Balancing Fairness and Efficiency in Transport Network Design through Reinforcement Learning

Extended Abstract

Dimitris Michailidis
University of Amsterdam
Amsterdam, The Netherlands
d.michailidis@uva.nl

Sennay Ghebreab
University of Amsterdam
Amsterdam, The Netherlands
s.ghebreab@uva.nl

Fernando P. Santos
University of Amsterdam
Amsterdam, The Netherlands
f.p.santos@uva.nl

ABSTRACT

Designing well-functioning and fair transport networks is not a trivial task, given the large space of solutions and constraints one must satisfy. Moreover, different spatial segregation sources can render some transportation network interventions unfair to specific groups. It is thereby crucial to optimize the transportation system while mitigating the disproportional benefits it can lead to. In this paper, we explore the trade-off between efficiency and fairness in the Transport Network Design Problem (TNDP), via the use of Deep Reinforcement Learning (Deep RL). We formulate different fairness definitions as reward functions — inspired by Equal Sharing of Benefits, Narrowing the Gap, and Rawl's justice theory. We apply our method to Amsterdam (The Netherlands) and Xi'An (China) and show that *vanilla* Deep RL can lead to biased outcomes. By considering different fair rewards, however, we can shed light on possible compromises between fairness and efficiency in the TNDP.

KEYWORDS

Transport Network Design; Deep Reinforcement Learning; AI for Social Good; Fairness; Equity

ACM Reference Format:
Dimitris Michailidis, Sennay Ghebreab, and Fernando P. Santos. 2023. Balancing Fairness and Efficiency in Transport Network Design through Reinforcement Learning: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Public transportation planning driven by economic principles of efficiency has exacerbated urban inequalities [12]. It is, therefore, crucial to look beyond optimizing the efficiency of public transport networks and to consider potential disparities they can lead to.

The Transport Network Design Problem (TNDP) is an NP-hard optimization problem, where the goal is to design a new line that maximizes the total citizens' travel demand satisfied. TNDP has traditionally been addressed through integer optimization and heuristic algorithms [1, 7, 11, 13, 15]. These methods, however, require a long list of constraints and are hard to generalize, entailing limitations to the search space in order to make the problem tractable [18]. Given the sequential nature of designing new transportation systems, formulating the TNDP as a Deep Reinforcement Learning (Deep RL) problem can inspire new solutions that enable non-myopic long-term decisions, as noted in recent work [5, 17].

Deep RL can effectively explore the search space in a variety of TNDP environments by optimizing a reward function, without requiring a long list of constraints to be formalized and imposed [18]. This suggested new prominent methods prone to be applied in real-world settings [14, 19]. Previous models ignore inequality issues, and as we show in this paper, can even worsen them.

To address disparities in the TNDP, we look into transport planning research, measuring inequalities along different dimensions [4, 6, 16]. We formulate the RL optimization problem of fairness and efficiency in transport network design and formalize fairness-based reward functions, evaluating their result along an efficiency-fairness trade-off. We apply this methodology in two real city environments: Xi'an in China [18]), and Amsterdam in The Netherlands. We show that different reward functions can be used to navigate the trade-offs between efficiency and fairness and to provide a new toolkit for decision-makers to test how different fairness-efficiency requirements can map onto different transportation lines.

2 FAIRNESS VS EFFICIENCY IN THE TNDP

We model a city as a two-dimensional grid environment $H^{n \times m}$. A transport line is a spatial graph $G(N, E)$, connecting a set of cells in the grid. The goal is to create lines that optimize the *total captured travel demand*, expressed as a function U_{od} of the estimated Origin-Destination (OD) matrix [8, 10]. To model TNDP as a Markov Decision Process (MDP), we adopt the formulation in [18]. An agent generates a transport line by taking sequential actions, i.e. select cells to be stops in the new line.At every time step t, the agent selects a cell $h \in H^{n \times m}$ to place a station on. At the end of the episode, the sequence of selected cells is the generated transport graph $G(N, E)$. The total selected locations are constrained by a budget B, a station number limit T, and direction-based constraints, so as to avoid unfeasible line shapes; for simplicity, we omit these constraints from the following discussion (see [18] for details). Formally, the MDP $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R} \rangle$ stands as:

- $\mathcal{S}$: the state; sequence of previously selected grid cells to include a transportation line stop.
- $\mathcal{A}$: the action; selected cell at each time-step to add a stop.
- $\mathcal{P} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \{0, 1\}$: state transition function.
- $\mathcal{R} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}$: the reward.

