PROCEEDINGS OF ELMAR-2021

63rd International Symposium ELMAR-2021

13-15 September 2021, Zadar, Croatia

EDITED BY

Mario Muštra
Josip Vuković
Branka Zovko-Cihlar

University of Zagreb
Croatia

Published by: Croatian Society Electronics in Marine - ELMAR

Edited by: Mario Muštra [1], Josip Vuković [2], Branka Zovko-Cihlar [2]
[1] University of Zagreb, Faculty of Transport and Traffic Sciences
Vukelićeva 4, 10000 Zagreb, Croatia
[2] University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3 / XII, 10000 Zagreb, Croatia

Front Cover: Painting by artist Mrs Ljerka Njerš

Printed by: Ispis Ltd., Zagreb

Print ISBN: 978-1-6654-4436-1, CFP21825-PRT

XPLORE ISBN: 978-1-6654-4437-8, CFP21825-ART

Printed in Croatia.

ELMAR

63rd International Symposium ELMAR-2021 is organised by:

Croatian Society Electronics in Marine - ELMAR, Zadar

University of Zagreb Faculty of Electrical Engineering and Computing,
Department of Communication and Space Technologies

ELMAR-2021 symposium Sponsors

Ministry of the Sea, Transport and Infrastructure of the Republic of Croatia

Ministry of Science and Education of the Republic of Croatia

HAKOM

Croatian Regulatory Authority for Network Industries

OIV Digital signals and networks

OIV – Digital signals and networks

Croatian Telecom Inc.

A1 Croatia Ltd.

Town of Zadar

County of Zadar

ELMAR-2021 symposium takes place under the co-sponsorship of:

IEEE Region 8
IEEE Croatia Section
IEEE Croatia Section Chapter of the Signal Processing Society
IEEE Croatia Section Chapter of the Antennas and Propagation Society
IEEE Croatia Section Chapter of the Microwave Theory and Techniques Society
Croatian Academy of Sciences and Arts
Croatian Academy of Engineering - HATZ
University of Zagreb
University of Zagreb Faculty of Electrical Engineering and Computing
University of Zagreb, Faculty of Transport and Traffic Sciences
University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology

ELMAR-2021 symposium is supported by:

ZADAR TOURIST BOARD
ERICSSON NIKOLA TESLA
OTP BANKA HRVATSKA
FALKENSTEINER BORIK ZADAR CROATIA

ELMAR-2021 SYMPOSIUM COMMITTEES

General Chair

Branka Zovko-Cihlar, University of Zagreb, Croatia

General Co-Chair

Mislav Grgić, University of Zagreb, Croatia

Program Chair

Mario Muštra, University of Zagreb, Croatia

ELMAR-2021 SYMPOSIUM INTERNATIONAL PROGRAM COMMITTEE

Juraj Bartolić, Croatia
Narcis Behlilović, Bosnia and Herzegovina
Jelena Božek, Croatia
Aura Conci, Brasil
Marek Domanski, Poland
Janusz Filipiak, Poland
Borko Furht, USA
Mohammed Ghanbari, United Kingdom
Dušan Gleich, Slovenia
Mislav Grgić, Croatia
Sonja Grgić, Croatia
Ho Yo-Sung, Korea
Bernhard Hofmann-Wellenhof, Austria
Ebroul Izquierdo, United Kingdom
Ana Katalinić Mucalo, Croatia
Ismail Khalil, Austria
Tomislav Kos, Croatia
Igor Kuzle, Croatia
Panos Liatsis, United Arab Emirates
Lidija Mandić, Croatia
Branka Medved Rogina, Croatia
Borivoj Modlic, Croatia
Marta Mrak, United Kingdom
Mario Muštra, Croatia
Zdeněk Němec, Slovakia
Miloš Oravec, Slovakia
Jarmila Pavlovičova, Slovakia
Fernando Pereira, Portugal
Jan Pidanič, Slovakia
Peter Planinšič, Slovenia
Pavol Podhradský, Slovakia
Darko Ratkaj, Switzerland
Fabiana Rodrigues Leta, Brasil
Gregor Rozinaj, Slovakia
Markus Rupp, Austria
Shiguang Shan, China
Thomas Sikora, Germany
Karolj Skala, Croatia
Ryszard Stasinski, Poland
Luis Torres, Spain
Frantisek Vejrazka, Czech Republic
Dijana Vitas, Croatia
Stamatis Voliotis, Greece
Josip Vuković, Croatia
Krzysztof Wajda, Poland
Branka Zovko-Cihlar, Croatia

