ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/328032816

Mobility Sharing as a Preference Matching Problem

Article *in* IEEE Transactions on Intelligent Transportation Systems · October 2018

DOI: 10.1109/TITS.2018.2868366

CITATIONS
0

READS
143

2 authors:

Hongmou Zhang
Massachusetts Institute of Technology
4 PUBLICATIONS **25** CITATIONS

SEE PROFILE

Jinhua Zhao
Massachusetts Institute of Technology
55 PUBLICATIONS **471** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project China's Urbanization and Motorization View project

Project Modeling Behavior Dynamics of Human Mobility using Transit Smart Card Data View project

All content following this page was uploaded by Jinhua Zhao on 12 October 2018.

The user has requested enhancement of the downloaded file.

Mobility Sharing as a Preference Matching Problem

Hongmou Zhang and Jinhua Zhao

***Abstract*—Traffic congestion, dominated by single-occupancy vehicles, reflects not only transportation system inefficiency and negative externalities, but also a sociological state of human isolation. Advances in information and communication technology are enabling the growth of real-time ridesharing to improve system efficiency. While most ridesharing algorithms optimize fellow passenger matching based on efficiency criteria (maximum number of paired trips, minimum total vehicle-time or vehicle-distance traveled), very few explicitly consider passengers' preference for their peers as the matching objective. Existing literature either considers the bipartite driver–passenger matching problem, which is structurally different from the monopartite passenger–passenger matching, or only considers the passenger–passenger problem in a simplified one-origin–multiple-destination setting. We formulate a general monopartite passenger matching model in a road network, and illustrate the model by pairing 301,430 taxi trips in Manhattan in two scenarios: one considering 1,000 randomly generated preference orders, and the other considering four sets of group-based preference orders. In both scenarios, compared with efficiency-based matching models, preference-based matching improves the average ranking of paired fellow passenger to the near-top position of people's preference orders with only a small efficiency loss at the individual level, and a moderate loss at the aggregate level. The near-top-ranking results fall in a narrow range even with the random variance of passenger preference as inputs.**

***Index Terms*—mobility sharing, matching, preference, social interaction.**

I. Introduction

ONLINE mobility on demand (MoD) services emerging with the advance of modern information technology enrich peoples' transportation mode choices. Ridesharing is an MoD service that connects the trips of passengers who can combine their trips with only small increases in travel time, such that a single vehicle can accommodate more than one passenger at a time. Examples of ridesharing services include Lyft Line, UberPool, and GrabHitch. By increasing the occupancy of vehicles, ridesharing has the potential to reduce the number of cars on roads, leading to reduction in traffic congestion and pollution. Lyft reported in 2014 that over fifty percent of their trips in San Francisco, and thirty percent in New York City were shared trips [1].

The current paradigm of mobility sharing is focused on optimizing system efficiency. The underlying assumption is that travelers place no value on the characteristics of their fellow passengers, or if they do, the relative importance is negligible compared to travel time and cost. Researchers have proposed various algorithms to maximize these efficiency benefits with different optimization objectives under different assumptions [2]–[7], and have carried out simulations with real-world data [8]. In the proposed algorithms, trip pairing is modeled as a graph-matching problem, and with different matching strategies, system optimization can be achieved. In a recent study, Alonso-Mora *et al.* investigated the minimum number of vehicles that would be needed by New York City if sharing possibility was optimized [6].

Whereas the efficiency benefits of ridesharing have been extensively studied, ridesharing is less understood as a sociological phenomenon [9]. In contrast to typical social interactions in public or private spaces (meeting rooms, streets, public squares, living rooms, etc.), the nature of shared car rides is impromptu, captive for a considerable duration, and remarkably more intimate—representing a unique juxtaposition of spontaneity and intensity. It is also distinct from mass transit modes such as buses and trains, in which most passengers refrain from engaging each other. For example, Uber promotes their UberPool service as a potential platform for business and job opportunities by meeting new people [10].

In previous studies, researchers have used preference to match drivers and passengers [11], [12], and formulated ridesharing as a bipartite matching problem, taking drivers and passengers as two separate sets of nodes. The mathematical structures of monopartite and bipartite matchings are different [13], [14]. For example, there is guaranteed to be stable solutions for bipartite matching problems, but there may not exist any stable solution for a monopartite matching problem. Therefore, the driver–passenger matching problem and the fellow passenger matching problem are not only operationally different, but also observe different system structures. Furthermore, these studies only considered travel distances and financial costs of trips to represent preference and may thus underestimate the variation and complexity of user preference resulting from social factors.

Thaithatkul *et al.* conducted the first research on using preference for fellow passengers matching [15], [16], and considered trip distance and randomly-generated personal utility for fellow passenger preference. They implemented the model in a one-origin–multiple-destination setting, using artificially generated trip features and synthetic preferences. The one-origin–multiple-destination setting is applicable to real-world scenarios such as trips from an airport, but does not capture the shareability of the general multiple-origin–multiple-destination situations. In this paper, we propose a more generic preference-based fellow passenger matching model, and examine the model properties with empirical taxi trips and the road network in Manhattan.

Berlingerio *et al.* quantified the enjoyability of passengers based on people's interests, social links, and tendency to connect to people with similar or dissimilar interests, mined from social network data, and formulated enjoyability together with the number of cars as a multi-objective linear

The authors are with the Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA. (e-mail: hongmou@mit.edu; jinhua@mit.edu).

programming problem [17]. Enjoyability is closely related to but different from preference, and the linear additivity of enjoyability of people does not fit in our paper, where we use Pareto comparison instead of quantified enjoyability.

Preference matching originated from, and has been widely studied in the context of college admission, marriage, roommate assignment, doctor residency assignment, and kidney exchange [13], [18]–[20]. Each of these problems is one case of matching people to people, or people to institution—students to colleges, doctors to hospitals, or kidney donors to receivers. Furthermore, in each of the settings, each individual—either person or institute—is assumed to have, explicitly or latently, a ranking of objects or people representing their preference. Within such a system, a stable outcome is expected, and can be achieved regarding the preference of all individuals in the system. A stable matching is a matching with which no individual could be better off by changing the match without making at least one other individual worse off. In other words, a stable matching is an optimal matching in a socially Pareto sense.

