Satisfying Social Preferences in Ridesharing Services

Sandhya Saisubramanian\textsuperscript{1} Connor Basich\textsuperscript{1} Shlomo Zilberstein\textsuperscript{1} Claudia V. Goldman\textsuperscript{2}

Abstract—Dynamic ridesharing services (DRS) play a major role in improving the efficiency of urban transportation. User satisfaction in dynamic ridesharing is determined by multiple factors such as travel time, cost, and social compatibility with co-passengers. Existing DRS optimize profit by maximizing the operational value for service providers or minimizing the travel time for users but they neglect the social experience of riders, which significantly influences the total value of the service to users. We propose DROPS, a dynamic ridesharing framework that factors the riders’ social preferences in the matching process so as to improve the quality of the trips formed. The trip formation is a multi-objective optimization that aims to maximize the operational value for the service provider, while simultaneously maximizing the value of the trip for the users. The user value is estimated based on compatibility between co-passengers and the ride time. We also present a real-time matching algorithm for trip formation. Finally, we evaluate our approach empirically using real-world taxi trips data, and a population model including social preferences based on user surveys. Our approach improves the user value and users’ social compatibility, without significantly affecting the vehicle miles for the service provider and travel time for users.

I. INTRODUCTION

Dynamic ridesharing services (DRS), such as UberPool and LyftLine, are becoming an increasingly popular means of commuting, especially in large cities [1], [2]. They are characterized by matching multiple requests that arrive in real-time, for a one-way and one-time trip. We consider a dynamic ridesharing setting where a service provider operates the vehicle fleet and requests arrive in real-time. Two important factors explain the growing attractiveness of DRS for customers: (i) cost effectiveness and (ii) ease of finding a ride in large cities where it is comparatively hard to find a taxi otherwise. For a service provider, dynamic ridesharing helps serve customers with possibly fewer vehicles, thus reducing their operational cost.

A common objective for optimizing riders’ satisfaction in existing ridesharing systems is to minimize their travel time [2]–[4]. In practice, however, there are many other factors that affect user satisfaction in DRS, apart from travel time. Since a user could be traveling with a stranger in the ride, their compatibility plays a major role in the user’s satisfaction. In fact, there is growing evidence that desire for personal space and security when riding with strangers pose a major barrier to using ridesharing for many users [4], [5]. For example, a female passenger may prefer to ride only with female co-passengers. The user may have a different set of preferences depending on the time of day and the location—preferences are trip-specific and not necessarily user-specific.

Consider a scenario with three requests where $r_1$ and $r_2$ are male and $r_3$ is a female passenger. Let these requests arrive at the same time (Fig. 1), such that optimizing the operational value for the service provider forms a trip with these requests (1(a)). However, this may violate the users’ social preferences and the trip may need to be altered to satisfy the preferences. If the passengers prefer riding with co-passengers of the same gender but are indifferent to riding with co-passengers of a different gender, then it is desirable to minimize their ride time overlap in the vehicle by altering the pick up and drop off order (1(b)). When the riders prefer co-passengers of the same gender and wish to avoid traveling with co-passengers of a different gender, then it is better to form two trips (1(c)). If the service does not provide a mechanism to express such social preferences and forms trips that violate these preferences (as in 1(a)), then the customers may not use the service. Current DRS, however, do not account for social preferences in their optimization, despite the fact that some users consider them to be crucial [4]–[7].

We present DROPS (Dynamic Ridesharing Optimization using Social Preferences), a dynamic ridesharing framework that facilitates incorporating social preferences of the users in trip formation process. A weight vector over preferences indicates the importance of each factor in determining the trip value to the user. The goal is to form trips that optimize both operational value for the service provider and value of the trip to the passengers, which incentivizes the users to continue using the service, thereby benefiting the service provider. The value of a trip to a user is calculated based on their social compatibility with other co-passengers, the ride time, and ride cost. We solve this bi-objective optimization problem.