ELMAR-2021 SYMPOSIUM
INTERNATIONAL REVIEW COMMITTEE

Goran Bakalar, Croatia
Alen Begović, Bosnia and Herzegovina
Marko Bosiljevac, Croatia
Jelena Božek, Croatia
Tonči Carić, Croatia
Aura Conci, Brasil
Emil Dumić, Croatia
Hrvoje Gold, Croatia
Sonja Grgić, Croatia
Edouard Ivanjko, Croatia
Niko Jelušić, Croatia
Juraj Kačur, Slovakia
Serđo Kos, Croatia
Tomislav Kos, Croatia
Hrvoje Leventić, Croatia
Panos Liatsis, United Arab Emirates
Časlav Livada, Croatia
Josip Lončar, Croatia
Sadko Mandžuka, Croatia
Marko Matulin, Croatia
Marta Mrak, United Kingdom
Štefica Mrvelj, Croatia
Mario Muštra, Croatia
Miloš Oravec, Slovakia
Jarmila Pavlovičová, Slovakia
Jan Pidanič, Slovakia
Peter Planinšič, Slovenia
Pavol Podhradský, Slovakia
Tatiana Privalova, Russia
Snježana Rimac-Drlje, Croatia
Gregor Rozinaj, Slovakia
Markus Rupp, Austria
Renata Rybárová, Slovakia
Dorsaf Sebai, France
Chakib Taybi, Morocco
Filip Turčinović, Croatia
Leo Vincelj, Croatia
Mario Vranješ, Croatia
Josip Vuković, Croatia
Ivana Žeger, Croatia

ELMAR-2021 SYMPOSIUM
ORGANISING COMMITTEE

Mislav Grgić, Croatia
Mario Muštra, Croatia
Jelena Božek, Croatia
Filip Turčinović, Croatia
Josip Vuković, Croatia

CROATIAN SOCIETY ELECTRONICS IN MARINE - ELMAR

President: *Branka Zovko-Cihlar*

Vice-president: *Ivan Pupovac*

Secretary: *Vladimir Roboz*

Members of the ELMAR Executive Committee:

Winton Afrić, Irena Galić, Mislav Grgić, Sonja Grgić, Irena Jurdana, Branka Medved Rogina, Snježana Rimac-Drlje, Vladimir Roboz, Sonja Zentner Pilinsky

Members of the ELMAR Supervisory Committee:

Tomislav Kos, Boris Kviz, Lidija Mandić

Address: Faculty of Electrical Engineering and Computing
Croatian Society Electronics in Marine - ELMAR
Unska 3 / XII, 10000 Zagreb, CROATIA

Phone: + 385 1 6129 567

Fax: + 385 1 6129 717

E-mail: elmar2021@fer.hr

URL: www.elmar-zadar.org

Combining Neural Gas and Reinforcement Learning for Adaptive Traffic Signal Control

Mladen Miletić[1], Edouard Ivanjko[1], Sadko Mandžuka[1], Daniela Koltovska Nečoska[2]
[1]Faculty of Transport and Traffic Sciences, University of Zagreb
Vukelićeva 4, HR-10000 Zagreb, Croatia
[2]Faculty of Technical Sciences St. Kliment Ohridski University – Bitola
Boulevard 1st of May b.b., Bitola 7000, Republic of North Macedonia
mladen.miletic@fpz.unizg.hr