In our paper, we explicitly consider passengers' preference for their fellow passengers as the trip matching objective, formulate ridesharing as a monopartite preference matching problem, and compare the matching outcome with those of efficiency-based methods. We use 301,430 taxi trips in Manhattan [21] (Fig. 1) on a randomly selected day in 2011 (April 24) to illustrate the model: a single set of passengers are paired with each other, and a maximum stable matching as the pairing objective is identified with regard to all passengers' preferences for their peers. We quantify the trade-off between efficiency-based ridesharing methods, and our proposed preference-based matching method.

We note that matching people with preferences and the elicitation of preferences are two distinct tasks. In this paper, we take preferences for fellow passengers as *given* in the form of rank orders, but do not address the issues of the preference elicitation. The matching algorithm we propose can operate with any complete or partial preference orders. At the end of this paper, we discuss the complexity in the preference for fellow passengers and its difference from the preference for other trip attributes. The complexity of preference for fellow passengers exists in its structure—being heterogeneous, dynamic, and more about compatibility than similarity—and in its elicitation. In the discussion section we propose possible preference elicitation methods with associated challenges.

We also emphasize that not all preferences are respectable. [22] and [23] demonstrated the ethical concerns surrounding ridesharing. Although an ethical discussion of preference in ridesharing is beyond the scope of this paper, we want to point out that it is critical for the society as a whole to draw the boundary between acceptable and unacceptable articulations of preferences. A positive understanding of the role of preference in ridesharing is an important prerequisite for addressing those concerns.

II. Methods

Formally, we consider the vehicle trips in a city to be in a shareability network [5]. A shareability network is an undirected graph $G(V, E)$ where each node in the node set V represents a trip, and $E = \{\{v_i, v_j\} | v_i, v_j \in V\}$ is the edge set indicating whether two trips are shareable. For example, for trips $v_i, v_j \in V$, if v_i and v_j can be shared, we have $\{v_i, v_j\} \in E$. The criterion used to determine whether two trips are shareable in this paper follows the "cap of maximal detour" rule [5]. Two trips are shareable if there exists a new route that can connect the origins and destinations of the two trips, and the additional travel time for either trip does not increase more than Δt, a predetermined parameter set by the system designer. In other words, in a shareability network with parameter Δt, if a trip would take a person time t to travel individually, the travel time would never be more than $t + \Delta t$ if the trip is shared with another allowed trip. In the scope of this paper, we limit our discussion to the ridesharing of two parties of passengers.

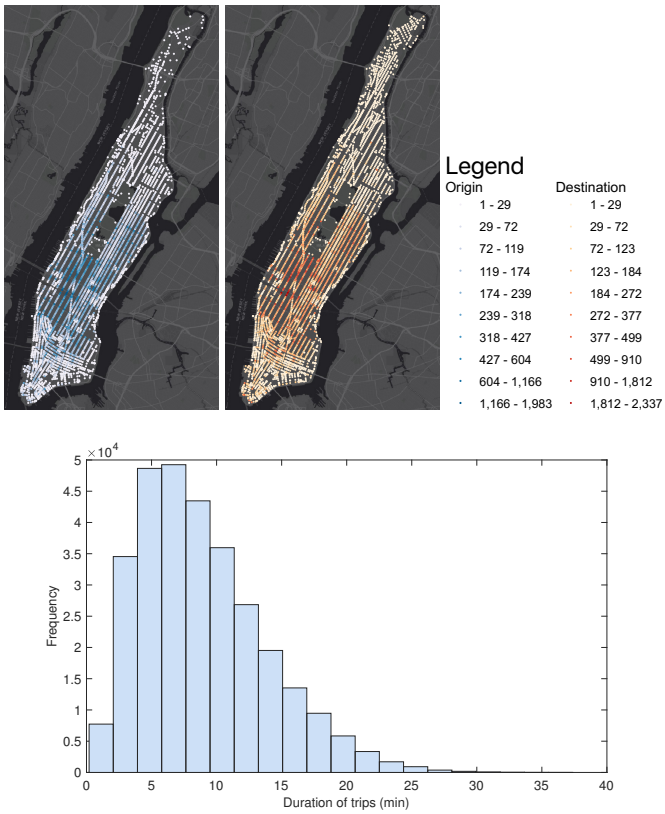


Fig. 1. Distribution of the origins and destinations of taxi trips in Manhattan and distribution of trip duration without sharing

The departure and arrival time constraints are also considered when building the shareability network. For any trip pairs, if the actual departure time of any of the two when sharing is later than the original departure time without sharing plus Δt, the trip pair is considered non-shareable, and similarly for arrival time [5].

Therefore, detour time and waiting time, or the difference between the original departure time without sharing and the pick-up time with sharing are unified as Δt, or the "passenger discomfort parameter", in the configuration of shareability networks. In other words, if two trips are shareable, neither

the detour time of each party, nor the difference between the actual pick-up time of the shared trip and the original departure time can be greater than Δt.

Fig. 2 illustrates the distribution of node degree of the shareability network built on the taxi trips in Manhattan in one day. For example, when Δt is 300 seconds, the majority of trips have approximately 100 shareable trips, and the maximum number of shareable trips that a trip has is approximately 900. Both numbers increase with Δt as sharing possibility increases with longer allowed detours.

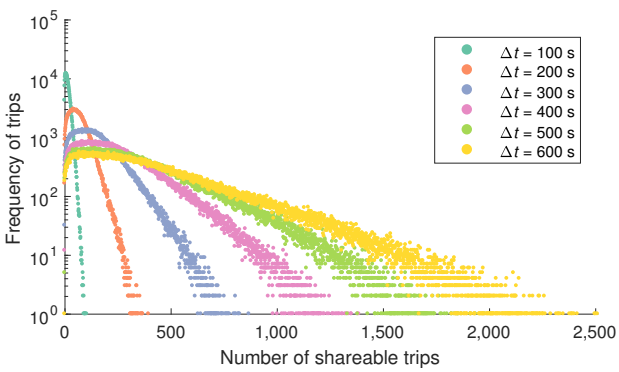


Fig. 2. Frequency of the number of shareable trips with a trip in shareability networks with Δt = 100–600 s (each point represents the frequency of trips with a certain number of shareable trips)

Upon the shareability network, we define a weight function $\omega : E \to \mathbb{R}$ indicating the efficiency benefits of sharing two trips. In this paper, we use the savings of vehicle-minutes (veh-min) and vehicle-kilometers-traveled (VKT) as the measures ω of efficiency. The travel time on each road link is estimated using the real-world taxi travel time data, and the travel route of each trip is inferred as the one with the closest total travel time to the actual time [5]. In this paper, we assume that the sharing of trips does not change the overall underlying traffic condition, i.e., we do not consider the feedback between traffic congestion and ridesharing, which is an important future research direction when the shared trips start to contribute to a high proportion of the overall traffic.

