



HAL
open science

Decentralized Traffic Signal Control with Deep Reinforcement Learning in a Multi-modal Network: Seeking for a Trade-off between Bus Service and Traffic

Jiajie Yu, Pierre-Antoine Laharotte, Ludovic Leclercq

► **To cite this version:**

Jiajie Yu, Pierre-Antoine Laharotte, Ludovic Leclercq. Decentralized Traffic Signal Control with Deep Reinforcement Learning in a Multi-modal Network: Seeking for a Trade-off between Bus Service and Traffic. TRB Annual Meeting 2023, Jan 2023, Washington, United States. hal-04638362

HAL Id: hal-04638362

<https://hal.science/hal-04638362v1>

Submitted on 8 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

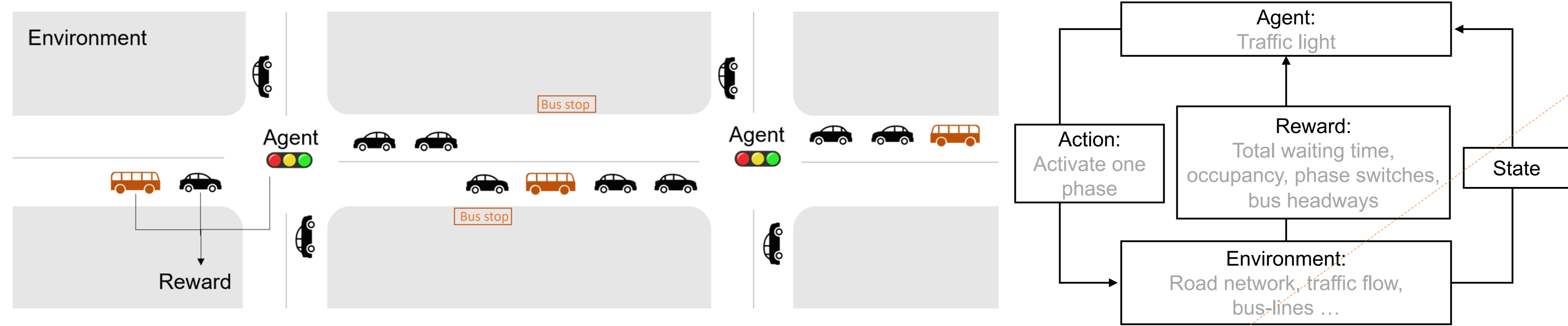
Decentralized Traffic Signal Control with Deep Reinforcement Learning in a Multi-modal Network: Seeking for a Trade-off between Bus Service and Traffic

Motivation:

- Unbalanced bus headway leads to long waiting time for passengers.
- Classical bus holding strategies may not be well-perceived by users.
- Holding can be achieved silently at traffic signals with proper control.

A decentralized Deep Reinforcement Learning (DRL) traffic signal control framework is proposed to increase traffic efficiency and equalize bus headways simultaneously.

Model formulation:



State

- Total waiting time $D_{i,t}$
- Occupancy $O_{i,t}$
- Forward and backward bus headways $h_{i,t}^f, h_{i,t}^b$
- Number of yellow phases in last 10 decision steps $n_{i,t}^y$

Action

- Predefined phase
- Yellow phase for switch

$$\text{Reward } r_{i,t} = r_{i,t}^{\text{traffic}_1} + r_{i,t}^{\text{traffic}_2} + r_{i,t}^{\text{agent}} + c \times r_{i,t}^{\text{bus}}$$

Reduction degree of total waiting time

$$r_{i,t}^{\text{traffic}_1} = \begin{cases} 1 & , D_{i,t} < D_{i,t-1} \\ 0 & , D_{i,t} = D_{i,t-1} \\ -1 & , D_{i,t} > D_{i,t-1} \end{cases}$$

Occupancy of incoming lanes

$$r_{i,t}^{\text{traffic}_2} = \begin{cases} 0 & , O_{i,t} < 0.5 \\ -1 & , O_{i,t} > 0.5 \end{cases}$$

Stability of decisions taken

$$r_{i,t}^{\text{agent}} = \begin{cases} 0 & , n_{i,t}^y \leq 2 \\ -1 & , n_{i,t}^y > 2 \end{cases}$$