***Abstract*—Travel time of vehicles in urban traffic networks can be reduced by using Adaptive Traffic Signal Control (ATSC) to change the signal program according to the current traffic situation. Modern ATSC approaches based on Reinforcement Learning (RL) can learn the optimal signal control policy. While there are multiple RL based ATSC implementations available, most suffer from high state-action complexity leading to slow convergence and long training time. In this paper, the state-action complexity of ATSC based RL is reduced by implementing Growing Neural Gas learning structure as an integral part of RL, leading to high convergence rate and system stability. The presented approach is evaluated on a simulated signalized intersection, and compared with self-organizing map RL-based ATSC systems. Obtained results prove that the reduction of state-action complexity in this manner improves the effectiveness of RL based ATSC not needing to have an a priory analysis of needed number of neurons for state representation.**

***Keywords*—Intelligent Transportation Systems; Adaptive Traffic Signal Control; Reinforcement Learning; Growing Neural Gas; Machine Learning**

I. Introduction

In today's urban environments, the problem of traffic congestion is ever increasing due to insufficient traffic infrastructure that is unable to cope with increasing demand. The constraint on available space for new infrastructure makes the traditional approach of building new roads and lanes not feasible (build-only). Modern approaches instead focus on maximizing the efficiency of existing infrastructure based on Intelligent Transport Systems (ITS). Implementation of Adaptive Traffic Signal Control (ATSC) systems can significantly increase the efficiency of signalized intersections by adapting the signal program according to the current traffic situation to improve the operational capacity of the signalized intersection [1], [2].

Efficiency of existing commercial ATSC systems are heavily model dependant, and need to be precisely calibrated to the traffic environment and driver behavior. Because of this the focus of researchers shifted towards Reinforcement Learning (RL) based ATSC systems, which can then learn the appropriate signal control policy according to the desired operational objectives [3]. To adapt RL to ATSC a set of states S and a set of actions A must be defined. The states describe the traffic environment, while actions include every possible change to the signal control regime. RL algorithms such as Q-Learning converge to optimal control policy given enough visits through each state-action combination. The need to describe the traffic environment in more detail increases state-action complexity. This results in the problem of maintaining the convergence rate of the algorithm within acceptable limits [4]. By implementing unsupervised clustering techniques such as Self-Organizing Maps (SOM) or Growing Neural Gas (GNG) the number of analyzed states can be reduced without noticeable negative impacts on the performance [5]. This is demonstrated in this paper by combining GNG with RL to learn the optimal control policy for ATSC.

This paper is organized as follows. The second section covers related work and identifies the contribution of this paper. The principle of using GNG to improve RL convergence is explained in the third section. The fourth section describes the used simulation framework and evaluation scenarios. The obtained results are presented in the fifth section. The last section covers the conclusion and future work discussion.

II. Literature Review

This paper is a continuation of [6] where a SOM was used to reduce state-action complexity in RL-based ATSC, and compared to traditional Q-Learning. The results showed that if the SOM is properly trained the same or even improved RL performance can be achieved with reduced number of states. The question on what is the optimal number of states remained open. Similar approach was used in [7] for the analysis of vehicle trajectories for the classification of intersections.

Combination of GNG and RL was implemented in [8] for car and ride sharing in which each vehicle was an autonomous vehicle controlled by a RL agent attempting to maximize profits by either moving to areas with more requests or remaining in the current area. Another example of using GNG with Q-Learning was shown in [9] where the GNG was used to approximate continuous state space to a discrete one.

Considering the above mentioned, the main contributions of this paper can be summarized as: (i) Application of GNG-RL to traffic signal control; and (ii) Comparison of GNG-RL with similar SOM-RL application to traffic signal control.

978-1-6654-4437-8/21/$31.00 ©2021 IEEE

III. Neural Gas and Reinforcement Learning

This section describes the fundamentals of RL based ATSC, GNG, and the principle on which GNG is implemented to RL in order to reduce state-action complexity in RL based ATSC.