We then consider the preference of the passenger in each trip for all other passengers. As discussed by [24] and [25], the reasons for a passenger preferring a certain fellow passenger to another vary. Here, we take the preferences as exogenous inputs to the problem, and the matching algorithm we develop can operate given any preference rank order.

If a shared trip consists only of individual passengers, the preference is individual preference. In other cases where there are more than one passenger in a party, we assume that all the passengers in the party are able to consent on an order; thus, we can consider each party equivalent to a single passenger. In the remainder of this paper, we treat trips and passengers as equal. Hence, for passenger v_i in V, his/her preference can be denoted as an ordered list $v_i : v_{k_1} \succ v_{k_2} \succ \ldots \succ v_{k_m}$, where $k_1, k_2, \ldots, k_m$ is the permutation of a sublist of $\{1, 2, \ldots, n\} \backslash i$, and n is the number of passengers in the system.

Note that the preference list for a passenger does not have to be a complete list containing all other passengers in the system, but only needs to contain all the *shareable* passengers, or the "neighbors" in $G(V, E)$. Nevertheless, we need to require that the preference lists be symmetrically compatible—if passenger v_i is on passenger v_j's list, passenger v_j also needs to be on passenger v_i's list.

A matching M is a subset of E, and it requires that $\forall e_i, e_j \in M, e_i \cap e_j = \varnothing$ such that no passenger is sharing with more than one other passenger. We denote the set of all feasible matchings of a shareability network G as $\mathcal{M}(G)$. A stable matching is a special matching M' such that there are no two passengers in the system who both prefer each other to their paired fellow passengers, or formally $\nexists e_i = \{v_{i_1}, v_{i_2}\}, e_j = \{v_{j_1}, v_{j_2}\} \in M'$, $v_{i_1} : v_{j_1} \succ v_{i_2}$ and $v_{j_1} : v_{i_1} \succ v_{j_2}$. The set of stable matchings is denoted as $\mathcal{M}'(G) \subset \mathcal{M}(G)$.

Gale and Shapley found that for such a system, referred to as the "stable roommate problem," a stable matching consisting of all passengers does not always exist [13]. Irving proposed a method that finds a stable matching if one exists [20]. When there is no stable matching for all people in the system, an alternative objective is to find a stable matching on a maximum subset of V—a "maximum stable matching" [14]—and Tan proposed an algorithm to find such a solution [26]. Therefore, we can look for the matching that is stable on the subset of V with maximal cardinality:

$$M_{\text{pref}} = \underset{M \in \mathcal{M}'(G)}{\operatorname{argmax}} \{|M|\}. \tag{1}$$

Because there may be more than one such matching, M_{pref} can be a subset instead of a single element in $\mathcal{M}$.

Further, we wish to determine the efficiency trade-off arising from the use of stable-preference matching; therefore, we compare it with two efficiency-based matching methods, maximum cardinality matching (MC), which maximizes the number of shared trips (n(shared trips)), and maximum weight matching (MW), which minimizes total system veh-min or VKT:

$$M_{\text{mc}} = \underset{M \in \mathcal{M}(G)}{\operatorname{argmax}} \{|M|\}, \tag{2}$$

$$M_{\text{mw}} = \underset{M \in \mathcal{M}(G)}{\operatorname{argmax}} \left\{ \sum_{e \in M} \omega(e) \right\}. \tag{3}$$

As $\mathcal{M}' \subset \mathcal{M}$, $\forall M' \in M_{\text{pref}}$ and $M \in M_{\text{mc}}$, there always is $|M'| \leqslant |M|$.

A. Synthetic preference orders

We use two types of preference inputs in this paper. 1) Random preference orders: for each node in the shareability network, we generated a random permutation of all the neighboring nodes and used the permutation as the preference order. We repeated the random permutation assignment 1,000 times to generate sufficient randomness in people's preference for others, and delineated the range of possible results, which turned out to be very narrow. 2) Group-based preference orders: we acknowledge that various factors can be related to people's preference for fellow passengers, ranging from gender, age, and income level, to personal interest, political affiliation, hobbies, and talkativeness; therefore, we implemented group-based preference assuming that each passenger belongs

to one of two groups based on one binary characteristic. S1–S4 are four scenarios that test different group shares, and different preference symmetricity assumptions between groups.

- **S0.** One group: randomly assigned preference orders;
- **S1.** Even, symmetric: 50% of passengers in Group 1, 50% in Group 2; people in both groups prefer fellow passengers from their own group to those from the other;
- **S2.** Even, asymmetric: 50% of passengers in Group 1, 50% in Group 2; people in Group 1 prefer fellow passengers from their own group to those from Group 2, whereas people in Group 2 are indifferent;
- **S3.** Uneven, symmetric: 20% of passengers in Group 1, 80% in Group 2; people in both groups prefer fellow passengers from their own group to those from the other;
- **S4.** Uneven, asymmetric: 20% of passengers in Group 1, 80% in Group 2; people in Group 1 prefer fellow passengers from their own group to those from Group 2, whereas people in Group 2 are indifferent;

We first randomly assigned trips into two groups based on the group share hyperparamter. We assumed that both groups are evenly distributed in space. Then, for each group, we assigned preference based on the group type. For groups in which people prefer same-group fellow passengers, we first determined whether the neighboring nodes were in the same group or the other group, and carried out random permutation on each of the two neighbor sets. We concatenated the two permuted lists as the preference lists. In Appendix A, the order of magnitude of the number of preference possibilities and the reduction of possibilities when groups exist are discussed.

We acknowledge that in a real-world ridesharing platform it is impossible to ask users to rank all the possible fellow passengers. This paper focuses on the matching process but not the the preference elicitation process. We develop a general matching algorithm which can operate given any complete or partial preference rank orders. In addition the complete or partial rank orders for fellow passengers do not have to be explicitly given by the users, but may also be derived based on user behaviors, travel history, or relevant information that the users provided to the service provider. We will comment on both the potential technical and ethical issues in the discussion section.

