Green Wave Coordination For Traffic Signal Control Using Deep Reinforcement Learning

Maxime Tréca[a,*], Mahdi Zargayouna[b], Dominique Barth[a] and Julian Garbiso[c]

[a] Université Versailles Saint-Quentin-en-Yvelines, Versailles, France
maxime.treca@gmail.com, dominique.barth@uvsq.fr
[b] Université Gustave Eiffel, Champs-sur-Marne, France
mahdi.zargayouna@univ-eiffel.fr
[c] Institut Védécom, Versailles, France
julian.garbiso@vedecom.fr
* Corresponding author

Extended abstract submitted for presentation at the 11th Triennial Symposium on Transportation Analysis conference (TRISTAN XI)
June 19-25, 2022, Mauritius Island

March 16, 2022

Keywords: Traffic Signal Control, Green Wave Coordination, Deep Reinforcement Learning

1 INTRODUCTION

Traffic signal control (TSC) designates the use of traffic lights to achieve two crucial tasks in urban mobility: optimizing traffic flows on a road network while ensuring the safety of its users. While various classical TSC methods have all aimed at solving these two challenges with different levels of success and complexity, the use of reinforcement learning (RL) algorithms for TSC, which first appeared in 1992 and has known an exponential rise since the years 2010 (Noaeen et al., 2021), have also provided excellent results in simulated traffic conditions.

The reinforcement learning for traffic signal control (RL-TSC) literature has developed and improved in recent years: recent advances in the field of deep learning has marked a shift from the use of classical RL algorithms such as Q-learning to more advanced deep reinforcement learning algorithms such as 3DQN (Gregurić et al., 2020) for TSC, notably increasing the performance of RL-TSC controllers. A common area of research in the RL-TSC literature nowadays is the study of coordination mechanisms used by the learning agents in multi-intersection road networks, which can improve the overall performance of RL-TSC systems. These modes of coordination are usually divided into three groups: *independent control*, in which intersections do not communicate with each other, *indirect coordination* methods, such as MARLIN-IC (El-Tantawy and Abdulhai, 2012), in which intersections model joint state and joint actions with their neighbors through indirect state observation and *direct coordination* methods, such as MARLIN-DC (El-Tantawy and Abdulhai, 2012), in which intersections directly communicate in order to decide on their next traffic-routing actions jointly. This paper presents a novel fourth mode of coordination directly inspired by the traffic engineering literature known as *green wave coordination*, in which multiple intersections maximize the vehicular throughput over an arterial. This extended abstract presents this novel coordinated RL-TSC method in section 2, before pitting it against a state-of-the-art RL-TSC method in a simulated setting in section 3. Finally, we conclude and present future research directions in section 4.

2 LEARNING AND TRAFFIC SIGNAL CONTROL

2.1 Reinforcement Learning for Traffic Signal Control

RL tasks for TSC are usually modeled using a Markov decision process (MDP). At each time step t, the agent (i.e. the intersection) observes the state of the environment s_t and chooses an action a_t. We define the state as $s_t = (\phi_t, d_t, c_t(l_1), \ldots, c_t(l_n))$, where ϕ_t represents the active green phase index, d_t the duration for which this phase has been active, and $c_t(l_i)$ the amount of vehicles present in lane l_i at step t. On the basis of this observation, the intersection chooses an action a_t influencing the current traffic signal settings in different ways depending on the RL-TSC method (see subsection 2.2). After application of the chosen action a_t, the traffic environment transitions to a new state s_{t+1} and returns a reward r_t to the agent. The reward of an agent taking two successive actions at steps t and $t+k$ is defined as $r_t = \sum_{l_i} \omega_{t+k}(l_i) - \sum_{l_i} \omega_t(l_i)$, where $\omega_t(l_i)$ is the cumulated waiting time of vehicles on lane l_i at step t,

2.2 Traffic Signal Control Methods

This paper features two RL-TSC methods based on deep reinforcement learning. The first method, I-2DQN, does not coordinate between agents and is hence an independent method. This method is both used as a benchmark and as a building block for the second method, GW-DQN, which features deep reinforcement learning for green wave coordination.