\textsuperscript{1}College of Information and Computer Sciences, University of Massachusetts Amherst, MA 01002, USA. \{saisubramanian, cbasich, shlomo\}@cs.umass.edu
\textsuperscript{2}General Motors, Technical Center Israel. claudia.goldman@gm.com

We thank Shannon Roberts and Fangda Zhang for conducting the user survey and for fruitful discussions on designing the population model.
using the scalarization technique [8], which solves a linear combination of the multiple objectives. The relative importance of each objective can be controlled using the weight vector for the objectives. Given a set of riders, candidate trips are formed using our real-time greedy algorithm that adds customers to a trip only if the newly produced trip’s value is above a certain threshold.

Two extensive user surveys were used to determine how users evaluate social ridesharing. Based on the 552 responses, we consider three basic social factors—age, gender, and user rating—along with a time preference indicating if the user is in a rush. The viability of factoring social preferences into the matching process is evaluated empirically. The experiments examine the impact of matching with social preferences (social matching) on users and the service provider. We test our approach on a real-world taxi trips dataset and compare the results with that of three baselines, each focusing on optimizing different components of the objective for trip formation. The population model and preferences used in our experiments are based on the survey results. Our results show that incorporating social preferences of users in the matching process improves the overall user value considering the social compatibility, without significantly affecting the operational cost for the service provider.

Our primary contributions are: (i) presenting DROPS, a system for dynamic ridesharing that uses social preferences (Sec. III); (ii) proposing a real-time greedy algorithm for trip formation (Sec. IV); and (iii) empirical evaluation showing the benefits of social matching using real-world taxi data and a population model based on user surveys (Sec. V).

II. RELATED WORK

Dynamic ridesharing services have gained popularity due to the cost benefits they offer to the users and service providers, apart from their contributions to sustainable environment resulting from efficient vehicle usage. DRS are characterized by user requests that arrive in real-time [7], unlike car-pooling where the requests are known a priori [1]. Optimizing DRS has been an active research area, attracting researchers from diverse fields such as operations research, transportation, and artificial intelligence [4], [9], [10], [18].

Existing literature on DRS can be classified broadly based on the objective function and the solution method employed. Optimization-based approaches are the common solution technique employed [3], [9]–[11] to optimize the route and travel time [4], [7], [10], [12], [13]. Specifically, the commonly used objectives for determining ridesharing matches are: (i) minimizing system-wide vehicle-miles; (ii) minimizing system-wide travel time; and (iii) maximizing number of participants. A critical missing component of these objectives is the in-ride user experience. Multiple surveys have acknowledged that it is essential to account for users’ social preferences to improve DRS [1], [4], [5], [7], [9], [14], [15], which is not currently handled. To address this discrepancy, we present a dynamic ridesharing framework that considers the social preferences of the users when matching them with other passengers for a trip.

III. PROBLEM FORMULATION

The DROPS framework facilitates customizing rides to improve user compatibility by incorporating the social preferences of users. Let \( \mathcal{R} \) denote the finite set of unsorted (non-dispatched) requests and \( \mathcal{V} \) denote the finite set of available vehicles, at time \( t \). Each request \( r \in \mathcal{R} \) is denoted by \( \{s, c, i, \bar{p}, \bar{w}, U\} \) and each vehicle \( v \in \mathcal{V} \) is denoted by the tuple \( \{ID, \gamma\} \). Refer Table I for the definitions of variables and constants employed in the formulation.

We consider social preferences in each request that correspond to three social factors: age, gender, and rating. Additionally, we consider a time preference to indicate if the user is in a rush. We identified these factors based on the results of our user survey, conducted specifically to determine user expectations in ridesharing services. The preferences \( \bar{p} \) are denoted as \(+1, -1, 0\), indicating the user’s desirability, undesirability, or indifference about a certain value of a factor. For example, a preference of \(+1\) for \textit{rating} \( \geq 4 \) denotes that the person prefers riding with co-passengers who have a minimum rating of 4, and a preference of \(-1\) for \textit{rating} \( \leq 3 \) denotes that the person wishes to avoid riding with co-passengers who have a rating of 3 or below. That is, if the rating on a scale of 1 to 5 is treated as a vector, then these preferences are denoted as \((-1, -1, 1, +1, +1)\). The weights \( \bar{w} = [w_1, w_2, w_3, w_4] \) correspond to the time, age, gender, and rating, respectively.