B. Irving–Tan algorithm

We build our stable-preference matching model following the Irving–Tan algorithm [20], [26]. There are two steps in the algorithm. In the first step (Algorithm 1), each passenger proposes to the top choice in their preference lists, and the proposed passenger could either hold the proposal if there is no better choice, or reject the proposal if a better choice is already present. If a passenger accepts a proposal while holding another proposal, the passenger rejects the originally held proposal. The rejected passenger will then have to re-propose. The process is repeated until each passenger is accepted by another passenger, or the preference list has been exhausted.

After the preference list reduction, the preference lists are significantly reduced. By symmetricity of each operation in step one, if there is only one fellow passenger v_j left in passenger v_i's preference list, v_i should also be the only one left in v_j's preference list. Therefore, for the passengers with reduced preference lists of length zero or one, the matching is complete. For passengers with preference lists of length at least two, there must be "rotations" that include at least three passengers. In the second step of the algorithm (Algorithm 2), one rotation is identified in each iteration. A rotation corresponds to a subset of the shareability network, in which the matching solution will be local to the nodes within it and independent of the rest of the graph. In other words, as long as a rotation R is identified from the network, the problem can be divided into a matching m_1 on R and a matching m_2 on $G \backslash R$, and the overall matching will be $m = m_1 \cup m_2$.

Algorithm 1 Preference list reduction: proposal and rejection

```
for each passenger v_i ∈ V do
    propose to the first one in the preference list
    if the proposed passenger does not have a better option then
        hold the proposal
    else
        reject the proposal and delete each other in the preference lists
    end if
    while rejected or preference list is not empty do
        propose to the next one
    end while
end for
for each passenger v_i ∈ V do
    delete all the choices with rankings larger than the proposal it holds
end for
```

There are two types of rotations, odd rotation and non-odd rotation. An odd rotation is a rotation with an odd number of elements. For example, consider the following case, in which no stable matching exists for all four elements.

$$\begin{aligned} &1\colon 2 \succ 3 \succ 4 \\ &2\colon 3 \succ 1 \succ 4 \\ &3\colon 1 \succ 2 \succ 4 \\ &4\colon 1 \succ 2 \succ 3 \end{aligned}$$

After step one, the preference lists will be reduced to the following table, and $\{1, 2, 3\}$ can be found as an odd rotation.

$$\begin{aligned} &1\colon 2 \succ 3 \\ &2\colon 3 \succ 2 \\ &3\colon 1 \succ 2 \\ &4\colon \varnothing \end{aligned}$$

To find a maximum stable matching for an odd rotation, one element needs to be eliminated from the list randomly [26].

In the other type of rotation, the non-odd rotation, there is an even number of elements. For example, in the following case, $\{1, 2, 4, 3\}$ is a non-odd rotation. In this case, the rotation can be further removed from the preference lists as described by [20] and [26]. After all rotations have been removed from the preference lists, each person should hold at most one choice in

1: $2 \succ 4 \succ 3$
2: $4 \succ 1$
3: $1 \succ 4$
4: $3 \succ 1 \succ 2$

Algorithm 2 Matching with the reduced preference lists

while there is rotation in the preference lists **do**
  if it is an odd rotation **then**
    Randomly delete an element from the odd rotation
    m_1 = matching(the odd rotation \ the deleted element)
    m_2 = matching(the rest of the preference lists)
    return $m_1 + m_2$
  else
    delete the rotation from the preference lists
  end if
end while

the corresponding preference list, which gives the maximum stable matching. The time complexity of this algorithm is $O(n^2)$ [26].

For the efficiency-based matching methods, we used the Python package NetworkX [27] for the maximum cardinality matching, and simple greedy algorithm for the maximum weight matching. The algorithm keeps pairing nodes with the highest weighted edges until no further matches can be made.

III. Results

First, we conducted the efficiency-based and preference-based matchings for the shareability network of taxi trips in one day in Manhattan with Δt = 300 s, and repeated the test for 1,000 random preference assignments to show the variability of the results. The solution time for each run was about eight minutes. Second, in order to estimate the impact of Δt on the convergence of preference matching, we performed sensitivity analysis for preference-based matching with a range of Δt = 100–600 s for 1,000 runs each. Third, for each of the four two-group scenarios, we conducted preference-based matching for 100 structured random assignments of preference.

A. Efficiency vs. preference

Table I summarizes the efficiency measures under different matchings. The table shows that preference-based matching results in only a marginally lower matching rate than MC matching. Further, the rate is 3.8% higher than those of MW matchings. For other efficiency measures, system-wise preference-based matching behaves similar to MC matching and has longer distances and travel times for each *vehicle trip* on average. However, for each *individual passenger*, the difference in travel time under preference-based matching and MW matchings is only about 40 seconds. Fig. 4 shows more details of the efficiency trade-off from a passenger's perspective. Fig. 4a shows the distribution of increase in travel time (detour); preference-based matching has more long detours than MW matchings, with all detours still capped by Δt. Fig. 4b depicts the cumulative distribution of the detour time as a proportion of the total travel time if the trip is not shared. MW matchings have 96.8% trips with detour of less than 10% of the non-sharing travel time, whereas for preference-based matching, the number is still as high as 92.8%.

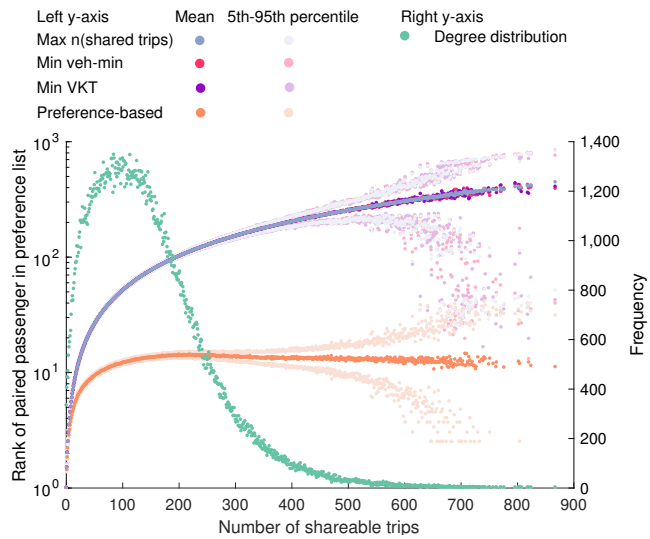


Fig. 3. Ranking of paired fellow passengers in respective passengers' preference lists (left y-axis). The darker dots show the average values of results, and the lighter dots are the 5th and 95th-percentile values; The three efficiency-based matchings almost overlap; Degree distribution: the frequency of the number of shareable trips of a given trip (right y-axis), a replicate of the Δt = 300 s curve of Fig. 2 for reference.