2.2.1 Independent Traffic Signal Control

The independent dueling deep Q-network (I-2DQN) method is a state-of-the-art RL-TSC method (Gregurić et al., 2020) which leverages modern reinforcement learning techniques. Our I-2DQN implementation features three fully connected hidden layers of 128 neurons, each associated with a ReLU rectifier and batch normalization layer, and a final double layer used for the computation of the advantage and value functions $A(s,a)$ and $V(s)$ typically found in dueling networks (Wang et al., 2016). This method uses experience replay by sampling mini-batches of 32 observations from a replay buffer $\mathcal{D}$ and uses a target network θ^T, which is updated every 100 learning steps. The implemented policy is an ε-greedy policy with a decaying search parameter ε which ensures a good balance between exploration and exploitation of the problem's state space. Intersections using I-2DQN choose between *step-based* actions: at each decision step t, the intersection has to choose whether to extend the current green phase by a single step (up to a parameter maximum duration $d_{\max}$) or to switch to the next green phase within the signal cycle (after a parameter mandatory amber and red safety phase of duration $d_{\min}$). Since they do not impose constraints on the total signal cycle duration, and since they provide a high level of signal control, step-based actions usually provide, *ceteris paribus*, superior performances in RL-TSC applications (Tréca et al., 2020).

2.2.2 Green Wave Coordination

Green wave coordination aims at maximizing the throughput over a major arterial by limiting the number of stops encountered by vehicles driving alongside it. This green wave phenomenon is achieved by computing *offsets*, which indicates the travel time of a vehicle driving between each intersection of the arterial in non-saturated traffic conditions. These offsets are usually computed manually using a time-space diagram, as shown on Figure 1. Interestingly, to the best of our knowledge, green wave coordination has not been applied in the RL-TSC literature.

The green wave dueling deep Q-network (GW-2DQN) method is a novel RL-TSC method featuring green wave coordination. Each intersection using this method features the same architecture as the I-2DQN method, with the exception that it implements *phase-based* actions. Phase-based actions, for which an agent chooses the entire length of a green phase at once, are

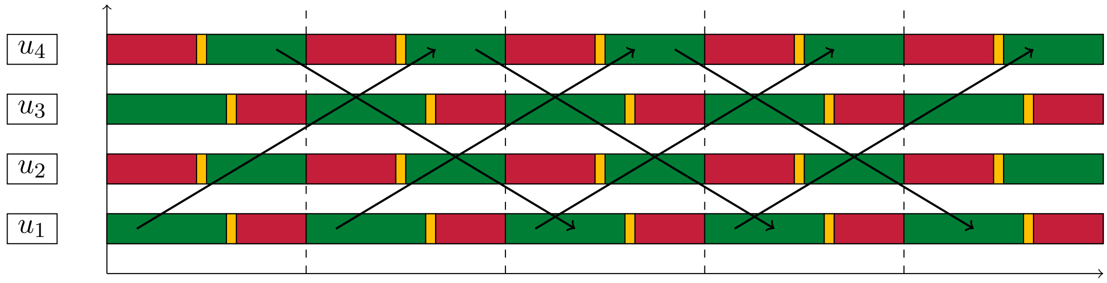

Figure 1 – *Example Time-Space Diagram on a four intersection arterial. The x-axis represents time and the y axis distance. The signal cycles of the four intersections are computed so that green waves, represented by black arrows, can occur in both directions of the arterial.*

needed to maintain proper offsets between signal cycles of the intersections along the arterial and allow for a green wave to occur. The signal cycle's duration, D, is divided between the green phase duration over the arterial d_1 and perpendicular to the arterial, d_2, and two mandatory amber and red phases of fixed duration d_{min}. Each intersection implementing the GW-DQN method hence has to pick a single action per signal cycle: the duration of green phase d_1, which can range between d_{min} and $D - 3d_{\text{min}}$, respectively the minimal green phase duration, and the maximal allowable phase duration which ensures that phase d_2 receives the required minimal phase duration. After choosing duration d_1, duration d_2 is automatically computed to fill up the entire signal cycle of total duration D.

The GW-DQN method is straightforward to implement since the only requirement is to maintain offsets between intersections of the arterial, which can easily be expressed as an additional set of constraints on action selection, regardless of the underlying RL method. Furthermore, the GW-DQN method represents a trade-off due to these additional constraints. On the one hand, using phase-based actions allows for green wave coordination, maximizing throughput over the arterial. On the other hand, using a phase-based action space reduces the adaptability of intersection to changing traffic conditions due to longer decision intervals, decreasing agent performance in the process (Tréca et al., 2020). Traffic simulations in section 3 comparing the I-2DQN and GW-2DQN methods should indicate which factors are more prominent.

3 EXPERIMENTAL RESULTS

3.1 Experimental Settings

We run our experiments on the `SUMO` traffic simulator and the `carmulator` RL-TSC library we have developed. The two methods are tested on a 4-intersection arterial road network. We use signal cycle values of $D = 60$ and $d_{\text{min}} = 5$. Green wave offsets are automatically computed by `SUMO`. Traffic is generated using a Poisson arrival process of parameter $\lambda_{u,v}$ for each entry and exit lane pair (u, v) of the network. The base value of parameter $\lambda_{u,v}$ is doubled if the entry lane u is located on the arterial or if the exit lane v is located on the arterial, and tripled if both u and v are located on the arterial, which promotes larger traffic flows on the main arterial while maintaining traffic flows on alternative routes.

