Adaptive Microtolling in Competitive Online Congestion Games via Multiagent Reinforcement Learning

Extended Abstract

Behrad Koohy
University of Southampton
Southampton, United Kingdom
behrad.koohy@soton.ac.uk

Sebastian Stein
University of Southampton
Southampton, United Kingdom
ss2@ecs.soton.ac.uk

Enrico Gerding
University of Southampton
Southampton, United Kingdom
eg@ecs.soton.ac.uk

ABSTRACT

Efficient urban traffic management remains a critical challenge, yet traditional congestion games fail to capture the dynamic and competitive nature of real-world transportation systems. We introduce the Multi-Market Routing Problem (MMRP), an online and oligopolistic extension that models competition amongst route providers utilising adaptive microtolling strategies to influence driver behaviour and mitigate congestion. We formally define the MMRP, highlighting the computational complexity of solving the MMRP, and use an adapted version of Proximal Policy Optimisation (PPO) to improve update stability in multiagent environments to address this problem in online settings. Our empirical analysis demonstrates that our PPO-based approach not only matches the performance of existing benchmarks but also significantly enhances equity, reduces travel times for users, and increases profitability for providers.

KEYWORDS

Multiagent Reinforcement Learning; Competitive Games; Congestion Games; Adaptive Pricing; Mechanism Design

ACM Reference Format:

Behrad Koohy, Sebastian Stein, and Enrico Gerding. 2025. Adaptive Microtolling in Competitive Online Congestion Games via Multiagent Reinforcement Learning: Extended Abstract. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

1 INTRODUCTION

Urban transportation systems are increasingly burdened by congestion, a challenge that significantly impacts economic productivity and quality of life [6]. Traditional Congestion Game (CG) [13] models provide valuable insights into individual route choices and their impacts on traffic flow, but neglect the competitive dynamics present in modern, dynamic, transportation networks, where multiple transportation providers compete in an oligopolistic manner. Effective modelling of competition in modern urban transportation networks provides valuable insights into how to maximise the efficiency of existing infrastructure and guide the strategic development of new transportation systems [19].

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

Traditional CGs are non-cooperative models where individual players select resources (i.e. routes in the context of transportation CGs) and an associated cost is incurred, which escalates with the popularity of that resource. In traffic management, CGs are essential for simulating user behaviours and decision-making in congested environments [10, 22], offering insights into traffic patterns and identifying bottlenecks for optimisation [1, 2, 7, 9, 15, 17]. A detailed discussion of existing applications of CGs to congestion management can be found in de Palma and Fosgerau [3], de Palma and Lindsey [4], Yang et al. [20]. While these models offer valuable insights into the impact of individual decisions on overall system performance, they are typically static and offline problems which assume perfect information. These simplifications fail to capture the dynamic and competitive interactions inherent in modern urban transportation networks, where multiple route providers continuously compete in real time. This limitation underscores the urgent need for more sophisticated modelling that more closely reflects the dynamic and competitive nature of urban transportation systems.

To overcome these limitations, we propose a novel framework that extends traditional CGs into an online and competitive setting, referred to as the **Multi-Market Routing Problem (MMRP)**. In our framework, transportation networks are modelled as systems where multiple route providers are able to utilise adaptive pricing to influence the behaviour of transportation users in response to fluctuating traffic conditions and competitive pressures. To solve the MMRP in practice, we propose a multiagent reinforcement learning based approach, utilising Proximal Policy Optimisation, to learn adaptive pricing strategies that effectively manage congestion in real time. Our framework bridges the gap between theoretical models and practical traffic management, and our empirical results demonstrate significant improvements in travel times, equity, and profitability, underscoring its potential impact on intelligent transportation systems.

2 ONLINE MULTI-MARKET ROUTING PROBLEM

We expand the definition of a Congestion Game [14, 19] to the Multi-Market Routing Problem (MMRP) M, where M is a 6-tuple: $M = (R, V, (\Phi_j)_{j\in V}, (\Theta_j)_{j\in V}, (D_i)_{i\in R}, (\Omega_i)_{i\in R})$ The set $R = \{R_0, \ldots, R_i\}$ is the set of available routes; the set $V = \{V_0, \ldots, V_j\}$ is the set of players. For each player $V_j \in V$, Φ_j denotes the strategy space of player V_j and Θ_j denotes the Value of Time (VoT) of player V_j. For each route $R_i \in R$, $D_i : \{0, \ldots, j\} \to \mathbb{R}$ represents the delay function of the route, mapping the number of players selecting a route to a travel time, and Ω_i represents the route cost strategy. For each player $V_j \in V, \Phi_j \subseteq 2^R$ defines the strategy space. We define MMRP as an optimisation problem, where the optimal assignment of an instance of MMRP is one in which players incur the lowest total travel time. When this problem is extended to online scenarios, we use the 7-tuple $\mathcal{OM} = (R, V, (\Phi_j)_{j\in V}, (\Theta_j)_{j\in V}, (\Lambda_j)_{j\in V}, (T_i)_{i\in R}, (\Omega_i)_{i\in R})$ where the variables (R, V, Φ) from the offline 6-tuple definition are not changed. In the online definition, Λ_i represents the entry time of a player V_i, and functions D, Ω are changed to become strategies that depend on time t, becoming $D(x, t)$ and $\Omega(x, t)$.