Fig. 3 shows the performance of the matchings in terms of preference. The y-axis on the left, in logarithmic scale, shows the ranking of the actually paired fellow passenger in the corresponding passenger's preference list, averaged across 1,000 runs, with regard to the x-axis—the number of shareable trips that the passenger has. For example, under both MC and MW matchings, for all passengers with 100 shareable fellow passengers, the average ranking of matched passengers are their 50th preferred choices, whereas under preference-based matching, the average ranking is slightly more than 10. The curves of both MC and MW approximate $k/2$, where k is the number of shareable fellow passengers—on average efficiency-based matchings pair a passenger to an *average* preferred fellow passenger as we assume preference is independent of space and time. Preference-based matching asymptotically approaches the tenth most favorable choice even with a k as large as several hundreds.

In the mean values and the gap between the 5th and 95th percentiles in Fig. 3, the range of results is very narrow across 1,000 repeat tests with randomly generated preferences, especially when the number of shareable trips is not sufficiently large. For example, for passengers with 300 shareable trips, the average rankings of their mean, 5th, and 95th-percentile-matched fellow passengers are 13.8, 12.6, and 15, respectively. The narrow range indicates that this near-top pairing performance is not dependent on specific preference orders.

TABLE I
EFFICIENCY MEASURES FOR EFFICIENCY-BASED MODELS AND PREFERENCE-BASED MODEL (Δt = 300 s)

	No sharing†	Max. n(shared trips) matching (MC)	Min. veh-min matching (MW_1)	Min. VKT matching (MW_2)	Preference-based matching
Avg. number of matched trips	0	297,818 (98.8%)	285,856 (94.8%)	285,902 (94.8%)	297,049.5 (98.6%)
Total system veh-min	2,686,933	2,391,910	1,689,484	1,689,465	2,341,199
Total system VKT	885,536	788,132	558,415	558,388	771,942
Veh-min per trip	8.91	7.94	5.60	5.60	7.77
VKT per trip	2.94	2.61	1.85	1.85	2.56
Passenger travel time per trip (min)	8.91	12.15	11.42	11.42	12.14
Total number of trips	301,430				
Trips with at least one shareable trips	300,776 (99.8%)				

†Assuming there is only one passenger in each trip.

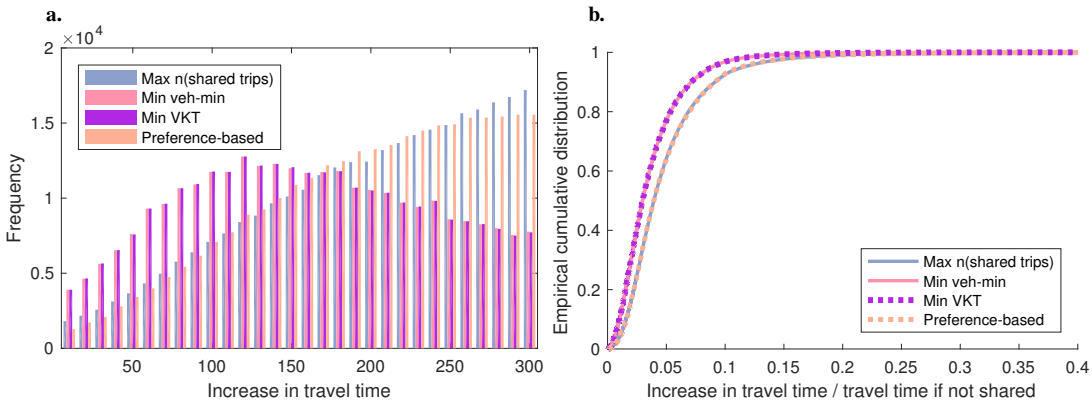


Fig. 4. Distribution of increase in travel time for all matched passengers under different matching methods (**a**) and empirical cumulative distribution of the percentage of increase in travel time in the travel time without sharing (**b**)

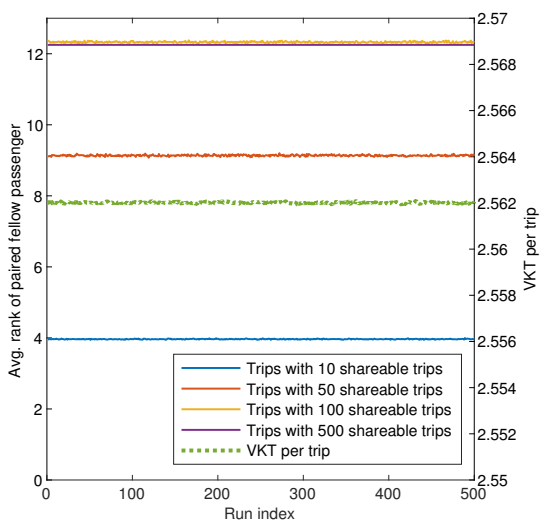


Fig. 5. Stability analysis of matching results

B. Result stability analysis

Since there is randomness in the preference–matching algorithm, we need to understand the stability of matching results in both the efficiency aspect and the preference aspect. To this end we ran the matching algorithm with fixed inputs of preference orders for 500 times, and eliminated odd-rotation elements randomly in each run. As shown in Fig. 5 both the efficiency performance of the algorithm, such as the VKT saving per trip, and preference performance, i.e., the average ranking of paired passengers for those with the same number of shareable trips are very stable across runs (four representative results are shown—the paired passenger's ranking for trips with 10, 50, 100, 500 shareable trips), and the ranges of variance of results are small. For VKT per trip, the range between the maximum value and the minimum value of results in the 500 runs is 2.4×10^{-4}, and for rankings of paired fellow passengers, the variance across runs is much smaller than one rank step in preference.

C. Sensitivity analysis for Δt

While considering five minutes (300 s) as a meaningful detour cap for shared trips, we are also interested in the impact of Δt on the performance of the preference-based matching model.

