VIA AI: Reliable Deep Reinforcement Learning for Traffic Signal Control

Matvey Gerasyov [1] Dmitrii Kiselev [1] Maxim Beketov [1] Ilya Makarov [2,3]

[1]HSE University, Moscow, Russia [2]Artificial Intelligence Research Institute (AIRI), Moscow, Russia [3]ISP RAS Research Center for Trusted Artificial Intelligence, Moscow, Russia

Abstract

Traffic signal control optimization is an integral part of any modern transportation system. However, modern traffic signal control systems often rely on predetermined fixed rules to adjust traffic signal timings. This paper presents VIA AI - an intelligent traffic signal control system that leverages deep reinforcement learning (RL) applied to count-based traffic data. Our solution offers additional adaptability and flexibility by allowing the system to learn and adjust its strategies based on real-time feedback and environmental changes. We test our approach using real-world traffic data and show that it outperforms classical methods of intersection control.

Importance of Intelligent Traffic Signal Control

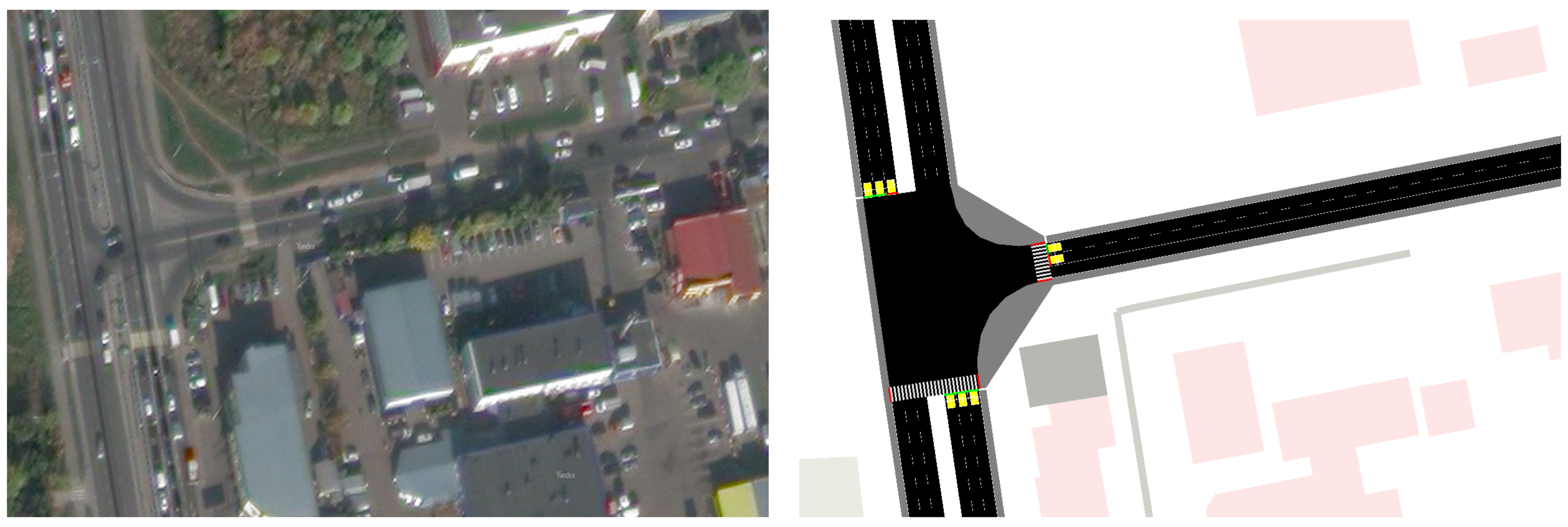

Figure 1. Comparison of the satellite view (left) of the modeled intersection with the SUMO view (right)

Modern cities frequently experience traffic congestion, particularly during peak commuting hours. Intelligent traffic light signals can reduce congestion on busy roads and intersections, make intersections safer by reducing collisions and accidents, reduce air emissions, and allow drivers to save time and fuel. In addition, it is possible to integrate intelligent traffic signal control with other smart city technologies, such as autonomous vehicles and public transit systems.