A solution to an instance of this problem is a set of trips \( \lambda \). Each trip \( \lambda \in \Lambda \) is a matching of requests to a vehicle, denoted by \( \lambda = (R, v, \tau) \) where \( \tau \) is the trip trajectory (route), and \( V(\lambda) \) denotes the value of the trip. The objective is to maximize the cumulative value of all trips dispatched in a given horizon \( H \).

\[
\max_{\tau \in H} \sum_{\lambda \in \Lambda} V(\lambda).
\]

\textit{a) Multi-objective formulation:} Since the goal is to form trips that maximize the operational value for the service provider as well as maximizing the overall user value, this is naturally a bi-objective optimization. To solve this, we employ the scalarization approach [8], [16], which projects

<table>
<thead>
<tr>
<th>Variables/ Constants</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A^t )</td>
<td>Set of trips formed at time ( t )</td>
</tr>
<tr>
<td>( V(\lambda) )</td>
<td>Value of trip ( \lambda )</td>
</tr>
<tr>
<td>( c^i_\lambda )</td>
<td>Cost of using the vehicle corresponding to ( \lambda )</td>
</tr>
<tr>
<td>( \gamma_r )</td>
<td>Maximum passenger capacity of vehicle</td>
</tr>
<tr>
<td>( ID_v )</td>
<td>Vehicle ID</td>
</tr>
<tr>
<td>( s_r, \tau_r )</td>
<td>Start (pick-up) and (drop-off) locations for ( r )</td>
</tr>
<tr>
<td>( \alpha_r )</td>
<td>User’s social utility</td>
</tr>
<tr>
<td>( x_r )</td>
<td>Amount charged to ( r ) for the trip</td>
</tr>
<tr>
<td>( d_r )</td>
<td>Discount offered to ( r ) for using ridesharing</td>
</tr>
<tr>
<td>( t_r )</td>
<td>Request initiation time</td>
</tr>
<tr>
<td>( \bar{p} )</td>
<td>Social and time preferences of ( r )</td>
</tr>
<tr>
<td>( \bar{w} )</td>
<td>User’s weights corresponding to preferences ( \bar{p} )</td>
</tr>
<tr>
<td>( U_r )</td>
<td>User demographics: {age, gender, rating}</td>
</tr>
</tbody>
</table>
a multi-objective value to a single scalar value by parameterizing the objectives using a weight vector. The weight for each objective indicates its relative importance, thus resulting in a single objective for optimization. Let \( \beta_o, \beta_u \) be the scalarization weights corresponding to operational value and user value, respectively. Then, \( \forall \lambda \), the trip value is:

\[
V(\lambda) = \beta_o \left( \sum_{r \in R_k} (x_r - d_r) - c_k^\lambda \right) + \beta_u \left( \sum_{r \in R_k} \alpha_r + d_r \right)
\]

The operational value and the user value are measured in dollars ($) and normalized to the same scale before scalarization. The operational value to the service provider depends on the cost of operating the vehicle for the trip \( (c_k^\lambda) \) and the amount paid by the riders, which is the difference between the amount charged for the trip \( (x_r) \) and the discount offered for using ridesharing \( (d_r) \). The value of the trip to a user depends on the user utility \( (\alpha_r) \) and the discount gained for using ridesharing \( (d_r) \). The user utility \( (\alpha_r) \) is the difference between the users’ social compatibility with their co-passengers and the extra travel time incurred by using ridesharing. The social compatibility for a request is calculated as the cumulative weighted difference between the preferences \( p_r^v \) and demographics of each co-passenger.

