Employee Ridesharing: Reinforcement Learning and Choice Modeling

Completed Research Full Papers

Wangcheng Yan
University of Tennessee Knoxville
wyan3@vols.utk.edu

Wenjun Zhou
University of Tennessee Knoxville
wzhou4@utk.edu

Chang Tan
iFLYTEK Bigdata Research
changtan2@iflytek.com

Lei Fan
iFLYTEK Bigdata Research
leifan@iflytek.com

Abstract

Many governments and companies offer incentive programs (e.g., allowance) to encourage employees to share rides for their daily work commute because ridesharing helps reduce travel costs, road congestion, and greenhouse gas emissions. For the program success, it is critical to understand employees’ motivation to use this service and effectively incentivize new users’ adoption and retain existing users by providing the best ridesharing experience. In this study, our goal is to learn user incentives and preferences as well as risks in employee ridesharing. We use a choice model to estimate the personalized utility of different choices among possible transportation modes and learn from longitudinal behaviors using reinforcement learning. The integrative model is derived theoretically and evaluated empirically using a real-world employee ridesharing data. We find that ridesharing is risky and company employees, especially drivers, adopt the service mainly because of social relationship with colleagues rather than the financial incentive.

Keywords
Incentive, choice modeling, user preference, risk analysis, driver-passenger matching

Introduction

Ridesharing, a form of multiple people using the same vehicle to arrive at a similar destination, has gained much attention in recent years (Saranow 2006). Ridesharing has become popular since it may help reduce travel costs, road congestion, and greenhouse gas emissions (Furuhata et al. 2013). Besides the well-known commercial ridesharing such as Uber and Lyft, employee ridesharing has emerged to help employees share rides for their daily work commute (Ferguson 1991) and has sparked much research and practice over several decades (Ferguson 1990).

Existing studies have reported motivating factors using survey responses (e.g., Ferguson 1990; Agatz et al. 2012). It was found that time-saving incentives and monetary incentives are critical for retaining users (Chan and Shaheen 2012), whereas the trust issue and effective matching risk are the major concerns for users to adopt the ridesharing (Agatz et al. 2011). These studies provide preliminary understandings of ridesharing incentives and users’ attitudes, but there is a lack of data-driven analysis towards the synthetic effect of various incentives.

In this study, we take the context of a company that has recently launched an employee ridesharing app. The employee ridesharing is different from commercial ridesharing in many aspects. Most importantly, rider offerers in the commercial ridesharing are typically incentivized by money, whereas those in the employee ride-sharing can be motivation by keeping a good social relationship with colleagues and they only receive a small reward as compared to commercial ride offerers. Moreover, though both offer a financial incentive to users, the latter requires passengers to meet with the driver in time and will charge them if late, whereas the former would not do so. Further, the origination and destination of a user in the employee ridesharing
are typically fixed, that is, they usually commute between home and company.

Our contributions can be summarized as follows. Our study is a pioneer work integrating the reinforcement learning and choice modeling in estimating user incentives and risk associated with each choice in the ridesharing. Our findings demonstrate that financial and time-saving incentives are two crucial factors for user adoption, which is aligned with existing literature. However, the social relationship with colleagues seems to be the primary incentive. Our results also confirm ridesharing involves more risk than public transit or solo driving and show that the extra driving distance is the primary concern for drivers. In a nutshell, our utility estimation approach could serve as a support for the ridesharing passenger-driver matching system.

Literature Review

Many studies have explored what factors play a role in the ridesharing adoption, in particular using survey data (e.g., Ferguson 1990). A critical incentive for the user adoption of ridesharing is reducing car-related expenses (Agatz et al. 2012). Other incentives include limited parking infrastructure in the company (Chan and Shaheen 2012) and the power of social aspects (Margolin et al. 1978). Despite its attractiveness to users, ridesharing is also faced with many barriers, such as regulation and criminal charges (Ballús-Armet et al. 2014). Further, users may be concerned about the flexibility and convenience of the ridesharing as compared to a private automobile (Stiglic et al. 2016). As stated in Agatz et al. 2011, users may be discouraged if they cannot find matches and then stop using the service. Such inconvenience comes from the risk of ridesharing; that is, users’ orders cannot be fulfilled. However, to the best of our knowledge, all the works above revealing user incentives are survey-based, and there is a lack of data-driven approaches to analyze the importance of each incentive quantitatively. Our work applies reinforcement learning and choice modeling to the employee ridesharing to measure users’ incentives quantitatively.