A. Reinforcement Learning in Adaptive Traffic Signal Control

In order to adapt RL to ATSC, the traffic signal controller is considered as agent in a Markov Decision Process (MDP) consisting of a $\langle S, A, T, R\rangle$ tuple, with S being the set of environment states, A the set of traffic signal changes, R as the reward or feedback from the environment after an action is taken, and T as the transition function which describes the change from one state to another [3].

To describe the state space S in RL-based ATSC multiple variables depending on the available data can be used. Most commonly used variables are the queue lengths on each individual intersection approach or lane. The state space is then defined as a continuous vector space $\mathbb{R}^n$ where n is the number of lanes approaching the intersection. As n increases the state space becomes more detailed, but becomes problematic for use in simple RL algorithms such as Q-Learning which deals with discrete states [10]. Possible solutions to this problem are the use of coarse coding, function approximation, or unsupervised clustering techniques such as GNG in this paper to group similar states in regions with each region then being considered as an individual discrete state.

Action set A is the set of possible traffic signal changes an RL agent can take. Several different approaches are used in literature: (i) agent can either increase the duration of the current signal phase, or switch to the next phase [11]; (ii) agent selects from a fixed set of possible signal programs [6]; (iii) agent changes the duration of each phase before signal program execution [12]; and (iv) agent chooses the next phase in sequence on predetermined decision points [13]. In this paper, the agent selects from a fixed set of 4 distinct signal programs.

The reward function R is the feedback the RL agent receives from the environment after completing an action. In RL-ATSC systems, the reward function is defined according to the desired operational objective of the controller. In most cases, the goal is to reduce overall delay of all vehicles, so the reward function can be expressed as the difference between the average delay before and after the action was taken.

With states, actions, and reward function defined the agent can learn the optimal control policy by estimating the value of each state-action pair using the following equation:

$$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha(r_i + \gamma \max_{a' \in A} Q(s',a')), \quad (1)$$

where $Q(s,a)$ is the value of taking action a in state s, r_i is the reward given from the environment after an action was taken, s' is the state in the next learning iteration, a' is the action that would give the highest estimated reward given state s'. Parameters α and γ are learning parameters tuned to specific task. The former being the learning rate, and the latter being the discount factor for future actions.

To select an action the deterministic greedy selection policy can be used to select the action which is estimated to return the highest reward given the current state. Additionally ϵ-greedy selection policy can be used to allow random actions to occur to allow the agent to learn new knowledge. This exploration-exploitation trade-off is usually modelled by gradually reducing ϵ value to promote more exploration in earlier stages of learning and promote more exploitation in the later stages.

B. Self-Organizing Maps and Growing Neural Gas

As mentioned, SOM and GNG can be used to create groups in continuous state spaces in RL to speed up learning. SOM is a type of artificial neural network which uses unsupervised learning to adapt to given input data by changing neuron weights to more closely match the given input signal. The input space is then discretized into n groups where n is the number of neurons, with each neuron serving as a group centroid. Neurons in SOM are interconnected in a grid, thus, creating a map. The topology of SOM adapts to the topology of the input data regardless of the number of dimensions, which makes the SOM a good tool for dimensionality reduction. Upon receiving an input signal the Best Matching Unit (BMU) is identified as the neuron with weights that are the most similar to the input, e.i. the neuron with the shortest Euclidean distance to the input. The input signal is then grouped under the BMU, and the network will then adjust its weights to move closer to the input signal using:

$$W_i(k+1) = W_i(k) + \Theta\alpha_{SOM}\big[X(k) - W_i(k)\big], \quad (2)$$

where $X(k)$ is the input signal, W_i is the weight vector of a neuron i, α_{SOM} is the learning rate, and k is the current learning step. Θ is the neighborhood function which scales the update based on the distance from the BMU using:

$$\Theta = \exp\Big(-\frac{d^2}{2R^2}\Big), \quad (3)$$

where d is the Euclidean distance between the neuron and the BMU, with R being used as the calibration parameter to determine the neighborhood distance.