We now explain the social utility calculation using a simple example. Consider two requests \( r_1 \) (female) and \( r_2 \) (male) that arrive at the same time and have the same source and destination coordinates, same age (30), and rating (4). \( r_1 \) prefers (+1) female co-passengers with age in the range 20-40 and rating \( \geq 4 \) and prefers to avoid \((-1)\) for all other values of social factors. Let the weight of these social preferences be \( w_r^v = [0.3, 0.3, 0.2, 0.5]^T \), corresponding to time, age, gender, and rating. The social compatibility for \( r_1 \) with respect to \( r_2 \) is \( 0.3 \cdot 0.2 + 0.5 = 0.6 \). Let the extra trip time be 2 minutes, then \( \alpha_r = 0.6 - 0.3 \times 2 = 0 \). When \( |R_k| > 2 \), \( \alpha_r \) is the sum of all pairwise values.

IV. SOLUTION APPROACH

Given a set of requests and vehicles, our solution approach (Fig. 2) consists of two components: (i) trip formation and (ii) trip dispatch. In each decision cycle, the trip formation component matches requests with each other and to vehicles, and the dispatch component decides which trips are dispatched. We restrict the scope of matching in this paper to requests and vehicles that have not been dispatched. That is, we do not consider adding a request to a vehicle (trip) en-route (already driving on the road). The route planner calculates the optimal trajectory for picking up and dropping off a given set of requests.

a) Trip Formation: In this phase, requests are matched with other requests and assigned a vehicle to form a trip. The matching is performed using a greedy approach outlined in Algorithm 1. The input to the algorithm is the set of requests and a trip value threshold \( \delta \) that indicates the required minimum improvement in trip value to form trips. The algorithm adds a request to the best trip (maximum improvement) that improves the trip value at least by a factor of \( \delta \) and if the trip size has not exceeded the maximum capacity of the vehicle (Lines 7-16). Standard hyperparameter tuning or sample average approximation [17] may be used to estimate \( \delta \). The trip value is estimated using Equation 1.

Each request is assigned to the best trip that satisfies the threshold improvement. If no such trip is found, then a new trip is created with the request (Lines 19-22). This ensures that all requests have a trip associated. The route planner computes trajectories that determine the pick up and drop off order for a given set of requests. The trajectory that maximizes the trip value for a given set of requests is selected as the route \( \tau \) for the trip. During the trip formation, the best route is computed optimally and updated whenever a new request is added to a trip (Line 8, 21). The output of this algorithm is the set of all trips formed, \( \Lambda^t \).

b) Trip Dispatch: The trips are dispatched in this phase if at least one of the following conditions is satisfied: (i) trip value is above the predefined dispatch threshold; or (ii) a request in the trip has remained unserved for a certain period of time since its arrival (queue time). The dispatch

\[
Algorithm 1: \text{Greedy Matching (}\mathcal{R}^t, \delta)\\
1: \Lambda^t = \emptyset\\
2: \text{foreach}\ r \in \mathcal{R}^t\ \text{do}\\
3: \quad \text{matched} = \text{false}\\
4: \quad \text{if } |\Lambda^t| > 0\ \text{then}\\
5: \quad \quad \lambda_{\text{best}} = \emptyset, \lambda_{\text{rem}} = \emptyset, \text{Best Value} = -\infty\\
6: \quad \quad \text{foreach}\ \lambda \in \Lambda^t\ \text{with } |R_k| < \gamma, \ \text{do}\\
7: \quad \quad \quad \text{Calculate best route for } \lambda = \Lambda + r\\
8: \quad \quad \quad \text{if } V(\lambda') - V(\lambda) \geq \delta\ \text{and } V(\lambda') > \text{Best Value}\\
9: \quad \quad \quad \quad \lambda_{\text{rem}} \leftarrow \lambda, \lambda_{\text{best}} \leftarrow \lambda'\\
10: \quad \quad \quad \quad \text{Best Value} = V(\lambda_{\text{best}})\\
11: \quad \quad \quad \text{matched} = \text{true}\\
12: \quad \quad \text{end }\\
13: \quad \quad \text{if matched = true then}\\
14: \quad \quad \quad \Lambda^t \leftarrow (\Lambda^t \backslash \lambda_{\text{rem}}) \cup \lambda_{\text{best}}\\
15: \quad \quad \text{end }\\
16: \quad \text{if matched = false then}\\
17: \quad \quad \text{Create new trip } \lambda \text{ with request } r\\
18: \quad \quad \text{Calculate best route for } \lambda\\
19: \quad \quad \Lambda^t \leftarrow \Lambda^t \cup \lambda\\
20: \quad \text{end }\\
21: \text{return } \Lambda^t
\]
threshold for trip value and the queue time for the requests may be guided by the average trip value and ride time of solo rides or determined by the service provider. In our experiments, trips that satisfy the queue time threshold are given a higher priority over the trips with lower queue time but higher trip value, ensuring that requests do not remain unserved forever due to lower trip value. The trips are then dispatched based on availability of vehicles, $V_t$. At the end of decision cycle $t$, all unserved requests—requests in trips that are not dispatched—are added to the requests set for the next decision cycle, $R_{t+1}$.

