ is the average speed of vehicles in a certain state and with selected actions.

The pressure function can be customized according to the specific characteristics of the crossroads and the optimization goal. This formula allows taking into account both traffic density and vehicle speed as factors that affect the pressure on the movement of vehicles at the crossroads.

Traffic density $Dns_t(St_t, Act_t)$ can be determined in a variety of ways, depending on the data availability and specifications of the crossroads control system.

Some solutions for determining traffic density:

- Installing counters on the entrance roads to count the number of vehicles entering the crossroads. Traffic density is defined as the number of vehicles per unit of time on each input road.
- Using modern transportation sensor technologies, such as cameras or sensors, to automatically determine traffic density. Video stream or sensor data is analyzed to determine the number of vehicles and their movement on the roads.
- Using information from transportation agents or transportation monitoring systems that can provide traffic density data.
- Traffic modeling, which uses mathematical models to simulate traffic flows and determine traffic density based on parameters such as speed, number of lanes, and others.

Depending on the conditions and availability of resources, it is possible to choose one or a combination of these methods to determine the traffic density at a particular time and the state of the crossroads.

The general approach to determining the schedule density is as follows:

$$Dns_t(St_t, Act_t) = \frac{Vlm_{total}}{Lng_{total}}, \quad (8)$$

where: Vlm_{total} is the represents the Q volume of vehicles entering the crossroads per unit of time;

Lng_{total} is the represents the Q total length of all entrance roads leading to the crossroads.

The total volume of vehicles Vlm_{total} is determined by taking into account the number of traffic flows and their characteristics on each input road. Let Flv_{ij} the traffic flow from direction i to direction j . Then the total vehicle volume is defined as:

$$Vlm_{total} = \sum_i \sum_j Flv_{ij}. \quad (9)$$

This dual sum represents the total number of vehicles entering the crossroads through all possible combinations of input and output directions. Each element represents the number of vehicles traveling from the respective directions.

The total length of all input roads Lng_{total} is determined by adding the width of each road, if it is different for different roads. Let be Wdt_i the width of the input road. Then the total length is determined by the formula:

$$Lng_{total} = \sum_i (Lng_i \cdot Wdt_i). \quad (10)$$

In order to take into account the factors for determining the volume of vehicles, we write down the total volume of vehicles Vlm_{total} , taking into account the factors related to the speed of vehicles Vls_{ij} on each input road between crossroads i and j . We also introduce the travel time factor $Time_{ij}$, which determines the time required to travel the distance between roads i and j and take into account the traffic density between the roads in terms of the number of vehicles per unit time per kilometer $Dns_{int_{ij}}$.

Then the formula for determining the total volume of vehicles will look like this:

$$Vlm_{total} = \sum_i \sum_j (Dns_{int_{ij}} \cdot Lng_i \cdot Wdt_i \cdot Vls_{ij} \cdot Time_{ij}). \quad (11)$$

Thus, we take into account not only the length of each road, but also its width, which is an important parameter in determining the volume of vehicles and traffic density at the crossroads.

3.1. Pressure-based Coordination

In order to optimize the traffic flow in the field of pressure management, the concept of flow pressure is becoming a staple. The main focus is on improving the efficiency of the traffic flow in general. The crossroads load is defined as the difference between the length of the queues of vehicles approaching the crossroads and those leaving it. This reflects an imbalance in the distribution of vehicles.

The main task is to minimize this pressure in order to achieve equilibrium in the distribution of vehicles along the network of directions and, as a result, increase the network capacity.

The maximum pressure control strategy aims to optimize stability by not only stabilizing traffic but also maximizing flow using local data from each crossroads.

The main aspect of this strategy is to optimize traffic signal performance by reducing the pressure in each state. In real-world maximum pressure control, a greedy approach is used to achieve a locally optimal decision.

Algorithm 1: Controlling the maximum pressure for each crossroads.