In this paper, we present a novel traffic signal control system powered by deep reinforcement learning. Our system is both flexible and designed with realistic input requirements, enabling it to adapt dynamically to fluctuating traffic conditions and improve overall efficiency. To evaluate our approach, we conduct a series of experiments using real-world traffic data collected from an isolated three-way intersection, shown in Fig. 1.

Data Collection and Processing

Traffic data enters our system via an API provided by our partners in compliance with local authorities. Initially, the system stores vehicle counts and average speed data for each detector in the database. These stored data are then utilized during training to construct a realistic traffic flow simulation using SUMO. To validate the obtained simulation, we compare simulated vehicle counts with the original data (see Fig. 2).

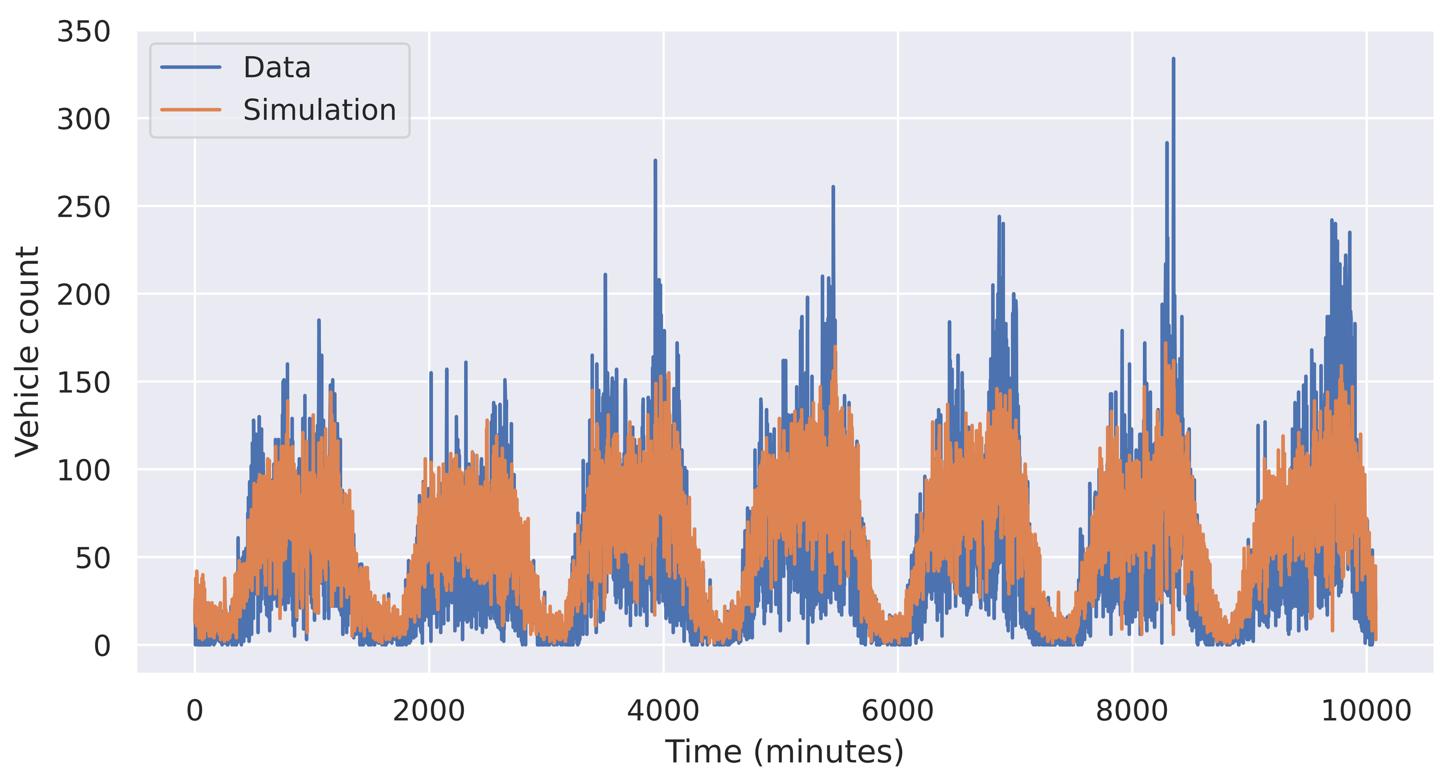

Figure 2. Comparison of the vehicle counts in simulation (orange) and in real traffic data (blue).

Training Details

The agent's input includes vehicle counts and average speeds recorded during the previous time period, along with the distribution of green time among phases in the current traffic light program. The agent selects a traffic light program from a predefined list, with each program varying only in cycle length and green time distribution. This action space design mirrors real-world intersection management constraints. The reward function, crucial for guiding the agent's performance, uses negative accumulated waiting time as feedback. This ensures the agent does not prioritize one approach over others. To train the agent, we employ the well-established deep reinforcement learning method, Deep Q-Network (DQN), in conjunction with the epsilon-greedy exploration strategy.