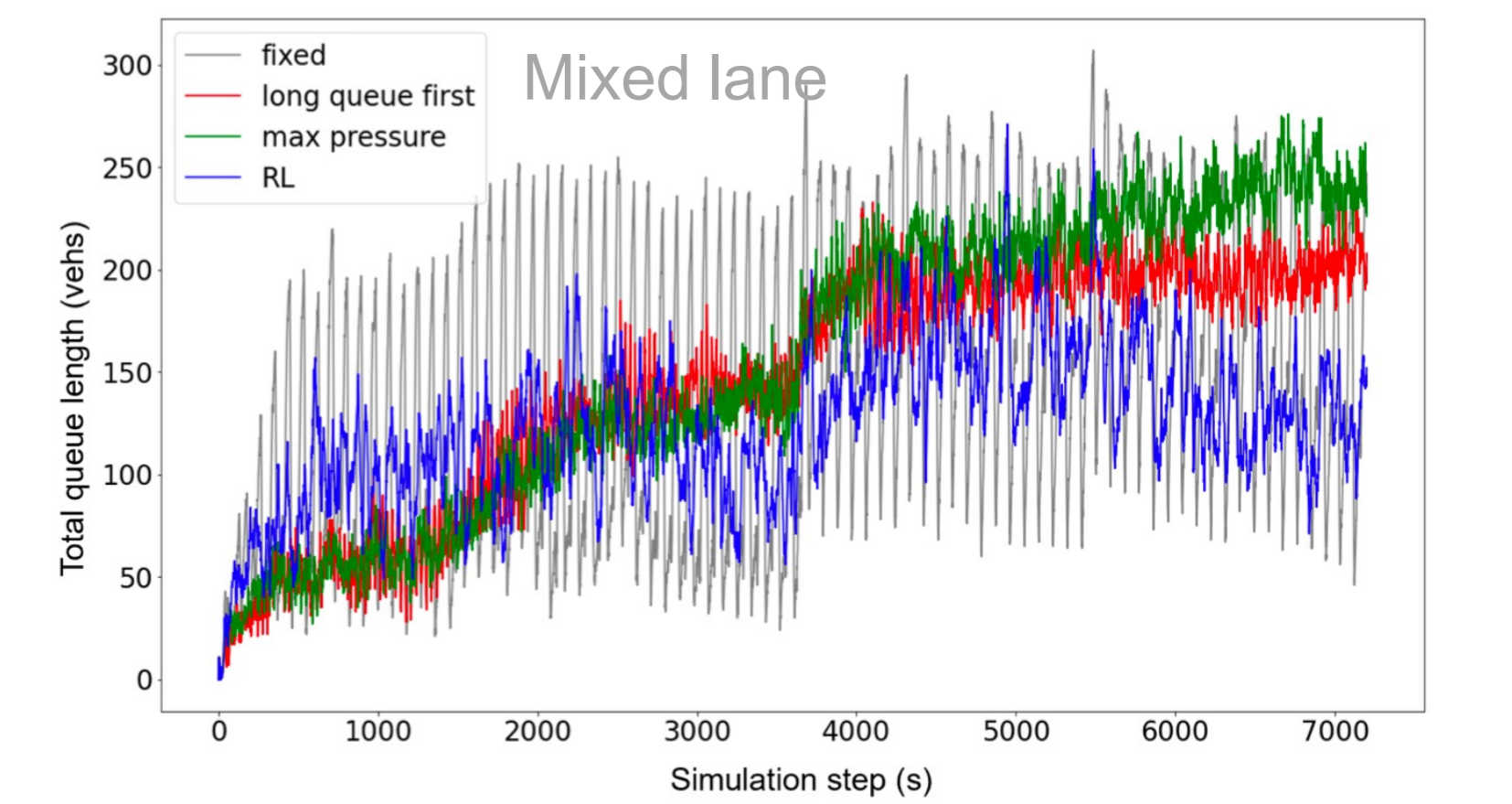
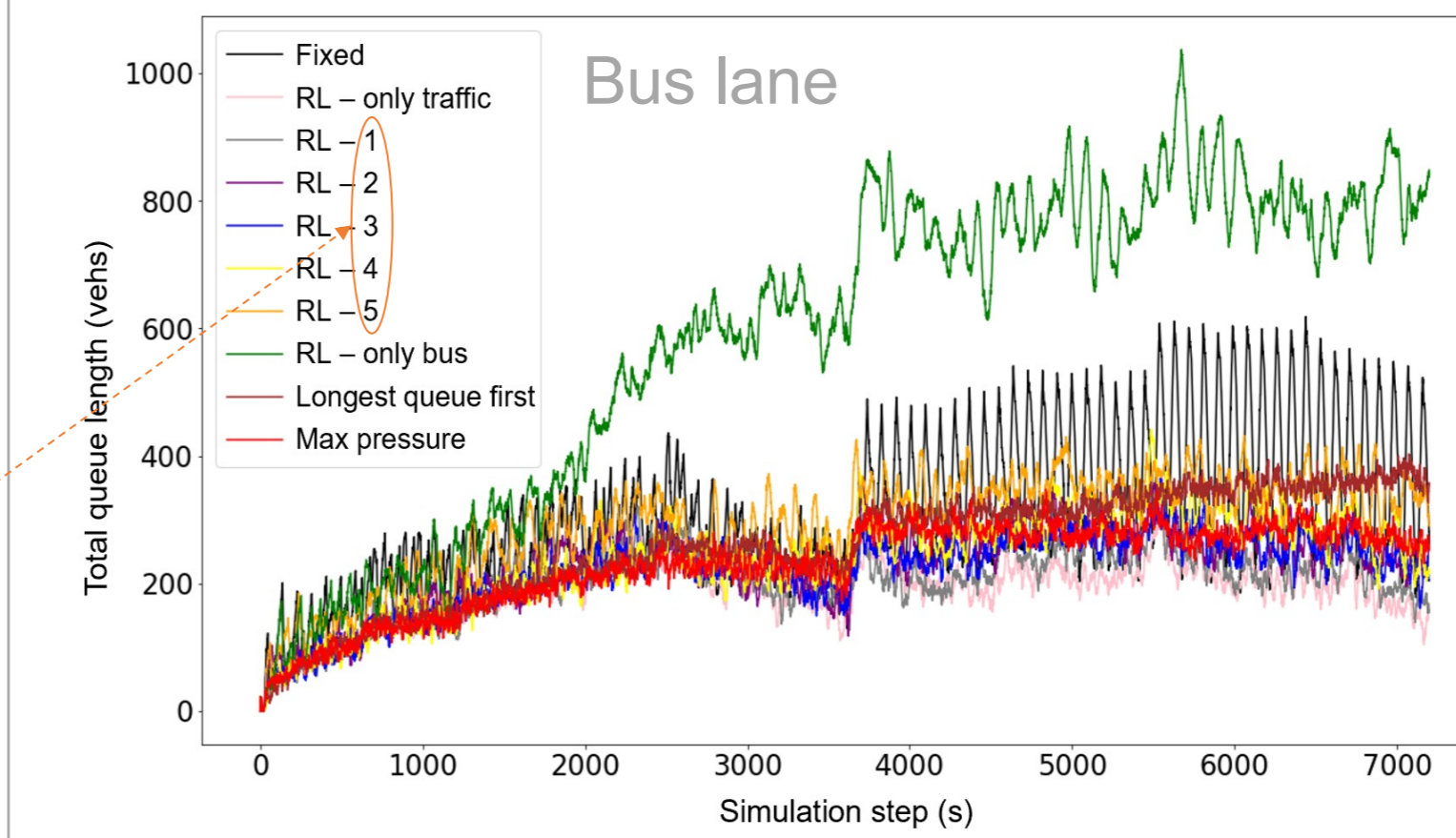
Bus headway control performance