1. Pressure initialization and estimation. For each state at the crossroads, the pressure $prs(st_i)$ is calculated, taking into account various aspects of the traffic flow, such as density, speed, and vehicle interaction.
2. Weight determination of the next state. Given the need to balance the various aspects of traffic, determine the next state st_{i+1} as the argument that maximizes pressure reduction while taking into account environmental considerations.
3. Adaptive pressure control. Taking into account the dynamics of the movement, the pressure calculation parameters are adaptively changed to ensure an effective response to changing road conditions.
4. Synchronization with other road intersections. Optimized state selection, taking into account common and interacting factors with other road intersections, to achieve harmonious traffic in the system.
5. Additional function of emergency states. Additional functions, such as emergency management or improved mobility of road users, are tested to ensure that a wide range of circumstances are taken into account.

3.2. DQN Agent

Agents in the reinforcement learning method seek to maximize their overall reward within the objectives of the maximum pressure control method. This increase in reward is proportional to the overall network throughput, subject to certain constraints.

Each agent is constrained to a certain subset of the overall system state. For example, for a typical crossroads that manages traffic flows, the agent's observation covers the active state and the pressure associated with the flows. In the case of fewer flows, the observation vector may contain zeros to maintain consistency.

The agent selects the state at any given time, determining the traffic light configuration. This approach allows for greater adaptability by allowing the agent to choose the optimal state to activate.

The reward for the agent is determined by the reduced pressure at the crossing. This pressure takes into account the difference in CO₂ emissions from vehicles waiting to enter and exit the crossroads.

3.3. Implementation of Deep Q-learning

Based on the chosen basic model, we apply the DQN method to solve problems related to various scenarios of traffic lights control at road intersections. The DQN implementation takes as input the state characteristics of different traffic flows and calculates the Q-value for each possible action, i.e., traffic state, based on the following Bellman equation:

$$Q(st_t, act_t) = Rrw(st_t, act_t) + \lambda \max Q(st_{t+1}, act_{t+1}), \quad (12)$$

where: $Q(st_t, act_t)$ is the represents the Q-value for state st_t and action act_t ;

$Rrw(st_t, act_t)$ is the reward for performing an action act_t in the state st_t ;

λ is the discount factor;

$\max Q(st_{t+1}, act_{t+1})$ is the maximum Q-value for the next state st_{t+1} and all possible actions act_{t+1} .

This equation estimates the Q-value for the current state and action using information about future rewards and maximum Q-values.

3.4. Synthetic test SUMO scenario GRID 4x4

The GRID 4x4 scenario in SUMO (Simulation of Urban MObility) is a synthetic test case used to simulate vehicle movements at road intersections in an urban environment. This scenario is used to test and evaluate traffic signal control algorithms, road safety, and transportation efficiency.

In the GRID 4x4 scenario, the crossroads consists of 4x4 road segments, creating a network of 16 road intersections. A large number of road intersections allow studying the interaction of traffic flows, conflicts, and optimal management strategies.

The main characteristics of the GRID 4x4 test scenario include:

- location of 4x4 road intersections;
- the number of roads is 16;
- a large number of road intersections to study the interaction of traffic flows.

Such a test scenario allows researching and testing traffic control algorithms at road intersections in an urban environment.

3.5. Evaluation of the quality of solutions

To determine the environmental impact of transport, in particular CO₂ emissions, we use the developed models with a traffic simulator. This makes it possible to model traffic in an urban environment and take into account its environmental impact.

To determine the environmental impact of CO₂ emissions in the simulation, we will take into account the following parameters:

- Identification of vehicle types, such as cars, trucks, buses, etc. Each type may have its own characteristics in terms of fuel consumption and CO₂ emissions.
- For each type of vehicle, it is necessary to specify characteristics such as average fuel consumption and CO₂ emission factor per unit of fuel consumed.
- Simulate the movement of vehicles in an urban environment, recording their routes, speeds, and fuel consumption while driving.
- Based on the recorded data, we calculate CO₂ emissions using the entered vehicle characteristics.

The calculation of CO₂ emissions is usually based on fuel consumption and vehicle-specific CO₂ emission data. To determine the CO₂ emissions, we defined the types of vehicles and specify their characteristics, such as average fuel consumption and CO₂ emission factor per kilometer. During the simulation, we determine the fuel consumption for each vehicle and the CO₂ emissions.

The fuel consumption is defined as follows:

$$FlCn = \sum_{i=1}^{N\Sigma} Fl_i \quad (13)$$

Each vehicle belongs to a certain category or type TrT that consumes a certain amount of fuel Fl .