Fig. 6 shows the results of preference-based matching with Δt from 100 s to 600 s. In Fig. 6a, the shapes of the curves are similar but the curves have different values. In the shareability network with a larger Δt, the average ranking of paired fellow passenger is worse for trips with a given number of shareable trips. For example, for trips with 200 shareable trips, with Δt = 200 s, the average ranking of paired fellow passenger is 9.4, whereas for Δt = 400 s, the ranking is 16.8, and for Δt = 600 s it is 20.3. However, it is worth noting that the trips with 200 shareable trips are no longer the same in the three cases. For any given trip, the larger Δt is, the more shareable trips it will have. This explains why the curves stretch to the right as Δt increases.

Further, as Δt keeps increasing, the gap between curves becomes increasingly smaller—leading us to expect a limiting curve when $\Delta t \to \infty$. This is because when Δt is sufficiently large, i.e., when very long detours are allowed, most trips are shareable with each other, and the shareability network becomes very dense, which eventually becomes a complete graph. In other words, all trips are shareable if no detour limit is applied. Increasing Δt at this time does not add any shareability, thus the curve reaches the limit position. Therefore, while increasing Δt makes the ranking worse, it has a limit.

At the same time, although in an absolute sense the ranking gets worse as Δt increases, the ranking still gets better relatively, as shown in Fig. 6b. In this graph, the ranking is normalized by the number of shareable passengers, and the number of shareable trips is normalized by the highest number of shareable trips in the graph. For example, if we denote the largest number of shareable trips that a passenger has as K, when Δt = 100 s, the people who have $20\%K$ shareable trips will be paired with their top 21.6% preferred fellow passengers, whereas for Δt = 300 s, the ranking improves to the top 8%, and for Δt = 600 s, it further improves to the top 4.5%.

D. Two-group scenarios

Table II and Fig. 7 summarize the results for preference-based matching under the four two-group scenarios with the one-group scenario S0 for reference. For scenarios S1–S4, the efficiency measures show that the preference for a specific subgroup of fellow passengers leads to fewer shared trips. The comparison between the symmetric scenarios (S1, S3) and asymmetric scenarios (S2, S4) show that the more people there holding group-based preferences, the lower the probability of sharing will be. However, the greatest difference between any two scenarios is only 1.1%.

Comparing the same-group pairing rates, in S2 we find that the groups that are indifferent (G2) have slightly lower chances of being paired with same-group passengers. Moreover, in S3 and S4, the chance of being paired with passengers from the same group is lower for the "minority" group (G1), even if the majority group is indifferent, although the difference is only 3–4%, and considering the group splits, 20% and 80%.

From Fig. 7, it is clear that the average ranking of paired fellow passenger is worse when both groups prefer to be paired with passengers from the same group (S1, S3), compared to if only one group does (S2, S4). However, this difference diminishes when the group split is uneven, as the gap between the curves of S3 and S4 is smaller than that of S1 and S2. The best-performing preference matching is observed in the even-group symmetric scenario (S1).

IV. Discussion

This paper formulated the passenger matching in a mobility sharing system as a monopartite matching problem, and examined the trade-off between this matching model with that arising from efficiency-based matching models. The results show that with only a small efficiency loss at the individual level, and a moderate one at the aggregate level, the improvement in preference ranking is substantial—from the average to the near-top. We also found that increases in the detour cap Δt lead to slightly—and boundedly—worse preference rankings in absolute sense, but to better rankings in relative sense. Based on the actual context and system design objective, ridesharing system designers can make the decision of which matching strategy to use, or how to combine multiple strategies.

We also modeled a two-group preference structure for fellow passengers in addition to the simple random case. The preference for a specific group of fellow passengers leads to lower pairing rates—the more people holding it, the lower the pairing rate will be. The same-group pairing rate is lower for the minority group.

It is important to distinguish the question of obtaining the ridesharing preferences from the question of developing the matching algorithm that utilizes such preferences. This paper focused on the latter question and developed the algorithm that can generate the optimal matching output with any set of input preference rank orders. The algorithm is independent of the input ridesharing preferences. We now comment on the former question. The challenge of obtaining the ridesharing preferences of the users is twofold: the complexity in the structure of preference for fellow passengers, and the complexity of eliciting such preferences.

We have identified at least three aspects of the structure of preference for fellow passengers in ridesharing that are different from the preference for typical travel attributes such as travel time and travel cost: 1) Heterogeneity—there is a higher degree of heterogeneity across individuals in the preference for fellow passengers: some people really enjoy the shared ride with a fellow passenger while others strongly prefer riding alone; 2) Dynamism—e.g. even for the same person he or she may want to be silent in the morning ride while hope to engage a conversation with a fellow passenger in the afternoon. The preference for fellow passengers is dynamic and transient; 3) Compatibility—since the pleasure or displeasure of a shared ride is a result of the co-production by fellow passengers, it is more about compatibility between both passengers than the absolute quality of each individual. An extreme example is smoking: a smoker may like to be paired with another smoker, even though smoking per se is a bad behavior for most people. (Of course smoking is typically banned in ridesharing in most cities for this example to be relevant.) Furthermore, compatibility does not necessarily require similarity. Some people prefer to be paired with people similar to them, but others may well like being paired with people different from them.

The elicitation of preference represents another layer of difficulty. There are several possible approaches to eliciting the preference for fellow passengers but each has its challenges: 1) directly ask users to identify the characteristics they prefer to see in their fellow passengers, such as profession, hobby, personal interests, etc. 2) train a machine learning model with the characteristics of paired passengers in previous trips and their post-trip ratings, and 3) start by giving users options of possible fellow passengers, and train models based on their

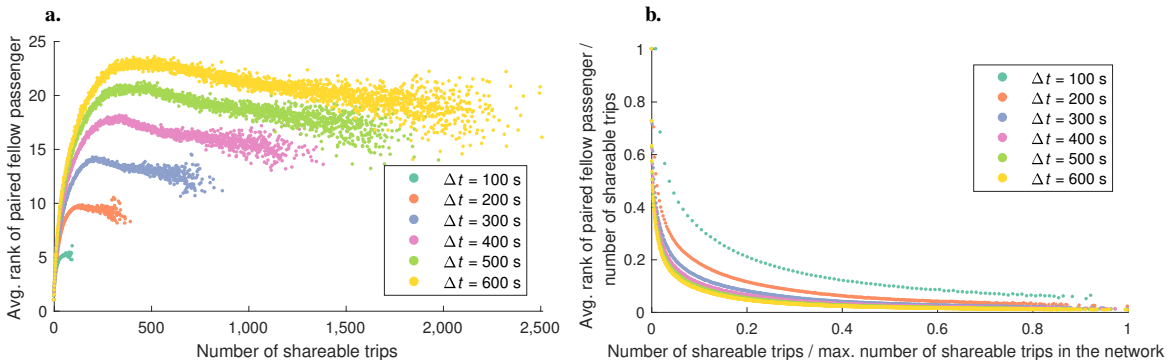


Fig. 6. Average ranking of paired fellow passengers with regard to the number of shareable trips for Δt = 100–600 s (**a**: absolute values; **b**: normalized values)