Reinforcement learning is a machine learning technique to model the behavior of a single agent in a stochastic environment. A basic reinforcement learning is the Markov Decision Process (MDP), of which the next state of an agent depends on his current state and action only. With the popularity and success of using reinforcement learning in computer science research and applications (e.g., AlphaGo), applying this technique in information systems (IS) and marketing research has attained more attention over the years. For example, Greenwald et al. 2010 adopted an MDP to evaluate the effect of different information revelation policy, including complete and incomplete ones, on the expected price paid by the procurer in e-marketplaces.

Choice modeling is a statistical tool that estimates users’ probability of choosing an alternative given the deterministic utility gain and uncertainty. For instance, Morency et al. 2012 proposed a two-stage model where the first stage is binary probit whether he is active, and then a random utility-based model for usage. This approach is useful for understanding and predicting user persistence, but did not consider the choice between being a passenger or driver. When deciding whether becoming a driver, users are faced with uncertainty, while has been widely studied in IS research. For example, Schwartz et al. 2017 studied the uncertainty in targeted market action’s effectiveness when a firm tries to acquire customers via display advertising. In this work, we aimed at applying the choice modeling in the ridesharing, where different choices may incur variant risk.

Problem Description

An anonymous company has adopted a ride-sharing app that allows employees to request and provide ridesharing. As demonstrated in Figure 1(a), after logging in the app, users are allowed to specify their roles as a passenger or driver by clicking on the corresponding tab. When they choose to be a passenger, they have to enter the origin, destination, and departure time. Upon submission of the request, the app will notify available drivers, and the requester has to wait for a driver to pick up the order (see Figure 1(b)). Meanwhile, users who elect to be drivers see a list of passengers seeking rides, and determine whether to offer a ride and to who request (see Figure 1(c)).

As we can see, in this system, passengers are not allowed to choose drivers (but are allowed to cancel an order), whereas drivers are allowed to select passengers. The selection process is manual as the app only sorts the requests simply by the distance of destinations. Both drivers and passengers are allowed to contact
each other for settling the transaction. Once the order is confirmed, they will meet and take the trip together. GPS traces with temporal-spatial information will be collected to confirm the occurrence of trips. Employees will only be subsidized for trips originating from or arriving at the company.

Given the past passenger-driver matching records, we would like to learn user incentives and preferences in employee ridesharing, which may serve as a basis for the passenger-driver matching system. Since the data have sequential choices and the current choice depends on the experience, we consider this problem in a reinforcement learning and choice modeling framework.

Formally, suppose that there are \( n \) users, each of which may choose a state from several alternatives (i.e., passenger, driver, and other) at each time point (morning or afternoon in a day). Anyone can only be in one state at a time point. We assume that preferences over these alternatives evolve based on the recent ride-share experience.

Users could also choose different actions. For the drivers, their actions could be submitting an order as a driver, submitting an order as a passenger, and no order submission in the app. For the passengers, their actions could be submitting an order as a passenger and no order submission.

For each user, they could obtain a reward being in a state. Since they cannot anticipate the reward before they are actually in the state, they will learn the reward with their tries. In this work, we will also use another concept - utility. The reward is a kind of utility. There are also other kinds of utility, such as the expected total reward by taking action \( a \) in a state \( s \). Meanwhile, the utility has some stochasticity due to the uncertainty of results from each action. In sum, the reward depends on the state and action, based on which we define the following terms.

**Definition 1 (Trajectory)** A trajectory \( \tau \) is a sequence of state-action pairs over time \( t \). We use \( s \) to represent state and \( a \) to represent action. Mathematically,

\[
\tau = \{(s_1, a_1), (s_2, a_2), \ldots, (s_T, a_T)\}.
\]

**Definition 2 (Feature)** A state-dependent feature set \( \Phi(s) \) is composed of \( J \) features,

\[
\{\phi_1(s), \phi_2(s), \ldots, \phi_J(s)\}.
\]

**A Utility Maximization Framework**

In this section, we learn users’ personalized incentives and concerns in the ridesharing via a utility maximization framework. We first provide some preliminaries about choice modeling and reinforcement learning. Then we define the personalized utility in our ride-sharing system and propose a new model named HEV-IRL to estimate the reward function.
Travel Means Choice Modeling