The offline formulation of the MMRP is NP-hard[1], rendering exact optimisation methods computationally intractable for large-scale instances and online problems. To this end, we employ Multiagent Reinforcement Learning (MARL), specifically an adapted version of independent Proximal Policy Optimisation (PPO) [16]. The use of PPO over existing MARL approaches is deliberate; the inherently competitive nature of route providers suggests that approaches which require inter-agent coordination, such as centralised learning with decentralised execution, are impractical. Consequently, we employed independent PPO, adapted for increased training stability and suitability in our environment, which allows each provider to optimise its pricing strategy. To adapt PPO for use in the MMRP, we employed separate policy and value networks [21], proven effective in multiagent and highly stochastic settings [8], to enhance stability. To mitigate divergence, we normalised rewards by first running random-agent experiments to compute the mean (μ_R) and variance (σ_R) of rewards, and then applied $\tilde{R}_t = \frac{R_t - \mu_R}{\sigma_R}$. Finally, we enabled multiple parallel experiments to replicate the vectorised actor framework used in single-agent PPO [16].

This approach enables each route provider to dynamically adjust tolls in real time, approximating the complex equilibrium behaviour of the system and effectively managing congestion. Our method thus offers a scalable, adaptive solution that bridges the gap between theoretical complexity and practical real-world traffic management. We consider a two-route network ($R = \{R_1, R_2\}$), akin to the parallel two-link models in [11, 18]. Each route's delay is defined by the Bureau of Public Roads volume delay function: $D_i(x) = f_0(1 + \alpha(\frac{x}{f_c})^\beta)$, where x is the number of vehicles at time t, f_0 is the free-flow travel time, f_c is the route capacity, and the calibration parameters $\alpha = 0.68$ and $\beta = 2.73$ align the function with real-world data [12]. For the values of (f_c, f_0), we set R_1 as $(15, 20)$ and R_2 as $(30, 20)$ for routes 1 and 2 respectively.

We trained our PPO agent for 4×10^7 steps, with each episode lasting 1000 timesteps. In each episode, the number of players was sampled from a uniform distribution $\mathbb{U}(500, 1000)$ to generalise across varied traffic scenarios. Agents share the same architecture, enabling robust performance without environment-specific tuning. The reward function is defined as profit per timestep to reflect a route provider's objective, and the action space consisted of three discrete actions: increase, maintain, or decrease the price by 1.

For our evaluation, we measured the average travel time and profit per vehicle, and employed the Gini coefficient [5] to quantify inequality in travel times across our simulations. A lower Gini coefficient indicates a more equitable distribution of travel times, while a higher value reveals significant disparities . This multi-faceted evaluation framework not only demonstrates the efficiency of our adaptive pricing strategies, but also rigorously assesses their fairness, providing a comprehensive picture of system performance under our proposed solution.

Table 1: Adaptive Pricing Results for the Online MMRP

	V	500	600	750	900	1000
MMRP-PPO	Time	**26.66**	**30.85**	**117.1**	**469.12**	**1739.88**
	Gini Coef.	**0.14**	**0.14**	**0.33**	**0.18**	**0.13**
	Profit	13.38	24.65	**66.04**	**86.78**	**90.41**
Random	Time	32.50	34.87	151.83	553.88	2083.49
	Gini Coef.	0.18	0.15	0.44	0.20	0.14
	Profit	**46.13**	**50.8**	46.64	48.52	46.55

3 RESULTS

Our results (Table 1) demonstrate that our adapted PPO-based approach significantly outperforms a random pricing agent in a two-route synthetic environment. At 500 players, our method achieves an average travel time of 26.66 timesteps compared to 32.50 timesteps for the Random Agent, while consistently maintaining lower Gini coefficients and yielding higher profits (rising from 13.38 at 500 players to 90.41 at 1000 players). Moreover, under infinite capacity conditions, our agents converge towards equilibrium strategies, confirming that our approach effectively captures equilibrium-like behaviour[2].

These results underscore the potential of our adaptive pricing strategy to transform real-world traffic management, delivering not only reduced congestion, but also a fairer distribution of travel costs. The robust convergence towards equilibrium under infinite capacity further validates our approach, suggesting its applicability in more complex, real-world scenarios.

4 CONCLUSION

In this study, we introduced the **Multi-Market Routing Problem (MMRP)**, an online, oligopolistic extension of traditional congestion games that models real-world traffic competition through multiple route providers employing adaptive microtolling. We formally defined MMRP and, to overcome its computational complexity, we developed an enhanced **Proximal Policy Optimisation (PPO)** algorithm tailored for competitive multiagent settings. Our evaluations demonstrate that our approach significantly reduces travel times, promotes equity, and increases provider profitability compared to benchmarks. Future work will explore scalability, reduced training costs, advanced techniques such as opponent modelling, and improved explainability to further bridge theory and practical traffic management. Overall, our contributions advance congestion game theory and offer actionable strategies for developing intelligent, adaptive transportation systems.