TABLE II
MATCHING RESULTS FOR TWO-GROUP SCENARIOS

Scenario	S0	S1	S2	S3	S4
Avg. number of matched trips	297,050 (98.6%)	294,494 (97.7%)	296,040 (98.2%)	294,629 (97.7%)	296,530 (98.4%)
Veh-min per trip	7.77	7.77	7.77	7.77	7.77
Paired with passenger in the same group	n.a.	G1: 97.6% G2: 97.6%	G1: 97.6% G2: 96.1%	G1: 94.4% G2: 98.5%	G1: 94.3% G2: 97.4%

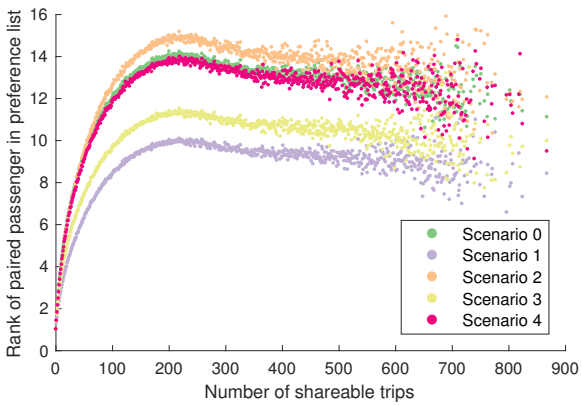


Fig. 7. Average ranking of paired fellow passenger under two-group scenarios

choices in each trip and their characteristics. There are two different types of challenges: first, passengers may not be able to articulate the preferences, the descriptions may be ambiguous, and disparity may exist between what people say and what they actually do, i.e., between stated preference and revealed preference (passengers may not want to express their actual preference for various reasons); second, a deeper challenge is that certain preferences for fellow passengers may not be appropriate or respectable, such as discriminatory attitude and behavior. Such preferences can either be explicitly expressed in the approach 1 or implicitly embedded in the past behavior and codified into machine learning algorithms in approaches 2 and 3. There are important questions to be addressed. For example, what are the boundaries between acceptable and unacceptable preferences? People may see less controversy when gender is used as a preference factor for security reasons, but race as a preference factor is definitely unacceptable. Often there are factors that are acceptable by some but not by others. Who (or which institutions) shall have the authority to determine which preference or preference factor is respectable and which is not? Since the transportation network companies are the designers of the ridesharing platforms, whether and how they shall be regulated in this regard? There remains a major gap in the development of both the social norms and the regulatory frameworks that can guide the ridesharing behavior and the associated system and policy design.

The complexity of preference structure and the process of eliciting preference for fellow passengers demand a thorough discussion on the behavioral, ethical, and institutional aspects, and are beyond the scope of this paper. We identify this as a critical direction for future research.

Further research may also examine 1) the spatial and social heterogeneity of sharing preferences; 2) the interaction between ridesharing and congestion; 3) the pricing of ridesharing services when preference is incorporated; and 4) the vehicle routing for preference-based ridesharing.

APPENDIX A
UNIVERSE OF PREFERENCE

In the shareability network $G(V, E)$, denote the degree of $v \in V$ as $\deg(v) = |\{u, \{u, v\} \in E\}|$. Then, the number of possible permutations as this passenger's preference for shareable fellow passengers is $\deg(v)!$. If the preferences of passengers are independent, the total possible number of preference assignments is

$$N = \prod_{v \in V} [\deg(v)!].$$

In the two-group scenarios, the preferences of passengers are not totally independent, but constrained by the group affiliation of neighboring nodes. If both groups can be considered as evenly distributed in space, like the case for gender, the

expected number of same-group passengers and passengers of the other group in the neighbors of a node are both approximately $\deg(v)/2$. Then, if the preference for one group is always better than that for the other, the total number of preference assignments would be

$$N' = \prod_{v \in V} \left\{ \left[\frac{\deg(v)}{2} \right] ! \right\}^2 .$$

Because when $x \to +\infty$

$$\frac{\Gamma(x/2)^2}{\Gamma(x)} \to 0,$$

the number of preference possibilities is much smaller than without group-based preference.

However, the above discussion is under the assumption that the preferences of passengers are independent, or are only constrained by group affiliation. In this case, the probability of each element in this possible set of preferences is uniform. If the preferences of people are not independent, e.g., some people are popular among all the others, or some are unwelcome by all the others, the universe would be much smaller. When interdependence of preference involves hard constraints, or some cases in the universe are extremely impossible, the probability of each preference assignment is stochastic with a distribution.

Acknowledgment

We thank Paolo Santi and Giovanni Resta for sharing the original shareability network results. We thank the TRB paper review coordinator Abolfazl Mohammadian, the anonymous reviewers for their invaluable comments. We thank the Singapore–MIT Alliance for Research and Technology (SMART) Future Mobility program and the MIT Institute of Data, System and Society Seed Fund for partial financial support for this research.

References

[1] T. Soper, "Lyft's carpooling service now makes up 50% of rides in San Francisco; 30% in NYC," Apr. 2015. [Online]. Available: http://www.geekwire.com/2015/lyfts-carpooling-service-now-makes-up-50-of-rides-in-san-francisco-30-in-nyc/

[2] N. Agatz, A. Erera, M. Savelsbergh, and X. Wang, "Optimization for dynamic ride-sharing: A review," *Eur. J. Oper. Res.*, vol. 223, no. 2, pp. 295–303, Dec. 2012.

[3] M. Furuhata, M. Dessouky, F. Ordóñez, M.-E. Brunet, X. Wang, and S. Koenig, "Ridesharing: The state-of-the-art and future directions," *Transp. Res. B: Methodol.*, vol. 57, pp. 28–46, Nov. 2013.

