THE OPEN JOURNAL OF MATHEMATICAL SCIENCES AND APPLICATIONS
ISSN 0000-0000
(https://openjournal.utar.edu.my/index.php/cmsojmsa)

Generative Adversarial Networks-Based Reinforcement Learning for Traffic Signal Control

Yong Siang Lok[1], Kok Lim Yau[*2], Hooi Ling Khoo[2], and Farank Nejati[2]

[1]*Postgraduate student (Master of Engineering Science), Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Malaysia*
[2]*Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Malaysia*

Received: *, accepted: *, published online: *.

Abstract

Reinforcement learning (RL) agents often struggles with generalization, performing poorly when encountering unseen state that do not exist during training. This limitation stems from their reliance on trial-and-error learning, where decision-making is guided primarily by reward from exprerienced states. In safety-critical application like traffic signal control, failure to handle unseen traffic scenarios can lead to suboptimal and unsafe actions. This study addresses this challenge by generating and incorporating unseen states into training of RL-based traffic signal controllers. By exposing agents with diverse and atypical traffic scenarios, the approach enchances roboustness and adaptibaility, enabling more reliable performance under real-world uncertainties. This paper presents the review of models, algorithms, applications, and open issues of both reinforcement learning and generative adversarial networks in addressing generalization to unseen state for RL model deployment.

Keywords: traffic signal control; reinforment learning; generative adversarial network.

Math. Subj. Class. (2020):

[*] Corresponding author: yaukl@utar.edu.my

1 Introduction

Reinforcement learning (RL) based traffic signal controller are typically trained using the historical traffic scenarios from real-world settings or standard simulation datasets in traffic simulators. However, these training traffic scenario are often limited in capturing rare but critical negative traffic scenarios (e.g. accidents, road blockages, or extreme congestions). This limitation creates challenges in generalization as RL-traffic signal controller during deployment as it may struggles to handle the state that fall outside the training distribution which predominantly composed of typical traffic conditions (Yau et al., 2017). Consequently, the applicability of RL-based traffic signal controller in the real-world settings is restricted, particulary whrn dealing with traffic uncertainties that deviate from the training distribution. For simplicity, we define such critical negative traffic scenarios that lie outside the typical training distribution as unseen states. To improve the generalization of RL-based traffic signal controller during deployment, it is essential to expose RL agents to these unseen states and incorporate them into the training state distribution.

A promising method to expose RL agents to unseen states is to formulate the problem as a data generation task, where the generative adversarial networks (GANs) model is employed to synthesize traffic scenarios that represent unseen state that differing from typical traffic scenarios. These synthetic unseen data serve as additional traning scenarios for RL-based traffic signal controller. To ensure effective learning, the generated scenarios must be designed with additioanl constraints, such that the synthetic unseen state not only about realism but also need to adapt to wide range of possible RL agent actions. This ensures that the training distributions sufficiently cover wide range of state-action distribution which support the generalization of RL-based traffic signal controller during real-world deployment.

The obejective of this research focuses on: a) to develop the generative adversarial networks based reinforment learning (GAN-RL) model for traffic signal controller under both typical and unseen states; b) to develop the generative adversarial networks (GANs) model for unseen state generation; c) to evaluate the performance of GAN-RL under typical and unseen states.

2 Literature Review

This section provides a review of models, algorithms, and open issues related to the application of reinforcement learning (RL) and generative adversarial networks (GANs) in traffic control and scenario generation.

The RL-based traffic signal controller functions as decision maker during deployment and as a trial-and-error model during training. In both cases, the aim is to optimizing the traffic sginal timing by selecting the high reward action from the enviroment which may represented by either real-world traffic or simulation model. RL-based traffic signal control model can be designed as single-agent (SARL) and multi-agent reinforcement learning model. A commom approach is to assign one agent to control signal timing of a single intersection, however, some studies have investigate the use of a singel agent to manage multiple intersection (Aslani et al., 2020). In MARL, traning can generally categorized into two training scenarios: a) centralized learning and b) decentralized learning. Centralied learning relies on a global memory that collects and store experiences, state and rewards from all agents and followed by synchronized updates. An example of such global memory is the knowledge container, which aggregates the agents experiecence, while there are approach to reduce the scalability issues using the computational global memory where each agent contributes to a local memory known as local cooperative factor (Li et al., 2021). In this case, weights are assigned to each agent to determine the extend of the contribution to the global memory. However, training speed may relatively slow as it required completion of agents. In decentralized learning, training is more faster because each agents updates its policy individually without relying on global synchronization. However, this create a non-stationary environment, as different update timing and limited communication among the agents. To address this challenge, techniques such as graph convolutional network, graph attention mechanisms (Wang et al., 2022) and feature embedding method have been employed to convert all the agent current or past experience to update signal. Previous RL algorithms proposed for traffic signal control focused on three aspects: exploration, exploitation and balancing between the two. Exploration can be archieve by policy based methods, which search across wide range of possible signal timing, while exploitation aims to identify optimal signal timing under specfic constraints, like certain road may need a minimun signal duration. Actor-critic algorithm applies to balance both exploration and exploitation.