To consider fairness, we define a set A, representing different groups, based on the house-price index (proxy for an area's development). Each cell h is associated with a group $a \in A$. We adjust

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

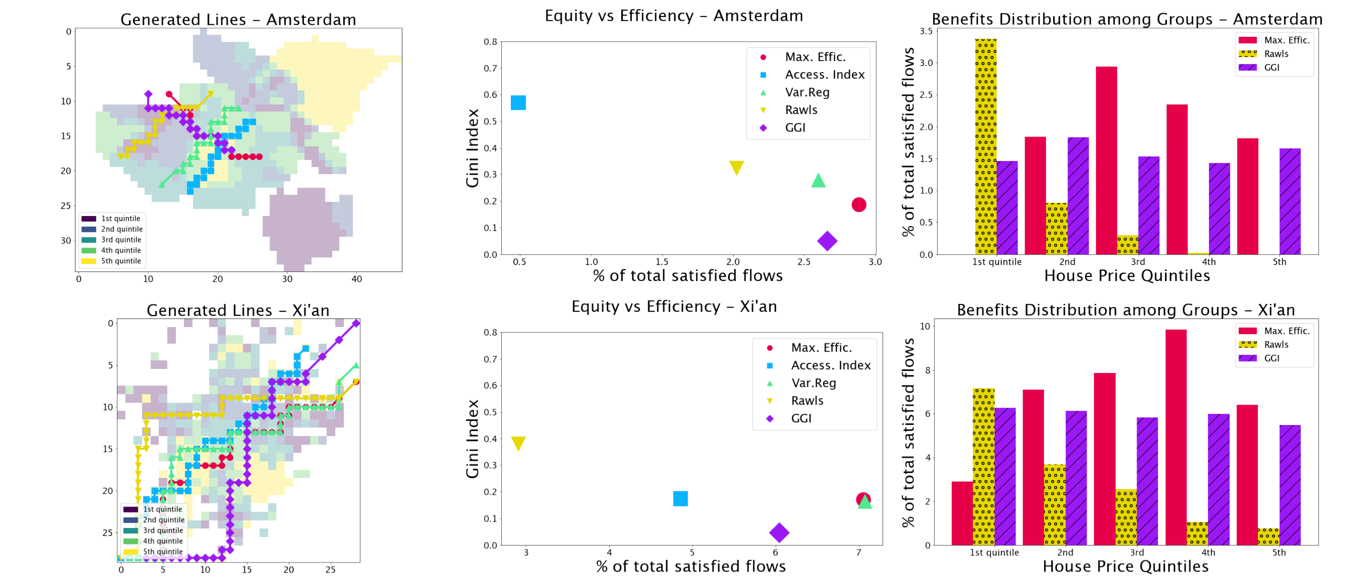

Figure 1: Results in Amsterdam and Xi'an. In the left column, we show the average generated lines for each city. In the middle, we present the fairness-efficiency trade-off(s). In the right column, we show the distribution of the satisfied demand between the five groups for selected models. Here we use house prices as a proxy for placing individuals in different wealth bins.

the objective function based on different notions of fairness [2]. We use $U(G)$ and $U_a(G)$ to denote the fraction of demand covered by the new transportation graph G, over all groups and for a specific group a, respectively.

Variance Regularization . This reward aims to achieve the *Narrowing the Gap* notion, by regularizing the variance of the groups' satisfied travel demand with a hyperparameter λ:

$$r = \sum_a U_a(G) - \lambda \mathrm{var}(U(G)). \tag{1}$$

Rawls . Reward based on Rawl's theory of justice that aims at maximizing the benefits received by the least privileged group [2]:

$$r = U_{a_0}(G), \tag{2}$$

where a_0 is the group with least utility before adding the new line.

