Analysis and mitigation of discriminatory behaviour in fleet management algorithms

Tarek Chouaki, IRT SystemX
Sebastian Hörl, IRT SystemX

Extended abstract submitted for presentation at the 12th Triennial Symposium on Transportation Analysis conference (TRISTAN XII)
June 22-27, 2025, Okinawa, Japan

October 31, 2024

Keywords: fairness, fleet management, on-demand mobility, simulation

1 INTRODUCTION

On-demand mobility systems consist of a fleet of vehicles that operate to serve travel requests arising during the day. One field of research related to these systems focuses on the design of operational strategies and evaluating them in realistic settings using agent-based mobility simulation tools (Jing *et al.*, 2020).

The core of the operation of on-demand systems lies in the dispatch algorithm, deciding which vehicle to allocate to which travel request. Most algorithms in the literature strictly aim at maximizing the number of transported passengers while minimizing the costs. Moreover, the common simulation setup assumes homogeneous users. Their interactions with the service are all simulated in the same way, which, in turn, avoids considering the potential discriminatory behavior that algorithms could show when deployed in a real-life system.

In this work, we study the discriminatory behavior of an insertion-based dispatch algorithm that is broadly used in the literature. We then propose and evaluate mitigation measures. These experiments are built using the MATSim agent-based mobility simulator Horni *et al.* (2016).

2 METHODOLOGY

2.1 Investigated dispatch algorithm

This work focuses on the DRT algorithm (Bischoff *et al.*, 2019). It is an insertion-based algorithm in which requests are processed in every simulated second of a day in their order of arrival. For new each request, the algorithm tries to insert new pickup and dropoff stops into the schedules of the fleet vehicles. Those contain the pickup and dropoff locations of already assigned requests. An insertion is a combination of a pickup and a dropoff index on the sequence of existing actions of the vehicle. For each insertion, it is checked whether neither the pickup time nor the dropoff time of any already assigned request would be shifted beyond promised thresholds, determined respectively by the maximum wait time (fixed to 10 minutes in this study) and the latest arrival time. If no insertion that fulfills these conditions exists, the new request is rejected. Otherwise, one insertion is chosen that causes the least additional drive time for the vehicle fleet. Furthermore, we use MATSim's recent prebooking functionality (Hörl *et al.*, 2024) to have all requests submitted *five* minutes in advance of their planned departure time.

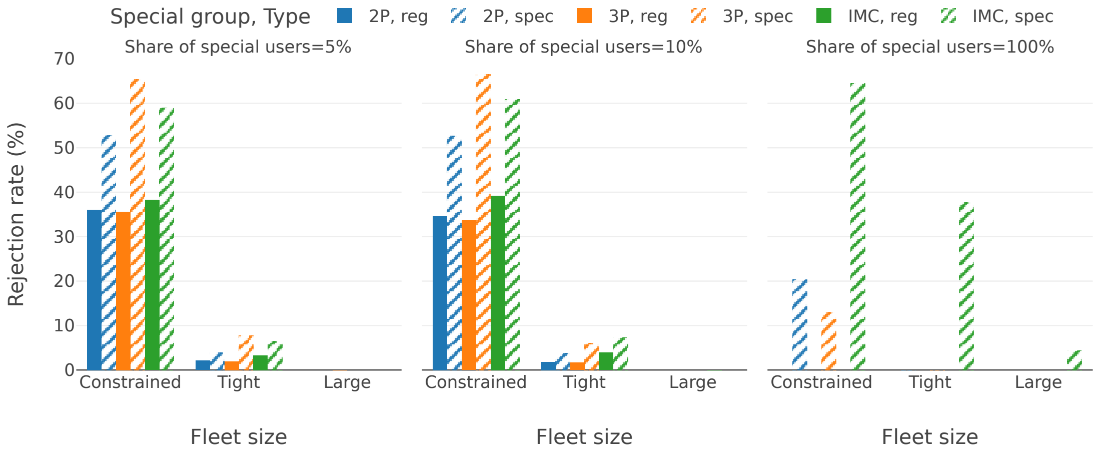

Figure 1 – *Assessment of rejection rates for regular and special users in the baseline case.*

2.2 Special users and demand scenarios

The time required by the users for pickup and dropoff is considered by DRT when scheduling the vehicle stops. Here, we consider that pickup and dropoff durations are the same for each request, and we call them *interaction time*. In most simulations in the literature with homogeneous users, all requests have the same *interaction time*. This assumption is relaxed here by considering, besides *regular* users with an *interaction time* of $30s$, the presence of *special* user types that have different *interaction time* configurations. We study *three* types of *special* users. **Individuals with mobility challenges (IMC)** refers to users that have special constraints when interacting with a vehicle, causing a higher *interaction time* of $240s$. They represent, for instance, users with wheelchairs. **Two-person groups (2P)** and **Three-person groups (3P)** have an interaction time of $60s$ and $90s$, respectively, and require the insertion of two and three persons, respectively, in the same vehicle at the same time. They represent, for instance, parents with children.