CO₂ emissions:

$$Ems = \sum_{i=1}^N Fl_i \times FlU_{CO_2}. \quad (14)$$

Emissions are thus determined based on the fuel consumed Fl by each vehicle unit and FlU_{CO_2} the CO₂ emission factor per unit of fuel consumed.

4. Results and Discussion

To evaluate the effectiveness of our approach, we chose the SUMO GRID 4x4 scenario. This scenario is characterized by a 4x4 road grid where each crossroads has the same settings and parameters.

In the SUMO GRID 4x4 scenario, traffic flows in a network consisting of a 4x4 grid of road intersections. Each crossroads has the same settings and parameters, making it ideal for testing the effectiveness of our approach.

Throughout the experiments, we analyze various aspects such as the environmental impact of transportation, fuel consumption, and traffic signal efficiency. These aspects are determined by the scenario parameters and the performance of our approach to optimizing traffic at road intersections.

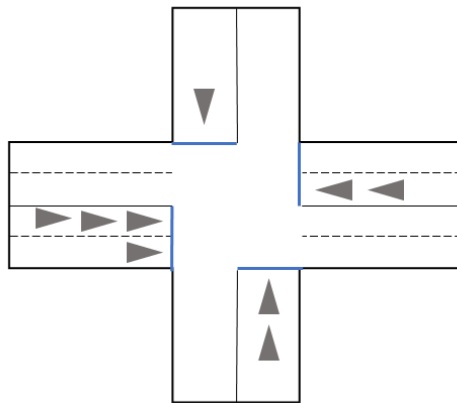


Figure 2: Scheme of the system validation scenario.

To test the performance of the developed algorithm, we used a traffic scenario similar to the synthetic 4×4 symmetric network shown in Figure 1, the 4×4 Grid. The study conducted a comparative analysis of the four algorithms:

- Maximum pressure control in which a combination of states with maximum joint pressure is enabled.
- MPLight algorithm, which uses traffic light control approaches to optimize traffic flow [4].
- Independent Deep Q-Network, i.e. independent DQN agents. For each intersection, a separate DQN agent is used, each of which uses the same convolutional neural network to aggregate information from different lanes. The hyperparameters are left by default in the Preferred RL library, except for the target network update interval, which was adapted to the environment.
- An extension of MPLight that takes into account the environmental impact, specifically CO₂ emissions from vehicles queuing to enter and exit the crossroads.

The algorithms were evaluated and compared across various environmental metrics, providing conclusions on the effectiveness and sustainability of their implementation. The diagrams below show the dynamics of queue length changes according to different numbers of training episodes.