Generalized Gini Index (GGI). This reward is used to achieve *Equal Sharing* of benefits. It is a weighted sum of the groups' utility, where the lowest weight is assigned to the group with the highest utility. The weights are assigned via a hyperparameter.

$$GGI_w = \sum_{\forall a \in A} w_d \ \sigma(U_a(G)). \tag{3}$$

σ is a permutation that sorts the utilities in descending order and weights w_d are non-increasing weights, $w_1 > w_2 > ... > w_{|A|}$.

We use a recently proposed Deep RL method to explore the trade-off described above and modify the reward function of the agent to achieve outcomes with different degrees of fairness [3, 18]. At each time step of the episode, the agent receives the current state as input and selects the next action based on a parametrized policy network $\pi(a_t|s_t, \theta)$. Action selection is stochastic during training and greedy during testing [3, 18]. The agent is trained via a policy gradient advantage actor-critic framework (A2C) [9]. Details on the architecture can be found in our supplementary code [1] and documentation, and in previous papers [3, 18].

3 EXPERIMENTS

We ran a vanilla Deep RL method (baseline) and our proposed modifications in two real-world case study cities: Xi'an and Amsterdam, evaluating them on total satisfied Origin-Destination (OD) % and the Gini index of satisfied OD % between groups — ranging from 0 (perfect equality) to 0.8 (maximum inequality — 0.8 due to there being five values, one for each group).

Max Efficiency Baseline. It maximizes the total efficiency of the line, without considering the benefits of different groups: $r = U(G)$.

Accessibility Index Baseline. The distance-decayed Wei et al. reward that aims to connect low and high-priced areas: $r = \sum_j D_j e^{-\beta t_{ij}}$, $i \neq j$, where D_j is the house-price index of j, t_{ij} is the distance between i, j and β is a tunable parameter.

In Figure 1, we show the generated transport lines, along with their performance on efficiency and fairness. We observe that utilitarianism can lead to disparities: the method assuming baseline reward (red circle) tends to achieve the best results on overall efficiency, but that comes at the expense of fairness. Disparities are mitigated by the proposed reward functions, which outperform the baselines in their respective fairness goals, providing a holistic view of the trade-off in the TNDP. In conclusion, our experiments show that due to segregation, different cities are characterized by different efficiency-fairness trade-offs, suggesting that policy-makers should consider multiple fairness criteria when applying these algorithms.

[1] https://github.com/sias-uva/fair-transport-network-design

ACKNOWLEDGMENTS

This research was carried on the SIAS group as part of the Civic AI Lab project (ICAI) — in collaboration with the Municipality of Amsterdam.

REFERENCES

[1] Renato Oliveira Arbex and Claudio Barbieri da Cunha. 2015. Efficient transit network design and frequencies setting multi-objective optimization by alternating objective genetic algorithm. *Transportation Research Part B: Methodological* 81 (Nov. 2015), 355–376. https://doi.org/10.1016/j.trb.2015.06.014

[2] Hamid Behbahani, Sobhan Nazari, Masood Jafari Kang, and Todd Litman. 2019. A conceptual framework to formulate transportation network design problem considering social equity criteria. *Transportation Research Part A: Policy and Practice* 125 (July 2019), 171–183. https://doi.org/10.1016/j.tra.2018.04.005

[3] Irwan Bello, Hieu Pham, Quoc V. Le, Mohammad Norouzi, and Samy Bengio. 2017. Neural Combinatorial Optimization with Reinforcement Learning. *arXiv:1611.09940 [cs, stat]* (Jan. 2017). http://arxiv.org/abs/1611.09940 arXiv: 1611.09940.