Our baseline demand scenario first reflects the state of the literature where all users are *regular*. Then, each of the three *special* types is separately added to represent a certain share of users (5%, 10%, 100%). The 100% share allows studying the behavior of the algorithm in the extreme case where all users have the same, but higher than the baseline, *interaction time*.

2.3 Mitigation strategies

We explore *two* methods to prevent the algorithm from discriminating against *special* users. The **User Agnostic Scheduling (UAS)** strategy makes the algorithm unaware of the actual *interaction time* of a specific request, which is not true in the baseline case. A parametrizable *user agnostic interaction time* is assumed instead (30s and 240s in the experiments below). The **Targeted Planning (TP)** strategy explicitly favors *special* requests by allowing them to send the trip requests earlier than other users, such that their requests can be scheduled with preference. The strategy is parametrized by the *targeted planning horizon* (10 to 30 minutes).

2.4 Fleet simulation

Our simulation experiments are conducted using synthetic travel demand data for the city of Paris (Hörl & Balac, 2021), by targeting 10% of the bus trips that take place in the city. This results in $50k$ daily requests. Fleet sizes ranging up to 600 vehicles were tested for serving the regular demand, yielding three configurations of interest. The *Constrained* fleet (150 vehicles) rejects 37% of requests throughout the day, the *Tight* fleet (300 vehicles) rejects 5%, mostly during peak times, and the *Large* fleet (550 vehicles) serves the whole baseline demand.

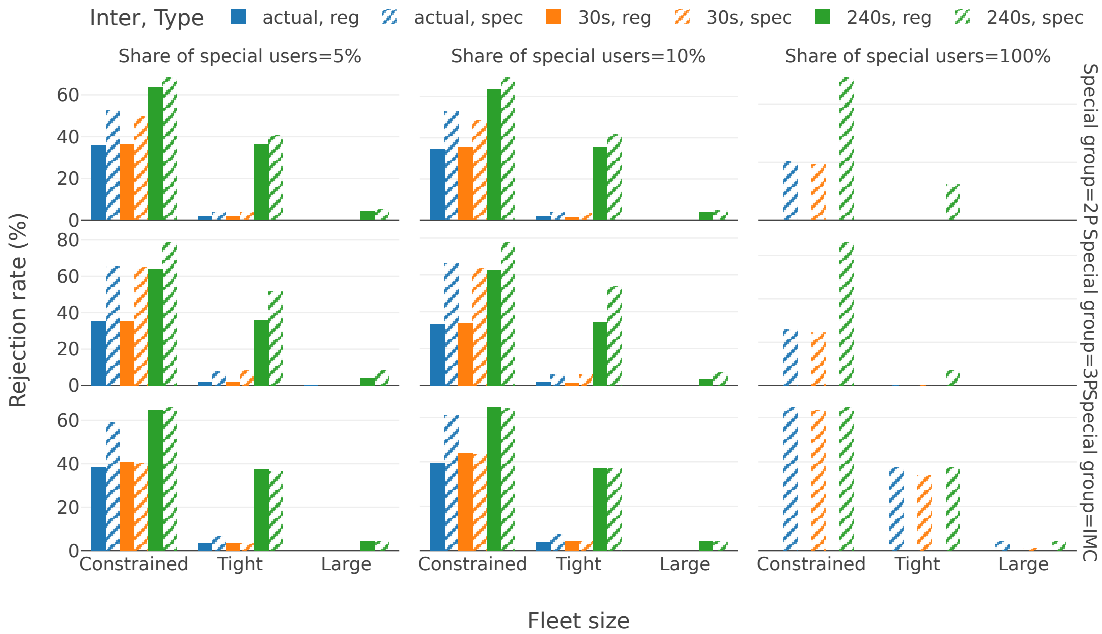

Figure 2 – *Rejection rates observed under the application of the UAS mitigation strategy.*

3 RESULTS

To test for discriminatory behavior of the algorithm, simulations for all fleet configurations combined with introducing each *special* user types separately, at a certain rate, are conducted. Figure 1 shows the observed rejection rates. The demand is characterized by the plot columns (share of *special* requests) and the color (type of *special*) requests. Each group of bars represents a fleet configuration, and the bar shades allow comparing between *regular* and *special* requests. From the height of the rejection rates, one can see that the algorithm yields substantially higher rejection rates for *special* users. The algorithm, hence, shows discriminatory behavior, and even at a *Large* fleet for **IMR** requests.

In the following, the potential of the two mitigation strategies is assessed. Figure 2 shows the rejection rates observed when applying **UAS**. The *special* request scenarios are now arranged in rows, while colors represent different *user agnostic interaction times*. The *actual* case shows the baseline, with each request being treated with its specific interaction time. In the case of a user agnostic interaction time of 30s, the algorithm is prevented from considering explicitly the higher interaction time of *special* requests. This leads to a slight decrease of their rejection rate, with the largest effect in a *Constrained* fleet for *IMC* requests. However, unexpected delays are introduced, which also lead to slight increases for *regular* requests. When 240s of interaction time are assumed for all requests, *special* requests equally benefit, but rejection rates of *regular* requests are heavily increased, indicating less optimal exploitation of the system.