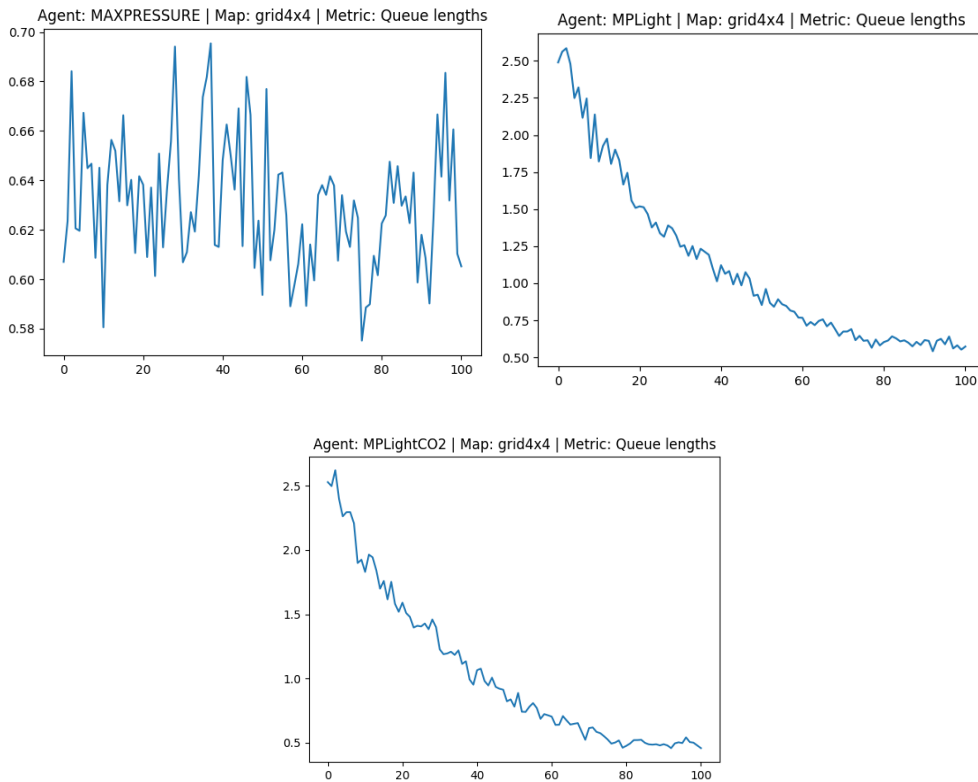


Figure 3: Dynamics of queue length change according to different number of training episodes.

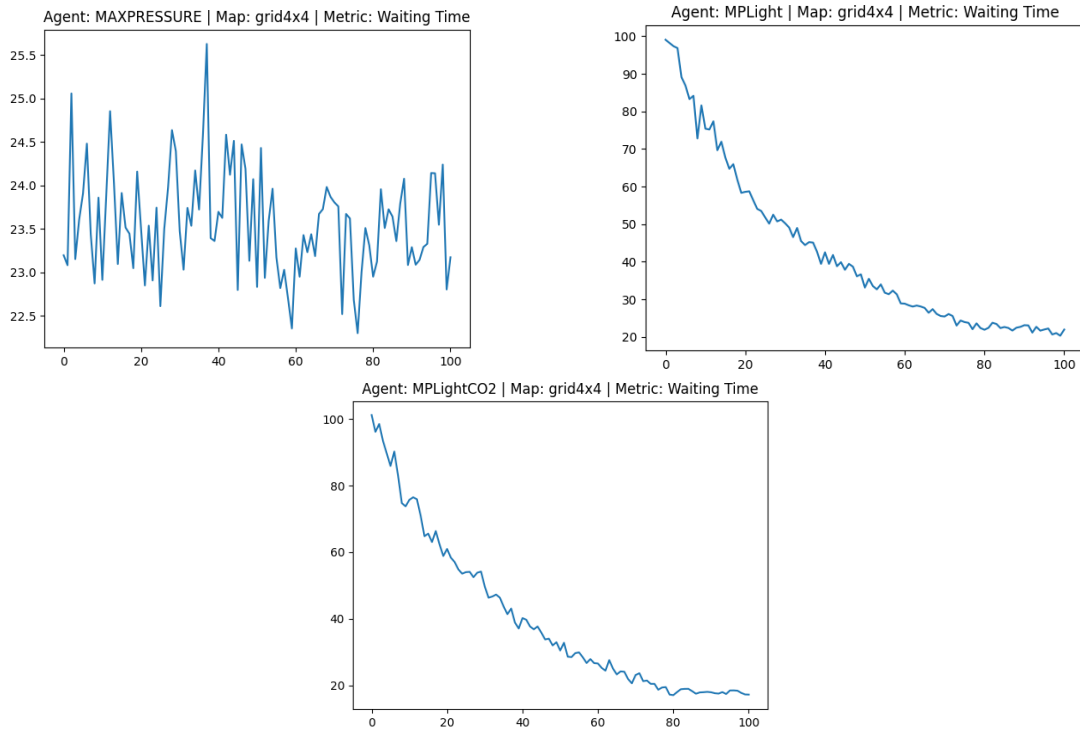


Figure 4: Average waiting time for a vehicle to cross the crossroads.