The generative adversarial network (GANs) model consists of two compoentns: a generator to generate traffic scenarios, and a discriminator to identify the input scenario is the real or fake (Goodfellow et al., 2014). The GAN model leverage the adversarial training where generator and discriminator are optimized in opposite way. The training objective is to minimize the statistical distance between real and synthetic traffic scenarios. GAN architecture can be build upon various deep learning networks, but in traffic application it is primarily designed to capture the spatio-temporal patterns of traffic state data. For example, convolutional neural networks (CNNs) for image based data, recurrent neural networks (RNNs) for temporal based data, graph neural networks (GNN) for graph structural data specifically the spatial based data and hybrid architecture that combined these deep learning model for spatial-temporal modeling. Variants of GANs can be distinguishing based on the loss function. For example, the traditional GAN employs a minimax loss function, Wassertein GAN (WGAN) using Wasserstein distance to improve training stability, Bayesian GAN incorporate Bayesian inference and maximum likelihood estimation, generative adversarial imitation learning (GAIL) uses the discriminator output as value function for policy optimization, relativistic GAN insights into how close between two distribution with respect to a reference, conditional GAN (cGAN) extends the framework with conditional generation, and cycle GAN employs cycle consistency loss to domian translation. In the traffic scenario generation, the application can divide into three task: a) data imputation, focusing on reconstruct missing traffic state data from neighbor information, b) data generation, focusing on converting noise to new data, c) data prediction, focusing on learning features from the past.

3 Open issues

This sections discussed the open issue related to unseen state training for reinforcement learning (RL) model and unseen state data generation for generative adversarial networks (GANs). In past research on RL-based traffic signal controller, most studies focused on developing different model and algorithm to achieve high performce in optimizing the trafifc signal under typical traffic condition. However, the consideration of unseen state has been limited. One possible reason is the lack of available dataset that capture such rare and critical scenarios, nevertheless the unseen satte could serve as valuable training scenarios for Rl agents to generalize better to handle real-world traffic uncertainties. The lack of available dataset is mainly due to expensive data collection in real world and limitation of rule-based traffic simulator. Consequently, these challenges also extend to GAN models, as no existing datasets are available to directly train a GAN for unseen states. To address such data limitations, it is essential to first identify what types of scenarios the GAN shoule generate, particularly in terms of critical and rare traffic conditions that are differering form typical traffic conditions. Unseen state can often classified as abnormal traffic condition, howeber, simply generate such abnormal traffic condition may not enough. For example, a trained RL-based controller may learn the rule to extending the green light duration to reduce overall waiting times on the road. Howeverm this actions become ineffective if the road is fully blocked by an obstacles, as simply extend the green light duration is does not resolve the congestion. This example highlights the generating abnormal traffic data is insufficient, and there is the need to generate the root cause (an incident) and its implication toward the traffic conditions. This leads to the conclusion that the GAN model should be capable of generating the local event along with its implication on the current traffic condition. Even after redefining the unseen state as local events and their implications, the challenge of absence of datasets for unseen state generations stilles remains and must be addressed. Drawing inspiration from driving behavior when facing obstacles, a driver typically take three keys action, rerouting, controlling vehicle to follow new path, and monitoring surrounding vehicles. This actions can be conceptually divided into three categories: path planning, trajectory planning and scheduling. Based on above concept, the pure generation task can divided into three subtasks: a) path planning, b) trajectory planning, and c) scheduling.

4 Conclusions

This paper highlights the generalization challenge of RL-based traffic signal controller when deployed in state that fall outside the training distribution. Providing diverse training scenarios for RL-based traffic signal controller can be an alternative solution proposed. To this end, generative adversarial network (GAN) is introduced to generate the unseen state as training scenarios for RL-based traffic signal controller. However, one major challenges lies in the not available dataset for unseen state need to be addressed. To address such challenges, the unseen state is redefined into local events and implications on traffic conditions. Next, drawing inspiration from driving behavior in obstalce-avoidance contexts, the generative task is divided into path, trajectory model and scheduling model.

References

Aslani, M., Mesgari, M. S. and Wiering, M. (2020) 'Adaptive traffic signal control with actor-critic methods in a real-world traffic network with different traffic disruption events', *Transportation Research Part C: Emerging Technologies* **119**, 102741.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014) 'Generative adversarial networks', *arXiv preprint arXiv:1406.2661 [stat.ML]* . Available at: https://doi.org/10.48550/arXiv.1406.2661.

Li, Z., Yu, H., Zhang, G., Dong, S. and Xu, C. Z. (2021) 'Network-wide traffic signal control optimization using a multi-agent deep reinforcement learning', *Transportation Research Part C: Emerging Technologies* **125**, 103059.

Wang, M., Wu, L. and Li, J. (2022) 'Traffic signal control with reinforcement learning based on region-aware cooperative strategy', *IEEE Transactions on Intelligent Transportation Systems* **23**(7), 3315–3325.

Yau, K. L. A., Qadir, J., Khoo, H. L., Ling, M. H. and ND Komisarczuk, P. (2017) 'A survey on reinforcement learning models and algorithms for traffic signal control', *ACM Computing Surveys* **50**(3).