For a user, there are several possible actions given a state. For instance, a user with state $s_1$ can choose $a_1$ so that the next state becomes $s_2$. To study the preferences of each alternative action, choice modeling with a multinomial logit specification has been commonly used. As the state is given, we leave it out for simplicity in exposition. When an individual is making a choice among alternatives, the utility of alternative $i$ is in the form of

$$U_i = Z_i + \epsilon_i,$$

where $Z_i$ is the deterministic component that helps define utility (e.g., a linear combination of input features), and $\epsilon_i$ is the stochastic term. $U_i$ can be understood as the satisfaction level from choosing alternative $i$ and ranges from $-\infty$ to $+\infty$. A key assumption in the multinomial logit model is that the random part of the utility function is independent and identically distributed (i.i.d.). This assumption is not always valid due to the fact that different alternatives may have different variations. For instance, for a passenger, choosing public transit has a small variation as one can always take a bus. However, choosing to take ride may carry the risk of no pick, which leads to a considerable variation.

The heteroskedastic extreme value (HEV) model is an extension of the multinomial logit model to solve this problem. In this model, $\epsilon_i$ follows a Gumbel distribution with location parameter zero and scale parameter $\theta_i$ (Details can be found in R.Bhat 1995). Given that the stochastic terms $\epsilon_i$ are independently distributed, the probability that an individual chooses alternative $i$, $P_i$, can be calculated from $U_i$ as follows,

$$P_i = \mathbb{P}(U_i > U_j, \forall j) = \int_{\epsilon_i = -\infty}^{+\infty} \prod_{j \neq i} \Lambda \left( \frac{Z_i - Z_j + \epsilon_i}{\theta_j} \right) \frac{1}{\theta_i} \lambda \left( \frac{\epsilon_i}{\theta_i} \right) d\epsilon_i,$$

where $\lambda(x)$ and $\Lambda(x)$ are probability density function (PDF) and cumulative distribution function (CDF) of the standard Gumbel distribution, respectively. In particular,

$$\lambda(x) = e^{-x}e^{-e^{-x}}, \text{ and } \Lambda(x) = e^{-e^{-x}}.$$

Reinforcement Learning

Reinforcement learning is often used for adaptively learn and maximize an agent’s utility over time. It is commonly studied with the assumption of a Markov Decision Process (MDP). A MDP is a process consisting of parameters $\{S, A, P, \gamma, R\}$, where $S$ is the state set, $A$ is the action set, $P$ is the transition probability from state $s$ to state $s'$ via action $a$, $\gamma$ is a discount factor with $0 < \gamma < 1$, and $R$ the reward function.
A policy $\pi$ is defined as a mapping from $S$ to $A$: $\pi(s,a) = P(A = a|S = s)$. We also have the long-term reward $G_t = R(S_{t+1}) + \gamma R(S_{t+2}) + \cdots + \gamma ^{k-1} R_{s_{t+k}} + \cdots = \sum_{k=1}^{\infty} \gamma ^{k-1} R_{s_{t+k}}$. Then the objective becomes finding an optimal policy $\pi^*$ that maximizes the long-term reward. A typical solution is introducing the value function $V$ and the Q-function, which are defined as follows,

$$V^\pi(s) = E_{\pi}[G_t | S_t = s], \quad Q^\pi(s,a) = E_{\pi}[G_t | S_t = s, A_t = a].$$

The definitions of $V$ and $Q$ are standard in reinforcement learning, which could solve the model-free reinforcement learning (Sutton and Barto 1998). These two value functions satisfy the Bellman equations,

$$V^\pi(s) = R(s) + \gamma \sum_{s'} P(s, \pi(s), s') V^\pi(s')$$
$$Q^\pi(s,a) = R(s) + \gamma \sum_{s'} P(s, a, s') V^\pi(s')$$

In the MDP process, the reward function should be given for each state $s$. However, in some cases, the reward function is not observable, but the sequence of state-action pairs is available. Inverse reinforcement learning (IRL) is an approach to find the reward function given the observed trajectory $\tau$. Following Ng and Russell 2000 and Babes-Vroman et al. 2011, we assume that the reward function is a linear combination of features,

$$R(s) = w^T \Phi(s) = \sum_{j=1}^{J} w_j \phi_j(s),$$

where $\phi_j(s)$ is the feature value related with state $s$ and the objective of IRL becomes to find the optimal weights $w$ such that the trajectory is close to the optimal policy.

