Congestion pricing in multimodal networks: an application of deep reinforcement learning

Nasser Parishad[a], Mehmet Yildirimoglu [a*]

[a] The University of Queensland, School of Civil Engineering, Brisbane QLD 4072, Australia

September 02-03, 2024, Crete, Greece

April 29, 2024

Keywords: Congestion Pricing, Trip-based MFD, Mode Choice, Reinforcement Learning, DDQN

1 INTRODUCTION

Traffic congestion is a major global problem that costs billions of dollars every year. Congestion pricing, which entails higher costs for car users, can incentivize the adoption of alternative departure times (Liu et al., 2023), routes (Li and Ramezani, 2022), or modes of transportation (Balzer and Leclercq, 2022).

To study congestion pricing on a large scale, Macroscopic Fundamental Diagrams (MFDs) offer a parsimonious regional-scale representation of urban traffic dynamics (Geroliminis and Daganzo, 2008). While traditional accumulation-based MFD models have been effective for real-time control applications, they lack granularity and do not reflect the varying circumstances across travellers. Trip-based MFD models address this limitation and offer improved modeling capabilities by providing different trip lengths for every traveler (Lamotte and Geroliminis, 2016). However, the absence of a closed-form analytical expression in trip-based MFDs besides the challenges in accurately estimating MFD parameters (Saffari et al., 2022) necessitates further research in this area.

To overcome these challenges, data-driven model-free approaches have been proposed (Wang and Paccagnan, 2022). However, they also require accurate estimation of network parameters such as critical accumulation. Given that the estimation of these parameters is susceptible to errors due to multivaluedness, instability, and hysteresis phenomena commonly observed in practical network scenarios (Daganzo et al., 2011; Mahmassani et al., 2013), further research is necessary to develop a more robust and reliable method.

Reinforcement Learning (RL) has emerged as a promising solution for addressing complex challenges (Sutton and Barto, 2018). Thanks to its model-free nature, it requires minimal prior knowledge of environmental dynamics. The introduction of the Deep Q-Networks (DQN) algorithm (Mnih et al., 2015) and Double Deep Q-Network (Van Hasselt et al., 2016) enable researchers to harness these algorithms for solving large and complex problems.

This study will examine the impact of pricing on travelers' mode choices, considering an elastic demand and modelling travelers' choices based on the perceived cost of all available alternatives. By utilizing a trip-based MFD as the environment for Reinforcement Learning (RL) and employing the DDQN as the RL agent, this paper aims to develop a data-driven method to determine an optimal dynamic toll profile. The objective is to maximize network throughput or minimize total time spent in a multimodal network by encouraging/discouraging the use of certain modes via an imposed toll price.

2 Traffic Dynamics and Modeling

2.1 Traffic Dynamics

A trip based MFD framework has been developed for a single region, representing a congested city center. All trips originate and end within the network. The simulation considers travelers' mode choice and offers three main modes: car, public transport (bus), and walking. Figure 1 illustrates the link-level representation of a segment of the network for each mode considered in this study. The network is structured as a grid, with dimensions of 10 by 10. While car and walking modes can effectively use the entire network, public transport stops are located every three nodes both in the horizontal and vertical directions. This enables a realistic representation of a multimodal network. Note that the link-level representation of the network will be used to determine the shortest path and calculate the trip distance and travel time associated with each mode.

A 90-minute demand profile projecting traffic congestion is considered in this study. The simulation iterates through the demand list with a step size of one minute. Traveler heterogeneity is incorporated by randomly generating origin, destination, and departure time for each traveler to be considered. The travel time (T_i^{car}) of the car user i, with trip length l_i^{car}, departure time t_i, and the speed V_i^{car} calculated by:

$$l_i^{car} = \int_{t_i}^{t_i+T_i^{car}} V_i^{car} dt \qquad (1)$$

The buses are assumed to operate on a designated line with a fixed headway. A reasonable constant speed is considered for both the bus and walking modes based on literature.