$$r_{i,t}^{\text{bus}} = \begin{cases} 1 & , h_{i,t}^f > h_{i,t}^b \text{ and bus gets green} \\ 1 & , h_{i,t}^f < h_{i,t}^b \text{ and bus gets red} \\ 0 & , h_{i,t}^f = h_{i,t}^b \\ -1 & , \text{otherwise} \end{cases}$$

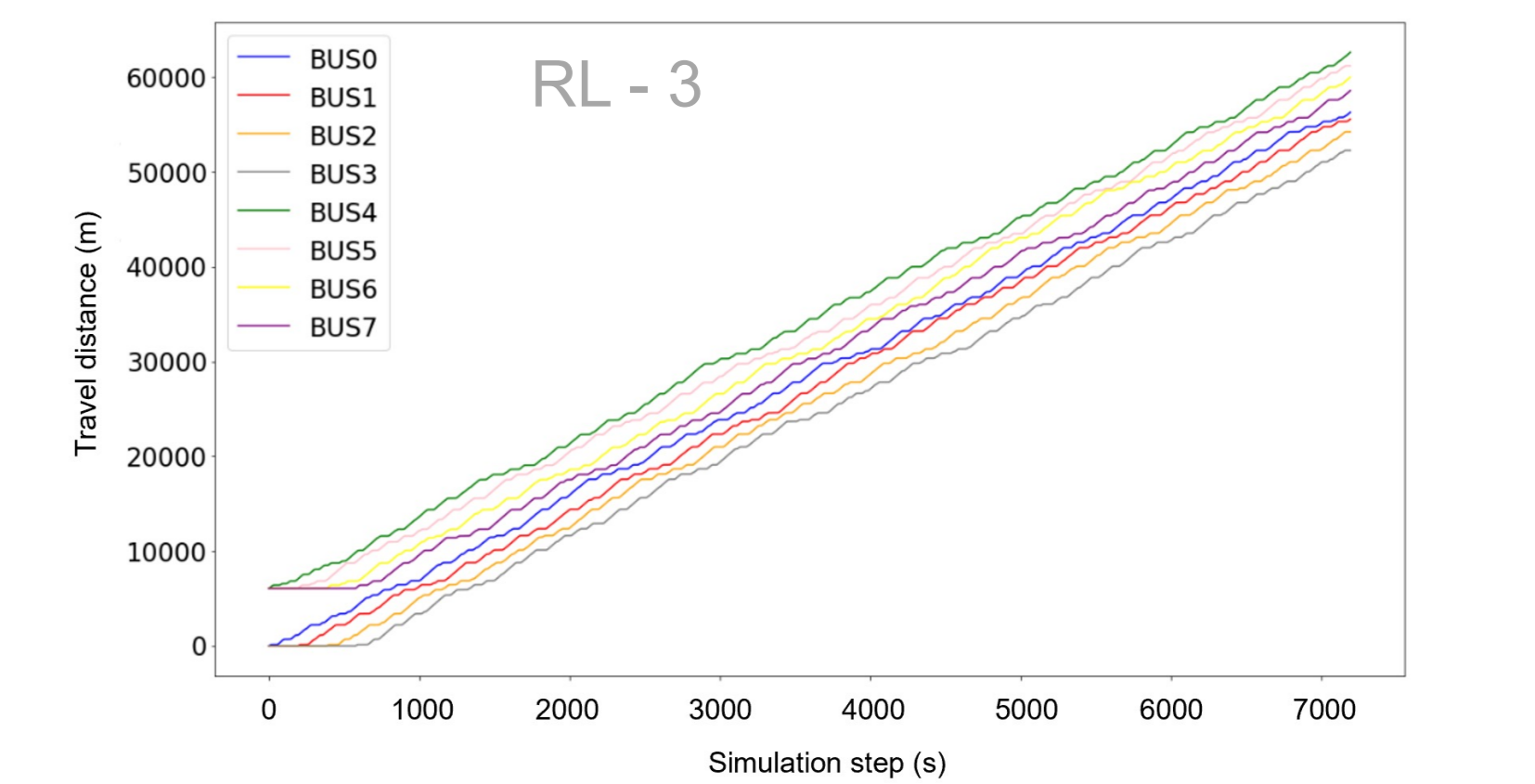
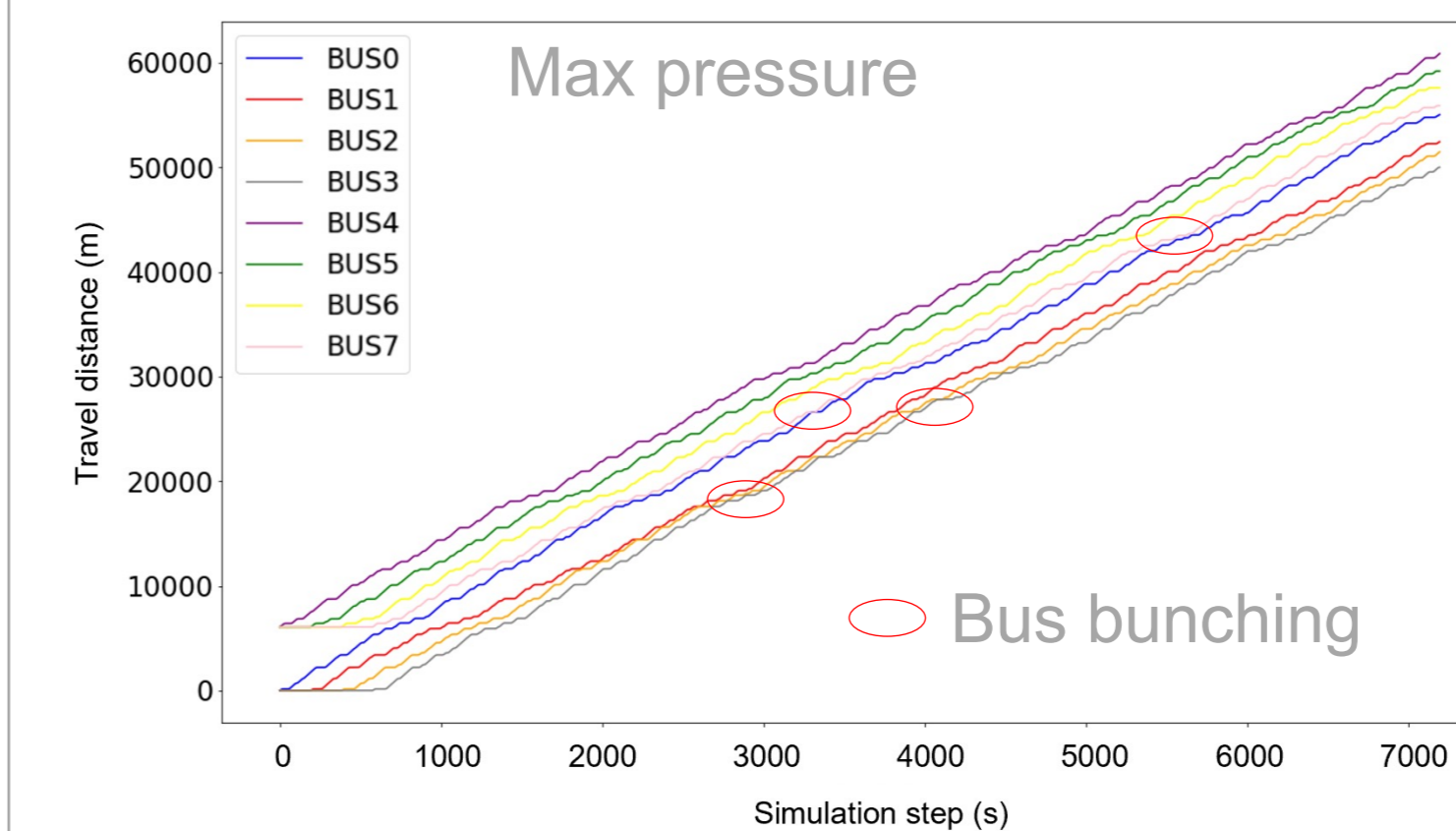
Numerical tests:



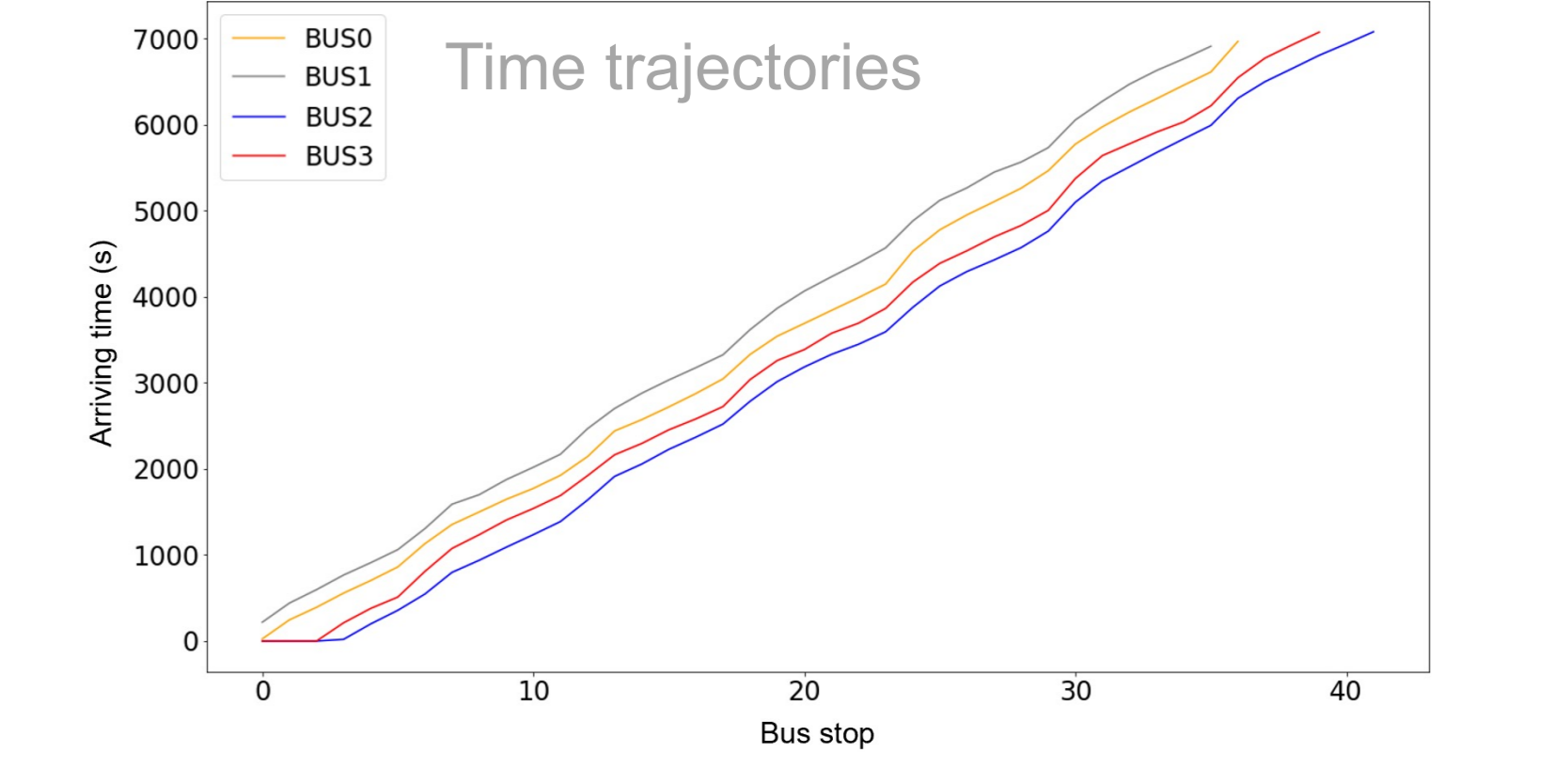
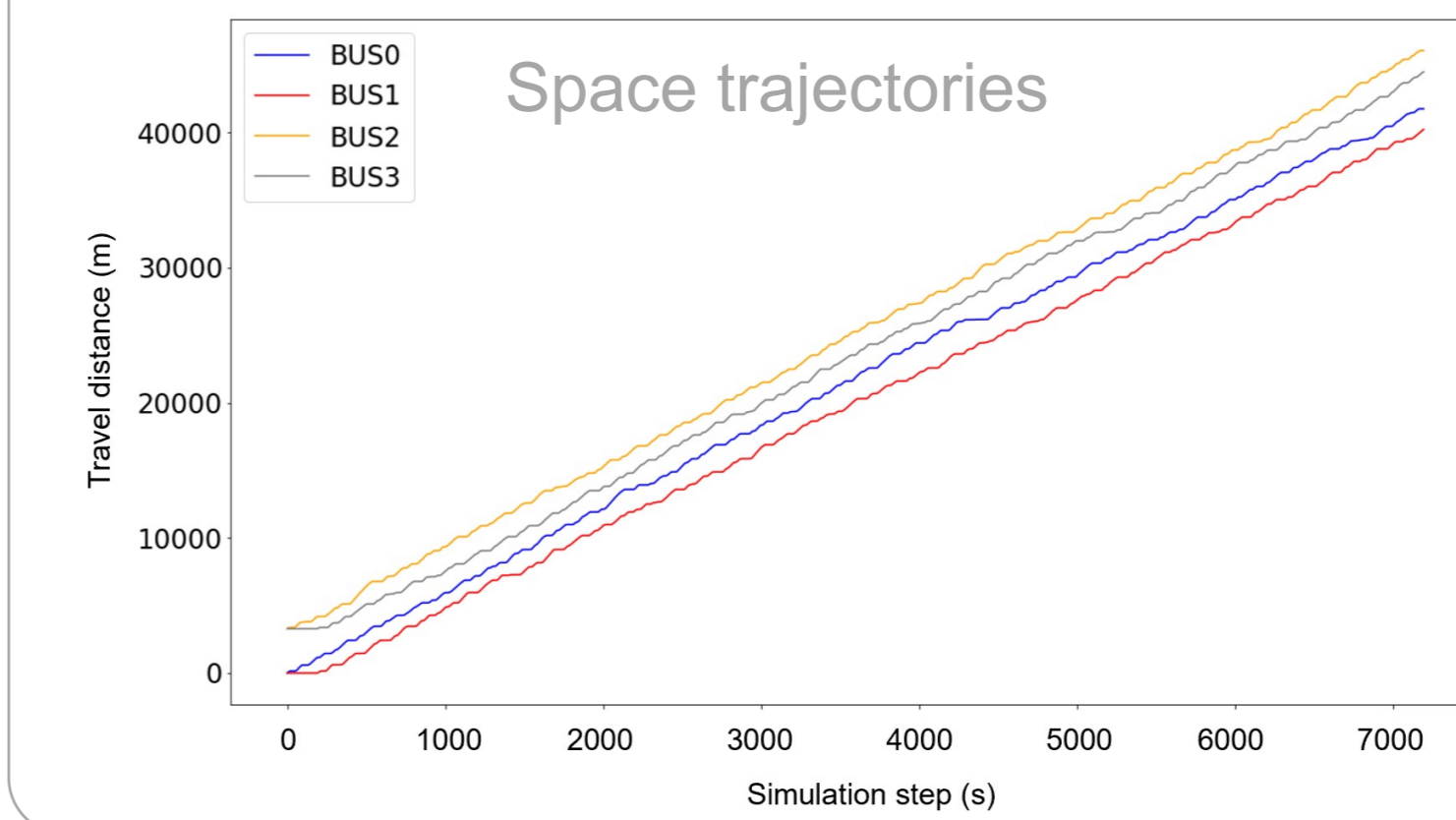
Training network: 5-signal arterial, 4 buses.
Testing network: 10-signal arterial, 8 buses.