[4] H. Hosni, J. Naoum-Sawaya, and H. Artail, "The shared-taxi problem: Formulation and solution methods," *Transp. Res. B: Methodol.*, vol. 70, pp. 303–318, Dec. 2014.

[5] P. Santi, G. Resta, M. Szell, S. Sobolevsky, S. H. Strogatz, and C. Ratti, "Quantifying the benefits of vehicle pooling with shareability networks." *Proc. Natl. Acad. Sci.*, vol. 111, no. 37, pp. 13 290–13 294, Sep. 2014.

[6] J. Alonso-Mora, S. Samaranayake, A. Wallar, E. Frazzoli, and D. Rus, "On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment," *Proc. Natl. Acad. Sci.*, vol. 114, no. 3, pp. 462–467, Jan. 2017.

[7] D. Pelzer, J. Xiao, D. Zehe, M. H. Lees, A. C. Knoll, and H. Aydt, "A partition-based match making algorithm for dynamic ridesharing," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2587–2598, Oct. 2015.

[8] N. Agatz, A. Erera, M. Savelsbergh, and X. Wang, "Dynamic ride-sharing: A simulation study in metro atlanta," *Transp. Res. B: Methodol.*, vol. 45, no. 9, pp. 1450–1464, Nov. 2011.

[9] M. Sheller and J. Urry, "The new mobilities paradigm," *Environ. Plan. A*, vol. 38, no. 2, pp. 207–226, Feb. 2006.

[10] T. Manion, "uberPOOL riders make a splash," Nov. 2014. [Online]. Available: https://web.archive.org/web/20170227011407/https://newsroom.uber.com/us-california/uberpool-riders-make-a-splash/

[11] X. Wang, "Optimizing Ride Matches for Dynamic Ride-sharing Systems," Ph.D. dissertation, Georgia Institute of Technology, 2013.

[12] X. Wang, N. Agatz, and A. Erera, "Stable matching for dynamic ride-sharing systems," *Transp. Sci.*, Aug. 2017. [Online]. Available: https://pubsonline.informs.org/doi/10.1287/trsc.2017.0768

[13] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *Am. Math. Mon.*, vol. 69, no. 1, pp. 9–15, 1962.

[14] D. F. Manlove, *Algorithmics of Matching Under Preferences*, ser. Series on Theoretical Computer Science. World Scientific, Apr. 2013, vol. 2, ch. 1.

[15] P. Thaithatkul, T. Seo, T. Kusakabe, and Y. Asakura, "A Passengers Matching Problem in Ridesharing Systems by Considering User Preference," *J. East. Asia Soc. for Transp. Stud.*, vol. 11, pp. 1416–1432, 2015.

[16] ——, "Simulation approach for investigating dynamics of passenger matching problem in smart ridesharing system," *Transp. Res. Procedia*, vol. 21, pp. 29–41, 2017. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S2352146517302120

[17] M. Berlingerio, B. Ghaddar, R. Guidotti, A. Pascale, and A. Sassi, "The GRAAL of carpooling: GReen And sociAL optimization from crowd-sourced data," *Transp. Res. Part C: Emerg. Technol.*, vol. 80, pp. 20–36, Jul. 2017.

[18] A. E. Roth, T. Sönmez, and M. Utku Ünver, "Pairwise kidney exchange," *J. Econ. Theory*, vol. 125, no. 2, pp. 151–188, Dec. 2005.

[19] A. E. Roth, "The evolution of the labor market for medical interns and residents: a case study in game theory," *J. Polit. Econ.*, vol. 92, no. 6, pp. 991–1016, Dec. 1984.

[20] R. W. Irving, "An efficient algorithm for the "stable roommates" problem," *J. Algorithms*, vol. 6, no. 4, pp. 577–595, Dec. 1985.

[21] NYC Taxi & Limousine Commission, "TLC Trip Record Data," Apr. 2011. [Online]. Available: http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml

[22] Y. Ge, C. Knittel, D. MacKenzie, and S. Zoepf, "Racial and gender discrimination in transportation network companies," *NBER Working Paper No. 22776*, Oct. 2016. [Online]. Available: http://www.nber.org/papers/w22776

[23] J. Morales Sarriera, G. Escovar Álvarez, K. Blynn, A. Alesbury, T. Scully, and J. Zhao, "To share or not to share: Investigating the social aspects of dynamic ridesharing," *Transp. Res. Rec.*, vol. 2605, no. 1, pp. 109–117, Jan. 2017.

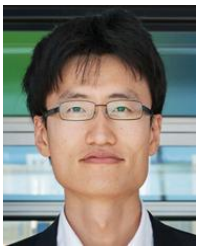
[24] K. J. Dueker, B. O. Bair, and I. P. Levin, "Ride-sharing: Psychological factors," *Transp. Eng. J. Am. Soc. Civ. Eng.*, vol. 103, no. 6, pp. 685–692, Nov. 1977.

[25] K. Ghoseiri, A. Haghani, and M. Hamedi, "Real-time rideshare matching problem," University of Maryland Department of Civil and Environmental Engineering, Tech. Rep., Jan. 2011.

[26] J. J. M. Tan, "A maximum stable matching for the roommates problem," *BIT Numer. Math.*, vol. 30, no. 4, pp. 631–640, Dec. 1990.

[27] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring network structure, dynamics, and function using NetworkX," in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Pasadena, CA USA, 2008, pp. 11–15.

Hongmou Zhang received the Master of City Planning degree from the University of Pennsylvania, and he is currently a doctoral student in the Department of Urban Studies and Planning at Massachusetts Institute of Technology. His current research interest is the social aspects of mobility sharing, and how mobility sharing systems could be improved and optimized. His prior research includes various topics in different modelable aspects of cities, e.g., land-use dynamics and impact of metro rail on urban mobility.



Jinhua Zhao is the Edward H. and Joyce Linde Associate Professor of City and Transportation Planning at the Massachusetts Institute of Technology (MIT). Prof. Zhao brings behavioral science and transportation technology together to shape travel behavior, design mobility system and reform urban policies. He develops methods to sense, predict, nudge and regulate travel behavior, and designs multimodal mobility system that integrates autonomous vehicles, shared mobility and public transport. Prof. Zhao directs the Urban Mobility Lab (mobility.mit.edu) at MIT.

View publication stats