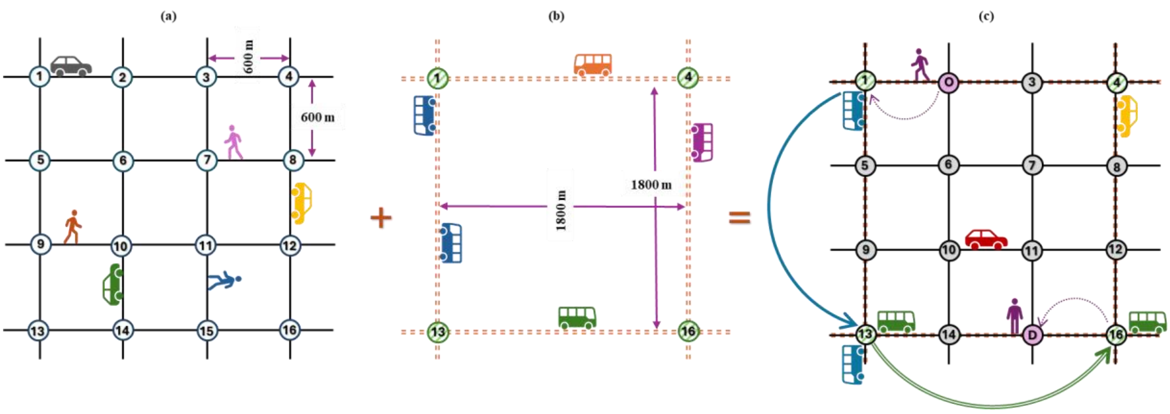

Figure 1. A portion of network layout, a) car/pedestrian network, b) public transport network, c) integrated network

2.2 Mode choice

Below are the monetary costs associated with using different modes. The decision-making process follows a multinomial logit model, which assumes that users perceive costs independently, with an added error term following a Gumbel distribution.

$$\begin{cases} C_i^{car} &= -\alpha T_i^{car} - \tau \\ C_i^{walk} &= -\alpha T_i^{walk} \\ C_i^{bus} &= -\alpha (T_i^{bus} + T_i^{wait} + T_i^{Transit} + T_i^{access}) \end{cases} \qquad (2)$$

Where α is the Value of Time (VoT) and τ represents the toll price. T_i^{wait}, T_i^{Transit} represent the required time that traveler i must wait for the bus and the required time for traveler i to change direction and board another bus, respectively. Both are uniformly distributed between 0 and the bus headway (Daganzo, 2010). Additionally, T_i^{access} equals sum of entrance and egress time.

3 RL Formulation

A double deep Q-network (DDQN) agent, along with a prioritized experience replay (PER) buffer, is used to learn a policy that maximizes the expected return from the start time. Figure 2 provides a schematic diagram, illustrating the operation of DDQN. The structure of the reinforcement learning (RL) problem is defined as follows:

- **State space:** at each time step the agent considers the demand value of the next time step, car accumulation, bus accumulation, pedestrian accumulation, previously applied toll as input data.
- **Action space:** the agent is allowed to take three different actions: increasing toll value by $\Delta\tau$, decreasing toll value by $\Delta\tau$, and keep the toll value constant.
- **Transition dynamics:** these are expressed by traffic dynamics described in section 2.1.
- **Reward function:** the scalar reward is defined as the normalized value of network outflow at each step.

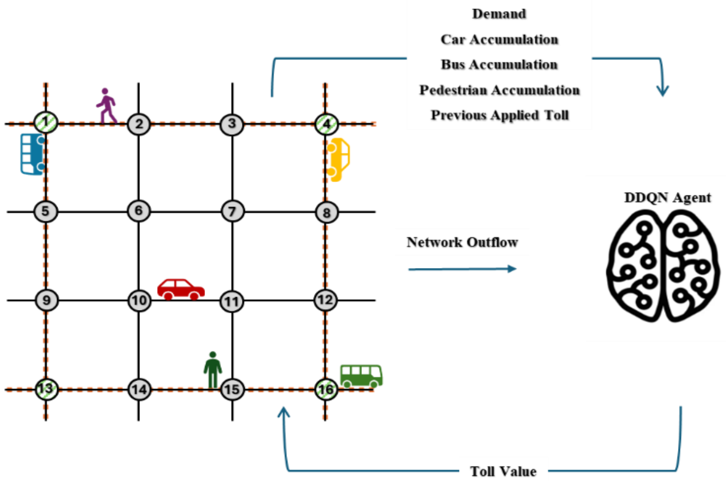

Figure 2. RL structure

4 Results and Discussion

Figure 3.a depicts the evolution of the agent. Over 250 episodes, the agent reduces total travel time by approximately 30%. To maintain car accumulation at critical levels and maximize network outflow in the car network (Figure 3.b), the agent applies a toll profile, as shown in Figure 3.c.