Removing neuron connections changes a SOM into a Neural Gas (NG) which can also group the data, but without the underlying topological structure. If NG is combined with Hebbian Learning it can learn to form connections between neurons creating a topology [14]. This flexible topology can also allow for simple growing mechanism, creating a GNG. New neurons will be added to areas where other neurons are too far to respond to the input signal allowing the network to grow in size and adapt its topology to the given data.

C. State-Action Complexity Reduction

The principle of data grouping with SOM and GNG can be used to enhance RL based ATSC systems in a structure shown in Fig. 1. Learning of SOM and GNG can be handled before (offline) or during (online) RL training. In the offline learning model, the network is trained using a sample input data set, and

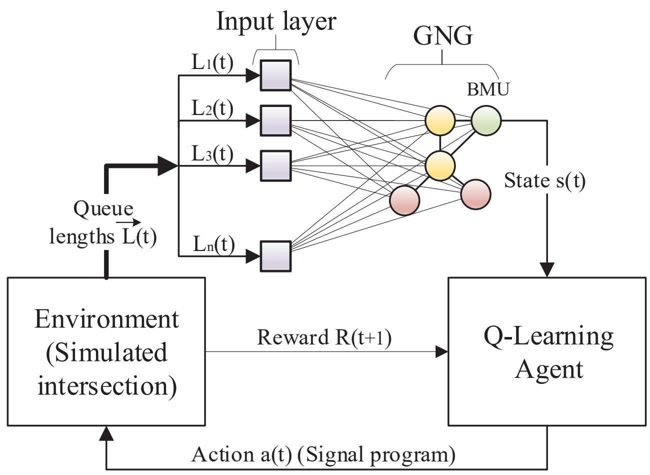

Figure 1. GNG-RL ATSC model

the trained network is then used to map continuous state space to discrete groups to be used in RL. In the online learning approach, the neurons are adapted at the same time as the RL agent learns the optimal control policy. The problem with this approach is that during initial epochs the neurons in SOM tend to change weights rapidly, which can slow down RL learning as neurons move to different areas in the state space. This approach becomes more feasible when used with GNG since only two neurons are present at the start of training, making it easy for the RL agent to find optimal actions. The growing structure of GNG is then exploited when the agent encounters an unfamiliar state by creating a new neuron with the weights equal to the input signal. This has a benefit of allowing the GNG-RL agent to adapt to previously unseen states.

IV. Simulation Model and Scenarios

This paper uses the same simulation model of an isolated intersection modelled in PTV VISSIM as in [6]. The simulated intersectionoperates with Fixed Time Signal Control (FTSC) based on 4 distinct realistic signal programs set to change during the day which is used as a reference point for the performance analysis of the RL based ATSC approaches. The simulation duration was 16.5 hours corresponding to a period from $05{:}30$ AM to $10{:}00$ PM.

Three different scenarios were tested using the same simulation model based on real data. The first tested scenario used the FTSC approach without any external controllers. The second scenario used SOM-RL with best case parameters from [6] with offline training of SOM consisting of 81 neurons arranged in a 9 x 9 grid. The third tested scenario used GNG-RL with online training of GNG. In both SOM-RL and GNG-RL scenario, the exploration-exploitation trade-off was modelled using:

$$\epsilon = 0.95 \cdot 0.98^n + 0.05, \tag{4}$$

where n is the current simulation number. The step time for action selection in both ATSC approaches was set to 300 seconds to allow the agent enough time to receive the reward for the previous action. The reward was modelled to be the change in vehicle delay using the following equation:

$$r_{t+1} = d_t - d_{t+1}, \tag{5}$$

where d_t is the total delay before the selected action was executed, and d_{t+1} is the total delay after the action was taken.

V. Results

In this section the presented approaches are evaluated by analyzing the obtained traffic parameters.

A. Obtained Traffic Parameters

The following Measures of Effectiveness (MoE) were used to analyze the performance of the proposed GNG-RL approach: TTT as the Total Travel Time of all vehicles in the network shown in Fig. 2; LT_{avg} as the average Lost Time of all vehicles in the network; and NS_{tot} as the total Number of Stops of all vehicles in the network shown in Fig. 3. Each analyzed scenario was run for 500 epochs, with each epoch consisting of 198 learning iterations. The average results of the last 100 simulations are presented in detail in Table I. In addition, the standard deviation (σ) of each analyzed MoE for the last 100 simulations is calculated in order to compare ATSC system stability.