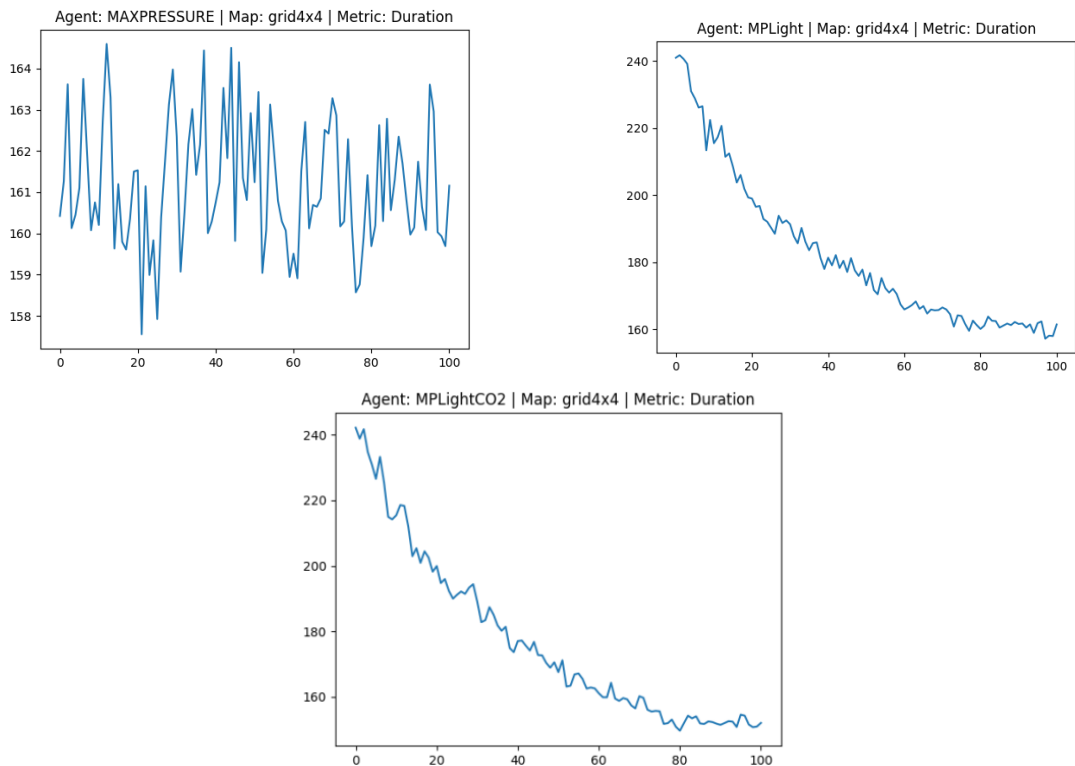


Figure 5: Average route travel time for all vehicles.

A comparative analysis of the impact of different traffic signal control algorithms on traffic efficiency and environmental impacts yielded the following results. The MPLight and MPLightCO2 methods proved to be the most effective, improving travel time and waiting time by about 33% compared to the MaxPressure algorithm. IDQN also showed improvement, but less significant, increasing travel time and waiting time by about 34%. MPLight and MPLightCO2 were effective in reducing CO₂ emissions by about 32-33%, making them environmentally friendly compared to MaxPressure and IDQN.