V. EXPERIMENTAL RESULTS

The experiments evaluate the impact of using social preferences in ridesharing, with respect to users and the service provider. We built a realistic ridesharing simulator using the Chicago taxi trips dataset\(^1\) and a population model based on user surveys. A map of Chicago divided into zones\(^2\) is shown in Fig. 4. We compare the results obtained using social preferences in dynamic ridesharing matching (SM) with that of three baselines: $(B_1)$ maximizing only the operational value, $\beta_0 = 1, \beta_u = 0$; $(B_2)$ maximizing only user value, $\beta_0 = 0, \beta_u = 1$; and $(B_3)$ maximizing the comprehensive trip value in Equation 1 but without considering user’s social preferences corresponding to age, gender, and rating $(w_u = 0, w_g = 0, w_r = 0)$ for the trip formation. Note that $B_3$ considers the operational cost for service provider, total cost of the trip for the users, and their cost of time in the trip formation. Trips are formed using Alg. 1 for each objective.

The algorithms and the simulation system were implemented by us on an Intel Xeon 3.10 GHz computer with 16GB of RAM, using a homogeneous vehicle fleet with a seat capacity of 4. Each decision cycle is 30 seconds in real-time and the horizon $H$ is one day. In our experiments, we assume that the number of vehicles is not bounded. We set the trip threshold $\delta$ to zero for the greedy algorithm—requests are added to the best trips possible as long as the current value of the trip is not diminished. This conservative value allows us to examine the benefit of social matching uniformly across zones. However, in practice this hyper-parameter may be tuned to further optimize performance subject to the service provider’s objective. The request queue time threshold for dispatch is set to 5 minutes. The travel time and distances are calculated using straight line distances between the coordinates and a vehicle speed of 30 miles per hour. While these experiments do not account for the actual routes and traffic conditions, these factors are not likely to change the relative merits of each approach and the conclusions of the study.

A. Population Model and Dataset

We consider a population model based on the results of online surveys conducted in North America. The surveys had 552 responses, of which $\sim 60\%$ used ridesharing at least once a month. The survey results indicated that users would like to be matched with people who are similar to them. The social factors, preferences ($\vec{p}$), and the weights ($\vec{w}$) were determined based on the survey results. The demographic information such as age and gender, for our experiments, is drawn from the actual Chicago demographic distributions\(^3\). The survey also indicated that some users are unwilling to use ridesharing when social preferences are not taken into account. To reflect this, certain users were marked as reluctant for ridesharing in the absence of social matching and these users were always dispatched in solo rides, when forming trips with the baseline objectives.

The Chicago taxi trips data consists of trip-specific information such as start and end time of the taxi ride, trip fare, and the pick up and drop off coordinates along with their geographic zones. Requests from each zone are partitioned into training and testing sets. We consider requests from two consecutive weeks in April 2015, originating in zones 8, 28, and 56 for testing and requests from three prior days in these zones for training. The average number of

\(^{1}\)https://data.cityofchicago.org/Transportation/Taxi-Trips/wrvz-psew

\(^{2}\)https://en.wikipedia.org/wiki/Community_areas_in_Chicago

\(^{3}\)http://chicago.areaconnect.com/statistics.htm
requests per day in each of these zones is 20000, 7000, and 1500 respectively. The scalarization weights for Eqn. 1 were estimated empirically using the training data. Fig. 3 shows the effect of different weights on the operational and user values. The weight for operational value ($\beta_o$) is shown in x-axis and the weight corresponding to user value ($\beta_u$) in y-axis. Varying the weights alters the relative contribution of each component, resulting in different trips. The weights that achieved the best trade-off in training data were selected for the test data, $\beta_o = 0.8$ and $\beta_u = 0.6$ for zone 8 and 28, and $\beta_o = 0.5$ and $\beta_u = 0.5$ for zone 56 data.