VIA AI Production Pipeline

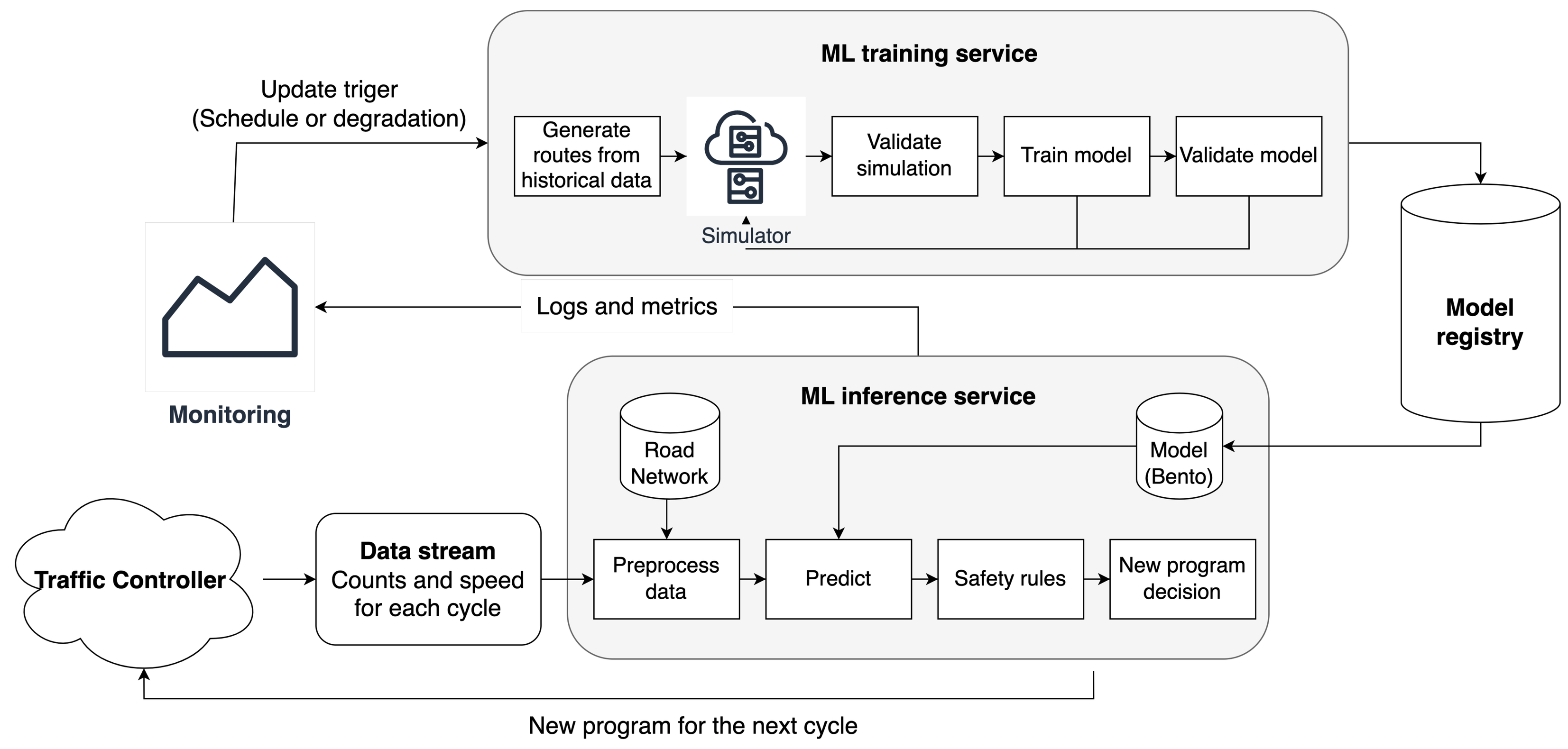

Figure 3. VIA AI production pipeline

After the model development, we deploy it in an automated pipeline to keep model performance at high levels over time (see Fig 3). It consists of two main blocks: Training and Inference services. The first generates routes from the historical data and conducts simulations to train and validate the model. The Inference service receives the actual data about vehicle counts and average speed from different lanes, preprocesses it for the model trained with the Training service, applies safety rules, and sends decisions about the selected program to the traffic controller. Inference service stores all predictions, metrics, and logs in monitoring databases. The model is updated when we detect model degradation or by schedule.

Experimental Results

To evaluate the agent, we conducted a simulation using one week of real-world data. The validation dataset was carefully selected to ensure no overlap with the data used for training. We compared the agent's performance against two baselines: the fixed program schedule implemented at the selected intersection and a program schedule generated using Webster's formula for green time allocation based on historical data.

We measured performance using a set of relevant metrics. The time loss metric indicates the average delay experienced by a vehicle due to stops when passing through an intersection. The trip time metric represents the average travel time from when a vehicle enters the simulation to when it exits. Lastly, the waiting time metric refers to the average duration a vehicle spends at low speed (< 0.1 m/s). The final evaluation results are presented in Table 1.