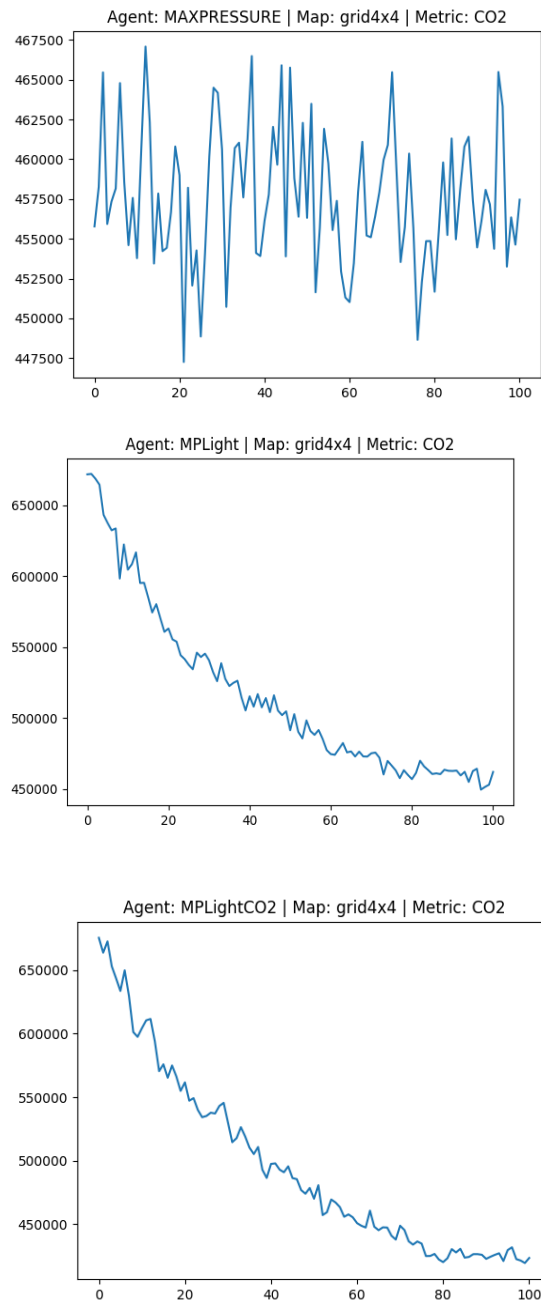


Figure 6: CO₂ emissions of vehicles crossing the crossroads.

Table 1
Quality indicators of models

| Model | Average duration of travel, s | Average queue length | CO ₂ , mg/s |
|-------------|-------------------------------|----------------------|------------------------|
| MPLight | 161,4 | 0,57 | 461908 |
| MaxPressure | 161,2 | 0,61 | 457459 |
| MPLightCO2 | 151,9 | 0,47 | 423419 |

Based on the results for travel time, MPLightCO2 performed the best, with a shorter average travel time compared to the other models. The average queue length in number of vehicle for MPLightCO2 is also the shortest, indicating more efficient traffic management and reduced congestion. MPLightCO2 showed the lowest CO₂ emissions of all the models, indicating its greater environmental efficiency.

MPLightCO2 performs better in terms of travel time, average queue length, and CO₂ emissions than both MPLight and MaxPressure. Compared to the baseline MaxPressure model, MPLightCO2 shows an improvement in travel time of 5.77%, in average queue length of approximately 29.51%, and in CO₂ emissions of approximately 7.43%.

Table 2
Results of traffic modeling

| Agent | Metric | First Value | Last Value |
|-------------|------------------------|-------------|------------|
| IDQN | Duration, s | 242,4 | 146,9 |
| MaxPressure | Duration, s | 160,4 | 161,2 |
| MPLight | Duration, s | 241,0 | 161,4 |
| MPLightCO2 | Duration, s | 242,2 | 151,9 |
| IDQN | Waiting time, s | 100,8 | 12,4 |
| MaxPressure | Waiting time, s | 23,2 | 23,2 |
| MPLight | Waiting time, s | 99,1 | 21,9 |
| MPLightCO2 | Waiting time, s | 101,2 | 17,2 |
| IDQN | CO ₂ , mg/s | 674575,8 | 408332,7 |
| MaxPressure | CO ₂ , mg/s | 455778,6 | 457459,8 |
| MPLight | CO ₂ , mg/s | 671710,8 | 461908,3 |
| MPLightCO2 | CO ₂ , mg/s | 675505,5 | 423419,3 |
| IDQN | Queue length | 2,53 | 0,34 |
| MaxPressure | Queue length | 0,61 | 0,61 |
| MPLight | Queue length | 2,49 | 0,57 |
| MPLightCO2 | Queue length | 2,53 | 0,47 |
| IDQN | Max queue | 1,67 | 0,28 |
| MaxPressure | Max queue | 0,47 | 0,47 |
| MPLight | Max queue | 1,59 | 0,44 |
| MPLightCO2 | Max queue | 1,64 | 0,36 |

MPLight, MPLightCO2, and IDQN showed similar improvements in queue length and maximum queue length, reducing them by about 75-76% and 70-71%, respectively, compared

to MaxPressure. Hence, MPLight and MPLightCO2 algorithms seem to be more effective from both a traffic improvement and environmental perspective than MaxPressure and IDQN in the studied scenario.

MPLight and MPLightCO2 performed significantly better than MaxPressure and IDQN in terms of travel time and waiting time. This may indicate the importance of considering not only traffic but also environmental aspects in crossroads management. Taking into account CO₂ emissions in the MPLightCO2 algorithm led to a significant reduction in the environmental impact of traffic. This is an important aspect in the context of urban sustainability. The reduction in queue length and maximum queue length in the MPLight and MPLightCO2 algorithms indicates their ability to effectively regulate traffic flow and provide better crossing capacity at road intersections. It is important to consider how adaptive the MPLight and MPLightCO2 algorithms are to different traffic conditions. Consider optimizing the parameters to improve adaptability in different scenarios.

