FairVizARD: A Visualization System for Assessing Fairness of Ride-Sharing Matching Algorithms*

Ashwin Kumar,[1] Sanket Shah, [2] Meghna Lowalekar,[†3] Pradeep Varakantham,[4] Alvitta Ottley,[1] William Yeoh[1]

[1]Washington University in St. Louis, [2]Harvard University, [3]Swiggy, [4]Singapore Management University
ashwinkumar@wustl.edu, sanketshah@g.harvard.edu, meghnal.2015@phdcs.smu.edu.sg, pradeepv@smu.edu.sg, alvitta@wustl.edu, wyeoh@wustl.edu

Abstract

There is growing interest in algorithms that match passengers with drivers in ride-sharing problems, and it is of interest to visualize pertinent information of the output of these algorithms in order to evaluate their performance across different metrics and tradeoffs between them. In this paper, we introduce a system, called FairVizARD, that visualizes the output of these algorithms across different geographic and temporal resolutions, allowing users to easily compare the performance and fairness of multiple ride-sharing allocation algorithms.

Introduction

On-demand taxi ride-sharing systems (e.g., UberPool and Lyft Line) have seen increased adoption over the past few years, as they provide several benefits to the various parties in the system. Passengers are incentivized by having a discount on the fare of their ride compared to a traditional non-shared ride. Drivers are incentivized by having an increase in earnings due to the larger number of passengers they can drive in each trip, and the Ride-Hailing Companies (RHCs) are incentivized by having an increase in the number of passengers requesting rides, leading to an increase in the revenue of the system.

However, the possibility of a taxi driver getting requests from multiple passengers at different pickup and dropoff locations introduces an additional layer of complexity to *matching algorithms* that RHCs employ to match requests by passengers to taxi drivers, like accounting for the necessary changes in the route of the taxi and, consequently, the additional delay added to the trip of passengers already in the taxi. Researchers have proposed a number of matching algorithms for ride-sharing systems by optimizing different objectives (e.g., minimizing the length of the detour of passengers, maximizing the average number of requests served by drivers, etc.) while keeping the runtimes of the algorithms small to ensure real-time computation (Alonso-Mora et al. 2017; Shah, Lowalekar, and Varakantham 2020). However, the matches proposed can be *unfair* (i.e, they have underlying biases that disadvantage certain parties of the system). For example, a matching algorithm that keep drivers in busy areas to maximize their revenue while ignoring requests from passengers in other areas is biased in favor of drivers and against passengers in non-busy areas. Consequently, there is growing interest in analyzing and improving the fairness of matching algorithms (Lesmana, Zhang, and Bei 2019; Nanda et al. 2020).

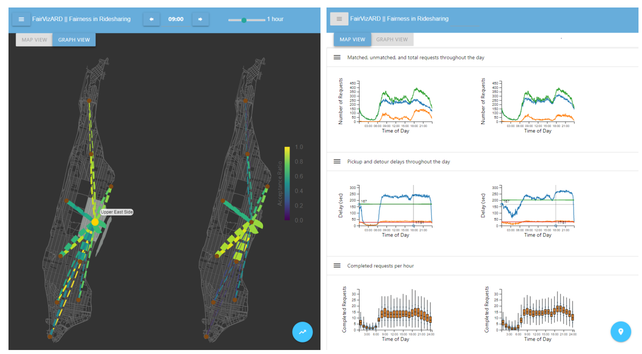

Figure 1: Left to right: (*i*) *Map View* of FairVizARD. (*ii*) *Graph View* of FairVizARD.

Towards this end, we present **FairVizARD**, a **Fair**ness **Viz**ualization for **A**nalysis of **R**ide-sharing **D**ata (see Figure 1).[1] Using a combination of map- and graph-based visualizations, FairVizARD shows spatio-temporal information of the passenger requests and taxi drivers as well as statistics on different factors that may be of interest to different parties in the system. To allow users to easily compare different matching algorithms and understand their tradeoffs (specifically on grounds of passengers served, delays and income distribution among drivers), FairVizARD presents this information for two algorithms side by side, where both algorithms are given the same initial problem configuration. To the best of our knowledge, FairVizARD is the first visualization system that enable users (as well as algorithm designers) to analyze and compare city-scale ride-sharing matching algorithms for fairness and other performance metrics.

*This research is partially supported by National Science Foundation grant #1812619.