For **TP**, Figure 3 shows the results when varying the targeted planning horizon H with 5 minutes representing the baseline. The solid and dashed lines compare between *regular* and *special* requests, while the plot rows indicate their type. We notice that the rejection rate for *special* requests decreases as H increases, and that a value $H > 10min$ is sufficient to a lower their rejection rates even below those of *regular* requests. For the *Tight* fleet, 10min are enough to achieve a near-zero rejection rate. With the *Constrained* fleet, *2P* and *3P* are nearly fully served from $H = 15min$ whereas $H \geq 20min$ is needed for *IMC* requests.

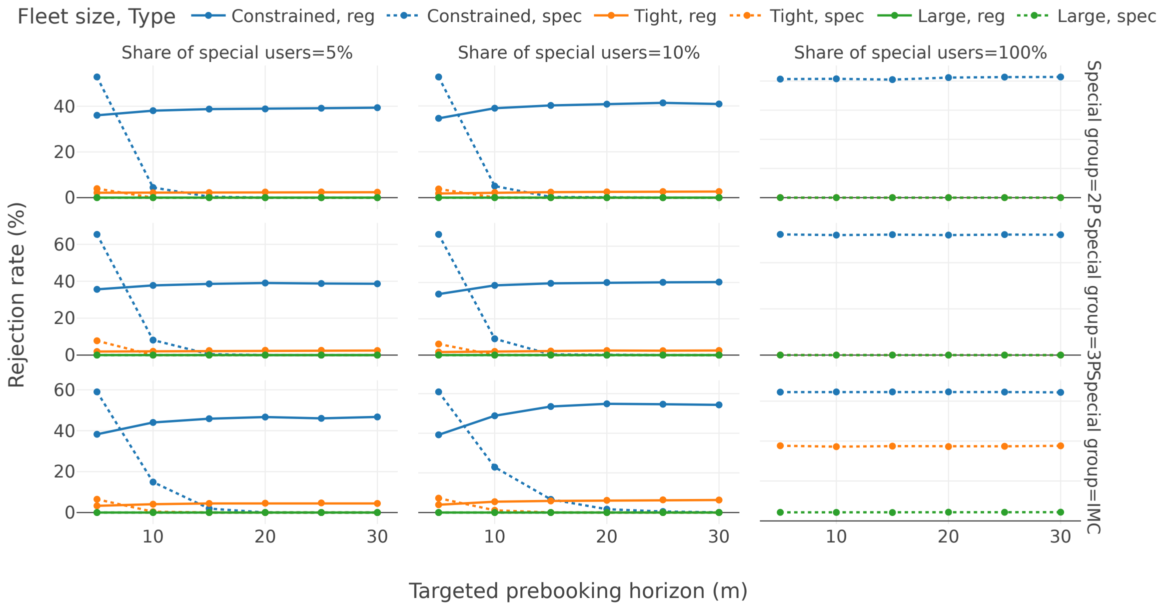

Figure 3 – *Rejection rates observed under the application of the TP mitigation strategy*

4 CONCLUSION

Our results exemplify how to uncover algorithmic discrimination in fleet management algorithms and explore the potential of two mitigation strategies. We show that fine-tuning the strategies to the specific use case is necessary, and that simulation-based analyses can help in doing so.

The practical implications of the tested strategies need further discussion in future work. The UAS strategy may imply forcing the operator to comply with nondiscriminatory regulation that lowers its global performance. On the other hand, the TP strategy puts the burden on the concerned user groups to plan their travels further in advance than others.

The strategies proposed here do not rely on structural changes in the algorithms, a pathway that has been followed in literature (Chouaki & Hörl, 2024). Our goal for future research is to map discriminatory behavior over various fleet management algorithms and propose a larger series of mitigation mechanisms.

References

Bischoff, Joschka, Führer, Karoline, & Maciejewski, Michal. 2019. Impact assessment of autonomous DRT systems. *Transportation Research Procedia*, **41**(Jan.), 440–446.

Chouaki, Tarek, & Hörl, Sebastian. 2024. Comparative assessment of fairness in on-demand fleet management algorithms. *In: hEART2024*.

Hörl, Sebastian, & Balac, Milos. 2021. Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data. *Transportation Research Part C: Emerging Technologies*, **130**, 103291.

Hörl, Sebastian, Chouaki, Tarek, Ludwig, Oliver, Rewald, Hannes, & Axer, Steffen. 2024. Evaluating prebooked on-demand mobility services using MATSim. *Procedia Computer Science*, **238**, 763–770.

Horni, Andreas, Nagel, Kai, & Axhausen, Kay W. (eds). 2016. *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press.

Jing, Peng, Hu, Hanbin, Zhan, Fengping, Chen, Yuexia, & Shi, Yuji. 2020. Agent-Based Simulation of Autonomous Vehicles: A Systematic Literature Review. *IEEE Access*, **8**, 79089–79103.