5. Conclusions and Future Work

The study presents an approach that uses reinforcement learning and traffic pressure to optimize vehicle travel time through a crossroads to reduce CO₂ emissions. The proposed method is based on the modern approaches MPLight, PressLight, but with an emphasis on the priority of the CO₂ emission metric as a key component of decision-making.

The developed algorithm has been experimentally tested on the synthetic test scenario SUMO GRID 4x4, which simulates the movement of vehicles at road intersections in an urban environment. The comparative analysis showed that the MPLight and MPLightCO2 algorithms, which take into account the impact on the environmental situation, proved to be more effective than MaxPressure and IDQN. They demonstrated an improvement in travel time and waiting time of up to 33%, a reduction in queue length by 75-76%, and a reduction in CO₂ emissions by 32-33%.

MPLightCO2 showed the best results among the algorithms compared, with the shortest travel time, shortest average queue length, and lowest CO₂ emissions, indicating its high performance from both a traffic improvement and environmental perspective.

The results obtained are preliminary and more testing on different models and configurations, as well as verification in real-world conditions, is needed. However, the proposed approach has shown satisfactory control results in a large-scale road network, which gives grounds for its further improvement and implementation.

6. Acknowledgements

The study was executed as a component of the Horizon Europe Framework Program, receiving support from the initiative aimed at aligning Ukrainian cities with the mission for Climate-neutral and smart cities (HORIZON-MISS-2023-CIT-02).

7. References

- [1] O. Pavlova, A. Bilinska, A. Holovatiuk, Y. Binkovskyi, D. Melnychuk. Automated System for Determining Speed of Cars Ahead. *Computer Systems and Information Technologies*. 2023. №3. Pp. 32-39.
- [2] C. Fetting, The european green deal. ESDN report, 2020, 53.
- [3] A. Haydari, Y. Yilmaz, Deep reinforcement learning for intelligent transportation systems: a survey. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 23 (1), 11–32.
- [4] S. Damadam, M. Zourbakhsh, R. Javidan, A. Faroughi, An intelligent iot based traffic light management system: deep reinforcement learning. *Smart Cities*, 2022, 5 (4), 1293–1311.
- [5] C. Chen, H. Wei, N. Xu, G. Zheng, M. Yang, Y. Xiong, K. Xu, Z. Li, Toward a thousand lights: decentralized deep reinforcement learning for large-scale traffic signal control. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34 (04) (2020) 3414–3421. <https://doi.org/10.1609/aaai.v34i04.5744>.
- [6] T. Wu, P. Zhou, K. Liu, Y. Yuan, X. Wang, H. Huang, D. O. Wu, Multi-agent deep reinforcement learning for urban traffic light control in vehicular networks. *IEEE Transactions on Vehicular Technology*, 69 (8) (2020) 8243–8256. doi:10.1109/TVT.2020.2997896.
- [7] O. Barmak, I. Krak, E. Manziuk, Diversity as the basis for effective clustering-based classification. *CEUR-WS*, 2020, 2711, 53–67.
- [8] E. Manziuk, I. Krak, O. Barmak, O. Mazurets, V. Kuznetsov, O. Pylypiak, Structural alignment method of conceptual categories of ontology and formalized domain. *CEUR-WS*, 3003 (2021) 11–22.
- [9] E. A. Manziuk, W. Wójcik, O. V. Barmak, I. V. Krak, A. I. Kulas, V. A. Drabovska, V. M. Puhach, S. Sundetov, A. Mussabekova, Approach to creating an ensemble on a hierarchy of clusters using model decisions correlation. *Przegląd Elektrotechniczny*, 96 (9) (2020) 108–113. doi:10.15199/48.2020.09.23.
- [10] L. Peng, L. Wang, D. Xia, Q. Gao, Effective energy consumption forecasting using empirical wavelet transform and long short-term memory. *Energy*, 238, (2022) 121756. doi:10.1016/j.energy.2021.121756.
- [11] Ü. Ağbulut, Forecasting of transportation-related energy demand and co2 emissions in turkey with different machine learning algorithms. *Sustainable Production and Consumption*, 2022, 29, 141–157. doi:/10.1016/j.spc.2021.10.001.
- [12] S. Jomnonkwao, S. Uttra, V. Ratanavaraha, Forecasting road traffic deaths in thailand: applications of time-series, curve estimation, multiple linear regression, and path analysis models. *Sustainability*, (2020) 12 (1), 395. doi:10.3390/su12010395.
- [13] T. Afrin, N. Yodo, A survey of road traffic congestion measures towards a sustainable and resilient transportation system. *Sustainability*, 12 (11) (2020) 4660. doi:10.3390/su12114660.