B. Discussion

The obtained results show that both SOM-RL and GNG-RL ATSC can significantly reduce the TTT. Both approaches converge in line with the specified exploration-exploitation trade off function, but it is observed that SOM-RL converges to lower values of TTT. The results for Lt_{avg} are proportional to the results for TTT which shows that the reduction is TTT is observed due to reduced vehicle lost time while interacting with the environment. The NS_{tot} was only slightly reduced in SOM-RL scenario, and slightly increased in GNG-RL scenario. The cause of this is that both agents learned to apply shorter signal programs which might cause more vehicles to stop before the intersection. However, the stopping duration was decreased since the phase changes occurred more frequently. The SOM-RL scenario had 91 neurons from the beginning of training, while the GNG-RL finished training with 63 neurons. Thus, it is possible to achieve similar, albeit a bit worse results with the lower number of states. One has to note that in the case of GNG-RL no a priory analysis regarding needed number of neurons has to be done making this approach easier to implement.

VI. Conclusion and Future Work

In this paper, the use of GNG for state complexity reduction in RL based ATSC is analyzed and compared with a similar approach using SOM. The results show that GNG-RL approach can achieve similar reduction in TTT and LT_{avg} as the SOM-RL approach. The final number of neurons in GNG was lower than the number used in SOM proving that the RL agent can achieve similar results even with a reduced state space. The addition of neurons in GNG was based on neuron

TABLE I. Mean results for all scenarios from simulation 400 to simulation 500

MoE	FTSC	SOM-RL			GNG-RL		
		Obtained	Change [%]	σ	Obtained	Change [%]	σ
TTT $[h]$	856.89	768.34	-10.33	7.28	777.29	$-$ 9.29	10.15
LT_{avg} $[s]$	29.59	23.72	-19.84	0.38	24.22	-18.15	0.50
NS_{tot}	48135	47875	$-$ 0.54	1027	48537	0.84	1381

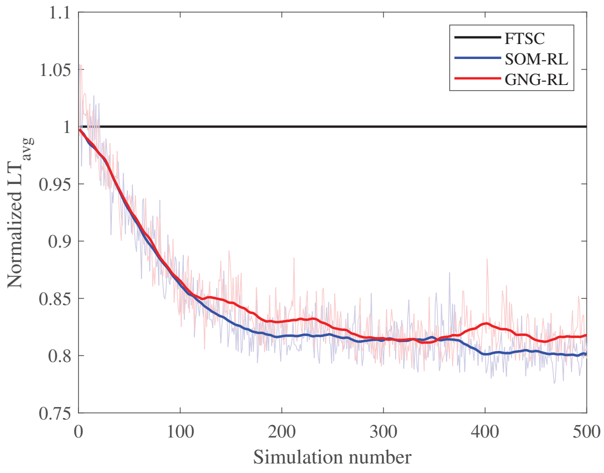

Figure 2. Normalized TTT for all tested scenarios

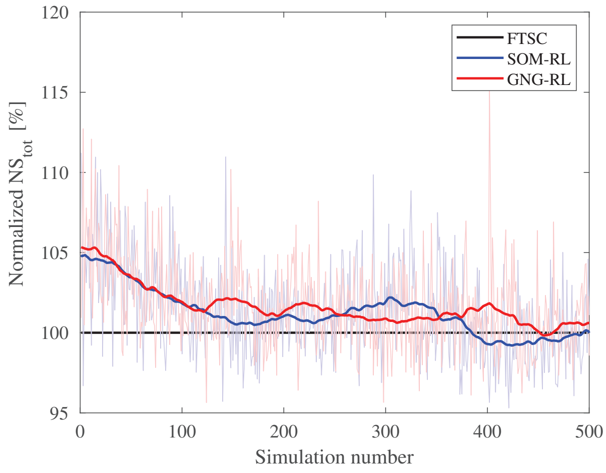

Figure 3. Normalized NS_{tot} for all tested scenarios

distances in the state space, resulting in multiple neurons with similar weights using the same identified optimal action. Hence, further analysis should be directed towards the transfer of learned control policy between connected neurons, since it is expected that similar states might use the same optimal action. Additionally the growth of GNG could depend not only on the observed states, but learned actions, since there is no need to add new neurons if the existing neurons adequately handle the state space area.