[4] Wenting Cheng, Jiahui Wu, William Moen, and Lingzi Hong. 2021. Assessing the spatial accessibility and spatial equity of public libraries' physical locations. *Library & Information Science Research* 43, 2 (April 2021), 101089. https://doi.org/10.1016/j.lisr.2021.101089

[5] Ahmed Darwish, Momen Khalil, and Karim Badawi. 2020. optimising Public Bus Transit Networks Using Deep Reinforcement Learning. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, Rhodes, Greece, 1–7. https://doi.org/10.1109/ITSC45102.2020.9294710

[6] Ahmed El-Geneidy, David Levinson, Ehab Diab, Genevieve Boisjoly, David Verbich, and Charis Loong. 2016. The cost of equity: Assessing transit accessibility and social disparity using total travel cost. *Transportation Research Part A: Policy and Practice* 91 (Sept. 2016), 302–316. https://doi.org/10.1016/j.tra.2016.07.003

[7] Wei Fan and Randy B. Machemehl. 2006. Using a Simulated Annealing Algorithm to Solve the Transit Route Network Design Problem. *Journal of Transportation Engineering* 132, 2 (Feb. 2006), 122–132. https://doi.org/10.1061/(ASCE)0733-947X(2006)132:2(122) Publisher: American Society of Civil Engineers.

[8] Reza Zanjirani Farahani, Elnaz Miandoabchi, W. Y. Szeto, and Hannaneh Rashidi. 2013. A review of urban transportation network design problems. *European Journal of Operational Research* 229, 2 (Sept. 2013), 281–302. https://doi.org/10.1016/j.ejor.2013.01.001

[9] Ivo Grondman, Lucian Busoniu, Gabriel A. D. Lopes, and Robert Babuska. 2012. A Survey of Actor-Critic Reinforcement Learning: Standard and Natural Policy Gradients. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42, 6 (Nov. 2012), 1291–1307. https://doi.org/10.1109/TSMCC.2012.2218595 Conference Name: IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews).

[10] Valérie Guihaire and Jin-Kao Hao. 2008. Transit network design and scheduling: A global review. *Transportation Research Part A: Policy and Practice* 42, 10 (2008), 1251–1273.

[11] Gilbert Laporte and Marta M. B. Pascoal. 2015. Path based algorithms for metro network design. *Computers & Operations Research* 62 (Oct. 2015), 78–94. https://doi.org/10.1016/j.cor.2015.04.007

[12] Karel Martens. 2016. *Transport Justice: Designing fair transportation systems*. Routledge. Google-Books-ID: m0yTDAAAQBAJ.

[13] Antonio Mauttone and María E. Urquhart. 2009. A multi-objective metaheuristic approach for the Transit Network Design Problem. *Public Transport* 1, 4 (Nov. 2009), 253–273. https://doi.org/10.1007/s12469-010-0016-7

[14] Grigory Neustroev, Sytze P. E. Andringa, Remco A. Verzijlbergh, and Mathijs M. De Weerdt. 2022. Deep Reinforcement Learning for Active Wake Control. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS '22)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 944–953.

[15] Mahmoud Owais and Mostafa K. Osman. 2018. Complete hierarchical multi-objective genetic algorithm for transit network design problem. *Expert Systems with Applications* 114 (Dec. 2018), 143–154. https://doi.org/10.1016/j.eswa.2018.07.033

[16] Rafael H. M. Pereira, David Banister, Tim Schwanen, and Nate Wessel. 2019. Distributional effects of transport policies on inequalities in access to opportunities in Rio de Janeiro. *Journal of Transport and Land Use* 12, 1 (2019), 741–764. https://www.jstor.org/stable/26911287 Publisher: Journal of Transport and Land Use.

[17] Naveen Raman, Sanket Shah, and John Dickerson. 2021. Data-Driven Methods for Balancing Fairness and Efficiency in Ride-Pooling. *arXiv:2110.03524 [cs]* (Oct. 2021). http://arxiv.org/abs/2110.03524 arXiv: 2110.03524.

[18] Yu Wei, Minjia Mao, Xi Zhao, Jianhua Zou, and Ping An. 2020. City Metro Network Expansion with Reinforcement Learning. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, Virtual Event CA USA, 2646–2656. https://doi.org/10.1145/3394486.3403315

[19] Ziyi Xu, Xue Cheng, and Yangbo He. 2022. Performance of Deep Reinforcement Learning for High Frequency Market Making on Actual Tick Data. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS '22)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1765–1767.