B. Analysis of Tradeoffs

**Impact on Users:** The impact on users is measured by the total user value (Fig. 5), average social utility per minute (Fig. 6), and the increase in ride time, relative to a solo trip. Trips formed by maximizing operational value ($B_1$) have the least user value across all zones, as expected. Our approach (SM) achieves user value close to that of optimizing for user value alone ($B_2$), and sometimes better than $B_2$. This is because, in some cases, the values of the trips formed by optimizing $B_2$ may not meet the dispatch threshold in which case the trips are dispatched after 5 minutes, which eventually reduces the user value. Our approach overcomes this drawback by optimizing for both the objectives, providing greater cumulative value for a given trip and enabling it to be dispatched more quickly.

The social utility ($\alpha_r$) per minute measures the average social compatibility of users with their co-passengers. To account for the different ride times, we normalize it to utility per minute, along with standard error (Fig. 6). We observe that SM consistently performs similar to or better than $B_2$, showing that the user value is improved through better matching, and not merely based on the ride time. We also evaluated the increase in ride time, compared to solo ride. SM increases the ride time by $\sim 3$ minutes, well within the range of users’ acceptable increase in ride time (at most 5 minutes), inferred from the survey results, and comparable to that of baselines. Due to space constraints, we do not plot these results. The social compatibility typically offsets the increase in ride time for the users, thus resulting in increased user utility when forming trips using our approach.

**Impact on the Service Provider:** The impact on service provider is determined based on the operational value and the total miles driven, to give a sense of degree of variation, induced by social matching, on the trip routes and quality of service. As expected, objective $B_1$, which optimizes the operational value only, achieves the highest operational value and objective $B_2$, which maximizes user value, has the lowest operational value (Fig. 7). The operational value achieved by our approach (SM) is close to that of $B_1$, with a slightly higher miles driven (Fig. 8) and higher user utility. The total number of trips formed by our approach is comparable to that of $B_1$. This shows that our approach improves the quality of trips without significantly affecting the total miles or the operational cost.

C. Runtime and Robustness

Since matching is performed every 30 seconds, it is important to ensure that the matching algorithm quickly forms trips so that it may be effectively used in real-time. The run time (in seconds) of our matching algorithm is on average $0.5$ in the zone with high request density (zone 8), $0.12$ in zone 28, and $0.003$ in zone 56.

We also compared our matching algorithm to a hindsight greedy matching with access to all the requests, including future ones. This helps evaluate the potential gain in operational value and user value, when knowledge of future requests is available. Our approach achieved at least $\sim 89\%$
of the operational value and up to $\sim 84\%$ of the user value compared to the hindsight matching in all zones. This indicates that any prediction method of future requests would yield very limited gains in the operational value, but some improvements in user value could be achieved by forming trips where the passengers have a higher social compatibility.

VI. CONCLUSION

Dynamic ridesharing is an increasingly appealing commuter option. However, numerous surveys have indicated that riders’ concerns, primarily about the social characteristics of co-passengers, pose a major barrier to using ridesharing for a segment of the population. We present the DROPS system for optimizing dynamic ridesharing with social preferences and present an efficient real-time matching algorithm that can handle effectively high density zones. Our results demonstrate that factoring social preferences into the matching process helps improve the user value, without significantly affecting the operational value to the service provider.

In the future, we aim to examine ways to extend the matching model to consider trips that are currently en-route. We also plan to employ a predictive model for future requests to improve the user value. While we anticipate some performance gains in that case, we do not expect the relative benefits of social matching to diminish.

REFERENCES