APCNCS2026: Asia-Pacific Conference on Networks and Complex Systems
Focus Session: Complexity of AI
10-12 June 2026, Nanyang Technological University, Singapore

Discovery of Bus Control Strategies with Reinforcement Learning to Optimize Commuter-Centric Objective

Andri Pradana[1], Lock Yue Chew[1], and Bo Yang[1]

[1]Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, 637371, Singapore

Keywords: Deep Reinforcement Learning, Multi-Agent Systems, Transportation Systems, Bus Scheduling, Bus Bunching

Extended Abstract

Motivation. Bus transportation systems exhibit complex dynamics arising from stochastic passenger demand, traffic variability, and interactions among buses. One major emergent phenomenon is bus bunching, where initially separated buses synchronize and cluster together, degrading service quality through increased commuter waiting and travel times. While dynamic bus control has been widely studied [3, 1, 4], most existing methods typically impose prior assumptions about the desired system behavior, for example, by optimizing headway regularity to explicitly avoid bus bunching. However, imposing headway regularity a priori assumes that evenly spaced buses necessarily produce the most efficient commuter outcomes. Our work [2] investigates whether modern deep reinforcement learning (RL) algorithms can discover efficient strategies directly from commuter-centric objectives without explicitly enforcing headway-based constraints. The results of this work may potentially challenge long-standing assumptions in transportation control and explore whether AI can uncover emergent coordination strategies that human-designed heuristics may overlook.

Approach and Methodology. We employ the Deep Reinforcement Learning algorithms for dynamic stay/leave decisions at bus stops in bus loop systems. Unlike conventional approaches, the reward functions directly optimize passenger waiting time or travel time, completely removing headway regularization from the optimization objective. The complexity of the multi-agent nature of the system due to multiple buses is reduced by requiring that similar buses in terms of speed/frequency follow the same strategy.

Results. The framework is evaluated using both simplified loop systems and a realistic model of the Nanyang Technological University shuttle bus network. For uniform bus frequency, RL rediscovers near-staggered configurations consistent with headway-based constraints (Table 1 & Fig. 1 left). However, for non-uniform bus frequency, the learned policies permit temporary bus bunching before rapid recovery, resulting in lower overall travel times than strict headway-maintaining control (Table 1 & Fig. 1 right). This confirms the potential of AI-driven discovery where optimal collective behaviors may be unintuitive.

Conclusions and Outlook. The results demonstrate that deep RL can autonomously uncover nontrivial real-time strategies in bus transportation systems by directly optimizing commuter-centric objectives. Due to scaling problem, we only manage to include two buses for non-uniform bus frequency. We plan to address this issue to involve more buses.

References

[1] Konstantinos Gkiotsalitis and Oded Cats. At-stop control measures in public transport: Literature review and research agenda. *Transportation Research Part E: Logistics and Transportation Review*, 145:102176, 2021.

[2] Andri Pradana and Lock Yue Chew. Discovery of bus loop scheduling strategies with reinforcement learning to minimize commuters' waiting and travel times. *Journal of Computational Science*, 95:102807, 2026.

[3] Jiawei Wang and Lijun Sun. Dynamic holding control to avoid bus bunching: A multi-agent deep reinforcement learning framework. *Transportation Research Part C: Emerging Technologies*, 116:102661, 2020.

[4] Yunqiang Xue, Meng Zhong, Luowei Xue, Haokai Tu, Caifeng Tan, Qifang Kong, and Hongzhi Guan. A real-time control strategy for bus operation to alleviate bus bunching. *Sustainability*, 14(13), 2022.

Table 1: **Average waiting time (AWT) and average travel time (ATT) for uniform bus frequency.** Results for various approaches for (left) 3 uniform-frequency buses with 4 bus stops and (right) 2 non-uniform frequency buses with 12 bus stops (NTU Blue route).

	Uniform frequency 3 buses 4 stops		Non-uniform frequency 2 buses 12 stops	
Approach	AWT (s)	ATT (s)	AWT (s)	ATT (s)
Naive (bus bunching)	474 ± 21	1039 ± 25	543 ± 43	1132 ± 49
Holding strategy	171 ± 4	754 ± 10	332 ± 7	1037 ± 16
RL	175 ± 4	729 ± 11	366 ± 11	942 ± 12

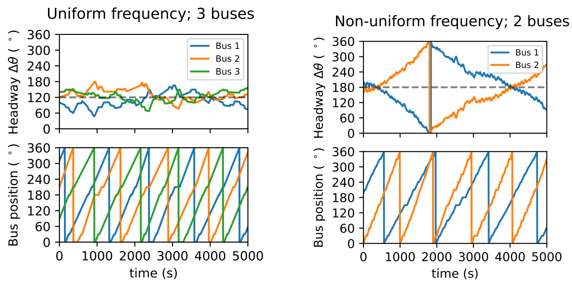

Figure 1: **Phase headway and position plots of buses.** Plots for the (left) 3 uniform-frequency buses with 4 bus stops and (right) 2 non-uniform frequency buses with 12 bus stops (NTU Blue route).