3.2 Agent Performance

We measure the performance of both methods after their training process of 500 simulation episodes of 1000 steps each. These performances are measured by running 20 traffic scenarios, each using distinct traffic demand patterns, and by measuring the performance spectrum of each method by plotting the worst and best vehicular cumulated waiting time evolution throughout the simulation scenario. These performance measurements are plotted in two cases: normal traffic conditions, using a near-saturating base arrival rate of $\lambda = 0.06$ (around 216 vehicles/hour) and saturated traffic conditions using a base arrival rate of $\lambda = 0.08$ (around 288 vehicles/hour), in order to observe how both methods fare in these scenarios.

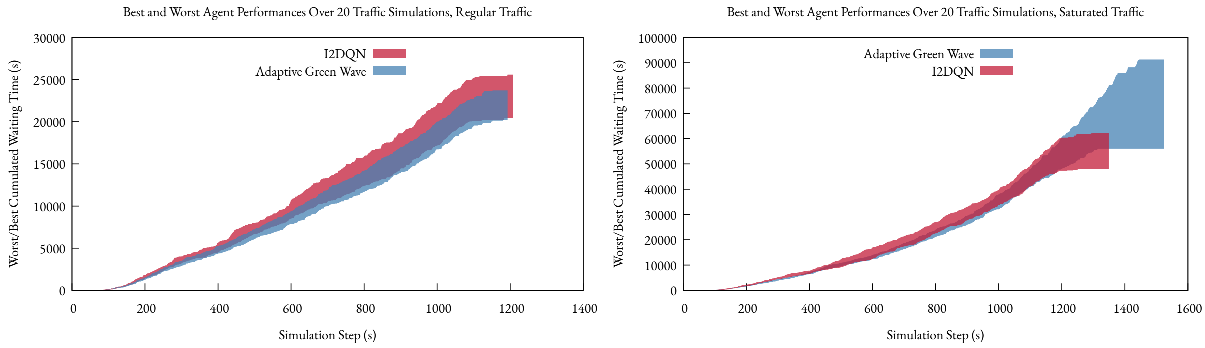

Figure 2 – *Performance spectrum of the I-2DQN and GW-2DQN methods in normal (left) and saturated (right) traffic conditions.*

Results show that the GW-2DQN method outperforms the I-2DQN method in terms of average and variance in cumulated waiting time in normal traffic conditions, showing that green wave coordination on arterials using deep reinforcement learning is beneficial for TSC. However, this relative superiority is entirely reversed when traffic conditions are saturated since the GW-2DQN performance drops significantly in the second experiment.

4 DISCUSSION AND FUTURE WORKS

This paper proposes a novel green-wave-based approach for a deep RL-TSC. The results show that the method outperforms I-2QN but underline the somewhat fragile nature of green wave coordination. If the GW-2DQN method allows for increased performance gains with minimal overhead in terms of complexity, these gains only hold if the green wave phenomenon can occur, implying that no traffic congestion is present along the arterial. Since this condition is not likely to be constantly verified in real-life scenarios, a hybrid RL-TSC method that could automatically switch between I-2DQN and GW-2DQN depending on congestion levels along an arterial could offer the best of both worlds. Such a mechanism could easily be constructed using vehicular congestion values along the arterial lanes, which are already used as a state feature of both RL algorithms.

References

S. El-Tantawy and B. Abdulhai. Multi-agent reinforcement learning for integrated network of adaptive traffic signal controllers (marlin-atsc). In *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*, pages 319–326. IEEE, 2012.

M. Gregurić, M. Vujić, C. Alexopoulos, and M. Miletić. Application of deep reinforcement learning in traffic signal control: An overview and impact of open traffic data. *Applied Sciences*, 10(11), 2020.

M. Noaeen, A. Naik, L. Goodman, J. Crebo, T. Abrar, B. Far, Z. S. H. Abad, and A. L. Bazzan. Reinforcement learning in urban network traffic signal control: A systematic literature review. 2021.

M. Tréca, J. Garbiso, and D. Barth. Guidelines for action space definition in reinforcement learning-based traffic signal control. In *ICAPS Conference, Special Track on Planning and Learning*. ICAPS, AAAI, 2020.

Z. Wang, T. Schaul, M. Hessel, H. Hasselt, M. Lanctot, and N. Freitas. Dueling network architectures for deep reinforcement learning. In *International conference on machine learning*, pages 1995–2003. PMLR, 2016.