Space trajectories of buses (bus lane):



Space and time trajectories of buses (mixed lane, RL - 3):



Value-function approximation:

Since the state-space is continuous, we adopt a process based on Deep Q-Learning to achieve value function approximation.
Optimal action value:

$$Q_i^*(s, a) = \max_{\pi} E \left[r_{i,t} + \gamma r_{i,t+1} + \gamma^2 r_{i,t+2} + \dots \mid S_{i,t} = s, a_{i,t} = a, \pi \right]$$

Update value function:

$$Q_i(s, a) = Q_i(s, a) + \alpha (r_{i,t} + \gamma \max_{a'} Q(s', a') - Q(s, a))$$

Loss function:

$$L_k(\theta_k) = E_{(s,a,r,s') \sim U(M)} \left[\left(r + \gamma \max_{a'} Q(s', a'; \theta_k^-) - Q(s, a; \theta_k) \right)^2 \right]$$

Strategies	Bus lane		Mixed lane	
	SD of space headway	Average queue length (vehs)	SD of space headway	Average queue length (vehs)
RL - 3	508.05	219.02	392.98	124.51
Actuated	748.32	258.18	539.51	144.02
MP	779.26	227.16	569.34	153.94
Fixed	778.64	305.35	738.18	137.26

Conclusions:

- Compared with benchmarks, the proposed method decreases the traffic delay by 28.27% and the standard deviation of bus headways by 34.75% simultaneously.
- Both scalability and portability are demonstrated by transferring locally learned strategies to similar intersection configurations.