Table 1. Evaluation results. All values in the table are expressed in seconds. DQN stays for the reinforcement learning agent. Webster refers to the schedule created using Webster's formula. Finally, Fixed Time is baseline schedule deployed at the real intersection.

Algorithm	Time Loss	Trip Time	Waiting Time
DQN	**73.4**	**131.9**	**52.9**
Webster	79.2	144.5	57.2
Fixed Time	97.2	164.8	73.1

As one can see, the proposed DQN agent is ahead of classical approaches in all considered metrics. Furthermore, it was observed that the baseline program schedules led to a significant increase in the average waiting time and other metrics during peak congestion hours and periods with rapidly varying traffic, whereas our agent demonstrated much more consistent results during the evaluation.

Conclusion

The problem of optimally controlling traffic signals at a given intersection is not new. However, most traffic signal control systems today still operate on fixed timing schedules preset by a transport engineer. While these fixed schedules can provide near-optimal performance under stable and predictable traffic conditions, they lack adaptability and struggle to respond to irregular traffic patterns. As can be seen in our experiments, this rigidity often exacerbates congestion and delays.

In this study, we presented VIA AI: a novel intelligent traffic signal control system. We investigated using count-based data to train deep reinforcement learning agents to control traffic lights at a signalized intersection. Our study suggests that, despite lacking complete information about vehicles in the system, the deep reinforcement learning approach demonstrates improved performance compared to classical baselines. In future studies, we plan to extend our results to multi-agent scenarios and verify whether it's possible to learn complex coordination between multiple controlled intersections using count-based traffic data.

msgerasyov@hse.ru