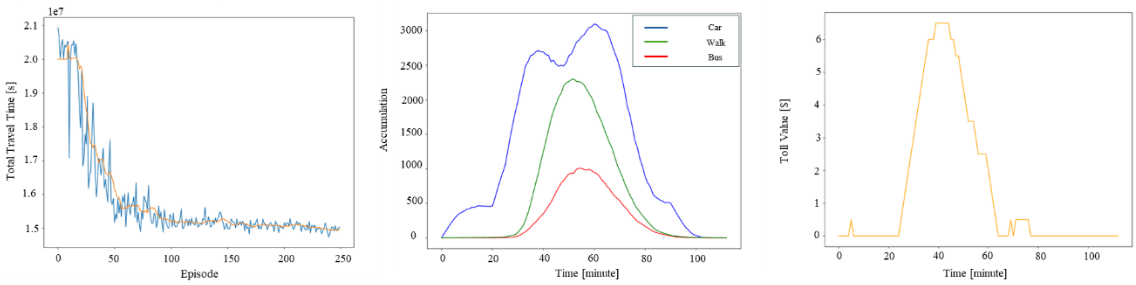

Figure 3. (a) the agent's evolution, (b) accumulation of each mode at the final stage, (c) toll profile.

After training, the agent was tested under various conditions, and a sensitivity analysis was conducted on all input data. The results indicate that even with a 20% error in the provided input data, the agent can still find the optimum toll profile without impediment. Furthermore, after training, the agent was tested under two different conditions: first, with a completely different demand profile, and second, with entirely different MFD coefficients. In both cases, the agent successfully reduced congestion and operated at critical accumulation levels.

5 REFERENCES

Balzer, L., & Leclercq, L. (2022). Modal equilibrium of a tradable credit scheme with a trip-based MFD and logit-based decision-making. Transportation Research Part C: Emerging Technologies, 139, 103642

Daganzo, C. F. (2010). Structure of competitive transit networks. Transportation Research Part B: Methodological, 44(4), 434-446.

Daganzo, C. F., Gayah, V. V., & Gonzales, E. J. (2011). Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. Transportation Research Part B: Methodological, 45(1), 278-288.

Geroliminis, N., & Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. Transportation Research Part B: Methodological, 42(9), 759-770.

Lamotte, R., & Geroliminis, N. (2016). The morning commute in urban areas: Insights from theory and simulation (No. 16-2003).

Li, Y., & Ramezani, M. (2022). Quasi revenue-neutral congestion pricing in cities: Crediting drivers to avoid city centers. Transportation Research Part C: Emerging Technologies, 145, 103932.

Liu, R., Chen, S., Jiang, Y., Seshadri, R., Ben-Akiva, M., & Lima Azevedo, C. (2023). Managing network congestion with a trip-and area-based tradable credit scheme. Transportmetrica B: Transport Dynamics, 11(1), 434-462

Mahmassani, H. S., Saberi, M., & Zockaie, A. (2013). Urban network gridlock: Theory, characteristics, and dynamics. Procedia-Social and Behavioral Sciences, 80, 79-98.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. Nature 518, 529–533.

Saffari, E., Yildirimoglu, M., & Hickman, M. (2022). Data fusion for estimating Macroscopic Fundamental Diagram in large-scale urban networks. Transportation Research Part C: Emerging Technologies, 137, 103555.

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction. MIT Press.

Van Hasselt, H., Guez, A., & Silver, D. (2016, March). Deep reinforcement learning with double q-learning. In Proceedings of the AAAI conference on artificial intelligence (Vol. 30, No. 1).

Wang, Y., & Paccagnan, D. (2022, December). Data-driven robust congestion pricing. In 2022 IEEE 61st Conference on Decision and Control (CDC) (pp. 4437-4443). IEEE.