†Work done while at Singapore Management University

[1]Video demo can be found at https://youtu.be/0-G-o4E0gKM

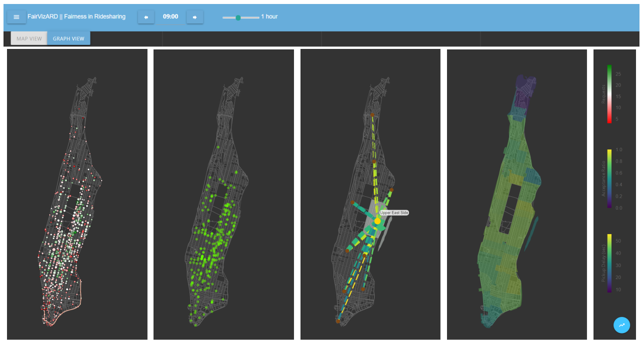

Figure 2: Map Views. Left to right: (*i*) Taxi Data; (*ii*) Request Data; (*iii*) Inter-Zone Data; (*iv*) Individual Zone Data; (*v*) Legend

FairVizARD Design

FairVizARD shows data for a single day (24 hours) with one-minute resolutions. This resolution is decided based on the decision epochs considered by the algorithms we run, and can be scaled as required. We demonstrate FairVizARD using Manhattan, NY as the network graph. All algorithms were run using demand data from the NY Yellow Taxi Dataset[2] with 1000 drivers operating across 24 hours. FairVizARD can be used to look at a single algorithm at a time, or to compare two or more algorithms side by side in a split-screen view.

FairVizARD is composed of two views: A *map view* that visualizes the spatial distribution of the data and a *graph view* that describe the temporal distribution of the data.

Map View: Figure 1(*i*) shows the map view displaying a specific type of information, which we will discuss later. In general, this view shows data at a particular time of day distributed across the map. When aggregated data is shown, then users can also adjust the *time window* over which the data is being aggregated by using a slider. Users can select different kinds of information to display on this view through the use of a sidebar. Each of these sub-views can be stacked, allowing users to look for correlations across different aggregation methods and data. Figure 2 illustrates the four possible sub-views.

- **Taxi Data:** Figure 2(*i*) shows the locations of all taxis at the current time as individual circles, with the size indicating the number of passengers associated with that taxi, and the color indicating the passengers picked over the previous time window, with a diverging color scale to more easily see the spread amongst drivers.
- **Request Data:** Figure 2(*ii*) shows the pickup locations of all the requests received at the current time step as green circles. Users can also toggle on the dropoff locations of those requests, which will be shown as red circles. Additionally, users can also filter the data to just show the matched requests, unmatched requests, or all requests.
- **Inter-Zone Data:** Figure 2(*iii*) shows the relationship between zones in the city, where zones are contiguous regions that are either defined by zip codes or by user-defined neighborhoods. Zones are visualized on the map as nodes located at the centroids of the zones. A directional edge from zone z_i to zone z_j shows trips from z_i to z_j in the previous time window. The thickness of the edge is scaled to the number of trips between those zones. The color of an edge can indicate inter-zone metrics of interest, like acceptance rate.
- **Pickup Delay:** Figure 2(*iv*) illustrates data related to zones, which are either defined by zip codes or user-defined neighborhoods. Each zone is colored according to its *average pickup delay*, which is the average delay incurred because of ride-sharing detours before the passenger is picked up.

[2] https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page

Graph View: The types of graphs are as follows:

- **Incoming Request Graphs:** The graphs at the top of Figure 1(*ii*) show examples, where they plot the number of matched, unmatched, and total requests received against the time of the day that the requests were received.
- **Pickup and Detour Delay Graphs:** The graphs in the middle of Figure 1(*ii*) show examples, where they plot the average pickup and detour delays for requests arriving at each time step as well as their 24-hour averages.
- **Completed Request Graphs:** The graphs at the bottom of Figure 1(*ii*) show examples, where they plot box plots of the distribution of the number of requests completed by taxi drivers throughout the day. Each box plot provides the range of requests completed (i.e., their minimum and maximum values), the median (shown by the horizontal line in the orange rectangle), the second and third quartiles (the ranges corresponding to the bottom and top portions of the rectangle, respectively), and the 10th and 90th percentiles (shown using ticks on the vertical line).

Conclusions

FairVizARD provides high-level information about ride-sharing algorithms at a glance, while allowing users to analyze them based on their notions of fairness. Since there are various stakeholders involved, such a system can aid users in finding metrics of their interest, possibly helping in decision making. We envision that FairVizARD can be used to communicate complicated metrics to the lay person, increasing transparency and providing directions for the improvement of ride-sharing algorithms.

References

Alonso-Mora, J.; Samaranayake, S.; Wallar, A.; Frazzoli, E.; and Rus, D. 2017. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences* 114(3): 462–467.

Lesmana, N. S.; Zhang, X.; and Bei, X. 2019. Balancing Efficiency and Fairness in On-Demand Ridesourcing. In *NeurIPS*, 5310–5320.

Nanda, V.; Xu, P.; Sankararaman, K. A.; Dickerson, J.; and Srinivasan, A. 2020. Balancing the Tradeoff between Profit and Fairness in Rideshare Platforms during High-Demand Hours. In *AAAI*, 2210–2217.

Shah, S.; Lowalekar, M.; and Varakantham, P. 2020. Neural Approximate Dynamic Programming for On-Demand Ride-Pooling. In *AAAI*, 507–515.