ACKNOWLEDGMENTS

We would like to acknowledge Dr Vahid Yazdanpanah for his feedback on this work. This research is supported by an ICASE studentship from EPSRC and Yunex Traffic, the EPSRC AutoTrust platform grant (EP/R029563/1), and an EPSRC Turing AI Acceleration Fellowship on Citizen-Centric AI Systems (EP/V022067/1).

[1] Proof omitted due to space constraints.

[2] Results omitted due to space constraints.

REFERENCES

[1] Philip N Brown and Jason R Marden. 2017. Optimal mechanisms for robust coordination in congestion games. *IEEE Trans. Automat. Control* 63, 8 (2017), 2437–2448.

[2] José Correa, Ruben Hoeksma, and Marc Schröder. 2019. Network congestion games are robust to variable demand. *Transportation Research Part B: Methodological* 119 (2019), 69–78.

[3] André de Palma and Mogens Fosgerau. 2010. Dynamic and Static congestion models: A review. *HAL Working Papers* hal-00539166 (2010).

[4] André de Palma and Robin Lindsey. 2011. Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies* 19, 6 (2011), 1377–1399.

[5] Robert Dorfman. 1979. A formula for the Gini coefficient. *The review of economics and statistics* (1979), 146–149.

[6] Kurt Dresner and Peter Stone. 2004. Multiagent traffic management: A reservation-based intersection control mechanism. In *Autonomous Agents and Multiagent Systems, International Joint Conference on*, Vol. 3. Citeseer, 530–537.

[7] Svenja M Griesbach, Martin Hoefer, Max Klimm, and Tim Koglin. 2022. Public Signals in Network Congestion Games. In *Proceedings of the 23rd ACM Conference on Economics and Computation*. 736–736.

[8] Shengyi Huang, Rousslan Fernand Julien Dossa, Antonin Raffin, Anssi Kanervisto, and Weixun Wang. 2022. The 37 Implementation Details of Proximal Policy Optimization. In *ICLR Blog Track*. https://iclr-blog-track.github.io/2022/03/25/ppo-implementation-details/ https://iclr-blog-track.github.io/2022/03/25/ppo-implementation-details/.

[9] Zahid Khan, Anis Koubaa, Bilel Benjdira, and Wadii Boulila. 2023. A game theory approach for smart traffic management. *Computers and Electrical Engineering* 110 (2023), 108825.

[10] Olivier Massicot and Cedric Langbort. 2019. Public signals and persuasion for road network congestion games under vagaries. *IFAC-PapersOnLine* 51, 34 (2019), 124–130.

[11] Igal Milchtaich. 1996. Congestion games with player-specific payoff functions. *Games and economic behavior* 13, 1 (1996), 111–124.

[12] Robert Neuhold and Martin Fellendorf. 2014. Volume delay functions based on stochastic capacity. *Transportation research record* 2421, 1 (2014), 93–102.

[13] Robert W. Rosenthal. 1973. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory* 2, 1 (01 Dec 1973), 65–67. https://doi.org/10.1007/BF01737559

[14] Tim Roughgarden and Éva Tardos. 2004. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and economic behavior* 47, 2 (2004), 389–403.

[15] Marco Scarsini, Marc Schröder, and Tristan Tomala. 2018. Dynamic atomic congestion games with seasonal flows. *Operations Research* 66, 2 (2018), 327–339.

[16] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).

[17] Chaitanya Swamy. 2012. The effectiveness of Stackelberg strategies and tolls for network congestion games. *ACM Transactions on Algorithms (TALG)* 8, 4 (2012), 1–19.

[18] Hamidreza Tavafoghi and Demosthenis Teneketzis. 2017. Informational incentives for congestion games. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 1285–1292.

[19] Cristian Toșa, Andrei Mitrea, Hitomi Sato, Tomio Miwa, and Takayuki Morikawa. 2018. Economic growth and urban metamorphosis. *Journal of Transport and Land Use* 11, 1 (2018), 273–295.

[20] Li Yang, Xinyuan Chen, and Zhiyuan Liu. 2019. Impact Analysis of Congestion Pricing Scheme in a Multimodal Transport Network. In *CICTP 2019*. ASCE, 3237–3248.

[21] Chao Yu, Akash Velu, Eugene Vinitsky, Jiaxuan Gao, Yu Wang, Alexandre Bayen, and Yi Wu. 2022. The surprising effectiveness of ppo in cooperative multi-agent games. *Advances in Neural Information Processing Systems* 35 (2022), 24611–24624.

[22] Jing Zhang, Jianquan Lu, Jinde Cao, Wei Huang, Jianhua Guo, and Yun Wei. 2021. Traffic congestion pricing via network congestion game approach. *Discrete & Continuous Dynamical Systems: Series A* 41, 7 (2021).