- [14] M. Nama, A. Nath, N. Bechra, J. Bhatia, S. Tanwar, M. Chaturvedi, B. Sadoun, Machine learning-based traffic scheduling techniques for intelligent transportation system: opportunities and challenges. *International Journal of Communication Systems*, 34 (9) (2021) e4814. doi:10.1002/dac.4814.
- [15] A. Navarro-Espinoza, O. R. López-Bonilla, E. E. García-Guerrero, E. Tlelo-Cuautle, D. López-Mancilla, C. Hernández-Mejía, E. Inzunza-González, Traffic flow prediction for smart traffic lights using machine learning algorithms. *Technologies*, (2022) 10, 5. doi:10.3390/technologies10010005.
- [16] G. Markowsky, O. Savenko, S. Lysenko, A. Nicheporuk, A. The technique for metamorphic viruses' detection based on its obfuscation features analysis, *CEUR-WS*, 2104 (2018) 680–687.
- [17] B. Zhurakovskiy, J. Boiko, V. Druzhynin, I. Zeniv, O. Eromenko, Increasing the efficiency of information transmission in communication channels. *Indonesian Journal of Electrical Engineering and Computer Science*, 19 (3) (2020) 1306–1315.
- [18] Z. Lei, D. Qin, L. Hou, J. Peng, Y. Liu, Z. Chen, An adaptive equivalent consumption minimization strategy for plug-in hybrid electric vehicles based on traffic information. *Energy*, 190 (2020) 116409. doi:10.1016/j.energy.2019.116409.
- [19] H. He, Y. Wang, J. Li, J. Dou, R. Lian, Y. Li An improved energy management strategy for hybrid electric vehicles integrating multistates of vehicle-traffic information. *IEEE Transactions on Transportation Electrification*, 7 (3) (2021) 1161–1172. doi:10.1109/TTE.2021.3054896.
- [20] A. Arbi, N. Tahri Stability analysis of inertial neural networks: a case of almost anti-periodic environment. *Mathematical Methods in the Applied Sciences*, 45 (16) (2022) 10476–10490.
- [21] Z. Sabir, A. Arbi, A. F. Hashem, M. A. Abdelkawy Morlet wavelet neural network investigations to present the numerical investigations of the prediction differential model. *Mathematics*, , 11 (21) (2023) 1–20.
- [22] A. Arbi, J. Cao, M. Es-saiydy, M. Zarhouni, M. Zitane Dynamics of delayed cellular neural networks in the stepanov pseudo almost automorphic space. *Discrete & Continuous Dynamical Systems-Series S* 15 (11) (2022).
- [23] F. Rasheed, K.-L. A. Yau, R. Md. Noor, C. Wu, Y.-C. Low Deep reinforcement learning for traffic signal control: a review. *IEEE Access*, 8 (2020) 208016–208044. doi:10.1109/ACCESS.2020.3034141.
- [24] S. Koh, B. Zhou, H. Fang, P. Yang, Z. Yang, Q. Yang, L. Guan, Z. Ji Real-time deep reinforcement learning based vehicle navigation. *Applied Soft Computing*, (2020), 96, 106694. doi:10.1016/j.asoc.2020.106694.
- [25] J. Ault, J. P. Hanna, G. Sharon Learning an interpretable traffic signal control policy. arXiv preprint arXiv:1912.11023, (2019).
- [26] X. Wang, L. Ke, Z. Qiao, X. Chai, Large-scale traffic signal control using a novel multiagent reinforcement learning. *IEEE Transactions on Cybernetics*, 51 (1) (2021) 174–187. doi:10.1109/TCYB.2020.3015811.