Acknowledgement

This research has been carried out within the activities of the Centre of Research Excellence for Data Science and Cooperative Systems supported by the Ministry of Science and Education of the Republic of Croatia, and the University of Zagreb Support Program for scientific and artistic research short-term support: “Innovative control strategies for sustainable mobility in smart cities". This work has been partly supported by the Croatian Science Foundation under the project IP-2020-02-5042, and by the European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATACROSS). The authors Daniela Koltovska Nečoska and Edouard Ivanjko received ERASMUS+ mobility grants for academic collaboration.

References

[1] D. Pavleski, M. Miletić, D. K. Nečoska, and E. Ivanjko, *In-Depth Evaluation of Reinforcement Learning Based Adaptive Traffic Signal Control Using TSCLAB*, pp. 49–64. Cham: Springer International Publishing, 2021.

[2] R. Marsetič, D. Šemrov, and M. Žura, “Road artery traffic light optimization with use of the reinforcement learning,” *Promet - Traffic & Transportation*, vol. 26, pp. 101–108, Apr. 2014.

[3] P. Mannion, J. Duggan, and E. Howley, “An Experimental Review of Reinforcement Learning Algorithms for Adaptive Traffic Signal Control,” in *Autonomic Road Transport Support Systems*, Autonomic Systems, pp. 47–66, Cham: Springer International Publishing, 2016.

[4] 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*, vol. 10, no. 11, 2020.

[5] L. Vojáček and J. Dvorský, “Growing neural gas – a parallel approach,” in *Computer Information Systems and Industrial Management* (K. Saeed, R. Chaki, A. Cortesi, and S. Wierzchoń, eds.), (Berlin, Heidelberg), pp. 408–419, Springer Berlin Heidelberg, 2013.

[6] M. Miletić, K. Kušić, M. Gregurić, and E. Ivanjko, “State complexity reduction in reinforcement learning based adaptive traffic signal control,” in *2020 International Symposium ELMAR*, pp. 61–66, 2020.

[7] Y.-C. Kao and C.-W. Wu, “A Self-Organizing Map-Based Adaptive Traffic Light Control System with Reinforcement Learning,” in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pp. 2060–2064, Oct. 2018. ISSN: 1058-6393.

[8] M. Guériau, N. Cardozo, and I. Dusparic, “Constructivist approach to state space adaptation in reinforcement learning,” in *2019 IEEE 13th International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, pp. 52–61, 2019.

[9] M. Baumann and H. K. Büning, “Adaptive function approximation in reinforcement learning with an interpolating growing neural gas,” in *2012 12th International Conference on Hybrid Intelligent Systems (HIS)*, pp. 512–517, 2012.

[10] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book, 2018.

[11] S. El-Tantawy and B. Abdulhai, “Multi-Agent Reinforcement Learning for Integrated Network of Adaptive Traffic Signal Controllers (MARLIN-ATSC),” in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 319–326, 2012.

[12] X. Liang, X. Du, G. Wang, and Z. Han, “A deep reinforcement learning network for traffic light cycle control,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1243–1253, 2019.

[13] I. Arel, C. Liu, T. Urbanik, and A. Kohls, “Reinforcement learning-based multi-agent system for network traffic signal control,” *IET Intelligent Transport Systems*, vol. 4, no. 2, pp. 128–135, 2010.

[14] Y. Prudent and A. Ennaji, “An incremental growing neural gas learns topologies,” in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, vol. 2, pp. 1211–1216 vol. 2, 2005.