Several approaches are available to solve the IRL problem of determining weight $w$. For instance, Babes-Vroman et al. 2011 derived the log likelihood function of the probability of observed trajectory and used maximum likelihood estimator. A max margin method was developed in Abbeel and Ng 2004 and a max entropy method was proposed by Ziebart et al. 2008.

**Personalized Utility Learning**

In reality, taking ride has extra stochasticity compared with public transit as the passenger needs to take the risk of canceling order and satisfy the time plan of driver. Drivers are also faced with a similar situation. Considering that different transit means have different stochasticity, we decide to use the HEV model. Specifically, we assume that for each individual $q$, the utility of taking action $a$ given state $s$ is,

$$U^q(s,a) = Q^q(s,a) + \epsilon^q(a)$$

where $U^q(s,a)$ is the actual utility, $Q^q(s,a)$ is the Q-function value evaluated at state $s$ and action $a$, and $\epsilon^q(a)$ is a action/choice dependent random term. Note that $Q^q(s,a)$ corresponds to the deterministic term $Z_i$ as specified in Equation 3.

Since different alternatives (driver or passenger or other means) may have different stochasticity, we cannot just use Boltzmann exploration to estimate $\pi(s,a)$. Instead, we need to combine HEV to IRL and estimate $w^q$ and $\theta^q$ for each individual $q$. Therefore, we propose a heteroskedastic extreme value - inverse reinforcement learning (HEV-IRL) model to maximize the probability of the observed trajectory with the defined utility function as specified in Equation 10.

**Definition 3 (HEV-IRL Model)** For each individual $q$, given feature set $\Phi(s)$ and trajectory $\tau$, HEV-IRL estimates the weight $w$ and scale parameter $\theta$. 

---

Twenty-fifth Americas Conference on Information Systems, Cancun, 2019 5
We describe an approach to find the maximum likelihood estimator for the HEV-IRL model. Given the trajectory $\tau$, the log-likelihood can be written as,

$$L = \sum_{(s,a) \in \tau} \log(\pi(s,a)).$$  

(11)

In the following part, index $i, j, k$ are used to represent different actions $a$. According to the HEV model, by substituting $u = e^{-\epsilon_i/\theta_i}$ in Equation 5, the probability of choosing action $i$ is,

$$\pi(s, i) = \int_{u=0}^{u=+\infty} G_i(u) \cdot e^{-u} du,$$  

(12)

where

$$G_i(u) = \prod_{j \neq i} \Lambda \left[ \frac{Q(s, i) - Q(s, j) - \theta_i \ln(u)}{\theta_j} \right].$$  

(13)

After the transformation, the integration can be estimated using the Gauss-Laguerre quadrature (Abramowitz and Stegun 1965; R.Bhat 1995).

To find the maximal likelihood estimators of the parameter set $\eta = (w, \theta)$ using the gradient descent (Good-fellow et al. 2016). The major step is to compute the gradients. For $w$,

$$\frac{\partial L}{\partial w_n} = \sum_{(s,i) \in \tau} \frac{1}{\pi(s,i)} \frac{\partial \pi(s,i)}{\partial w_n} = \sum_{(s,i) \in \tau} \frac{1}{\pi(s,i)} \int_{u=0}^{u=+\infty} \frac{\partial G_i(u)}{\partial w_n} \cdot e^{-u} du.$$

(14)

Note that $\pi(s, i)$ can be calculated using Equation 12, and $\frac{\partial \pi(s,i)}{\partial w_n}$ can be calculated by changing the order of integration and partial derivative of $G_i(u)$ \left( $\frac{\partial G_i(u)}{\partial w_n}$ \right). Given the initial values of $V(s), Q(s,a)$, and $\frac{\partial V(s)}{\partial w_n}$, we can update $\frac{\partial Q(s,i)}{\partial w_n}, \frac{\partial V(s')}{\partial w_n}, \text{ and } \frac{\partial \pi(s,i)}{\partial w_n}$ recursively. For $\theta$,

$$\frac{\partial L}{\partial \theta_j} = \sum_{(s,i) \in \tau} \frac{1}{\pi(s,i)} \frac{\partial \pi(s,i)}{\partial \theta_j}.$$

(15)

Similarly, We need to calculate $\frac{\partial G_i(u)}{\partial \theta_j}$, which is then integrated to obtain $\frac{\partial L}{\partial \theta_j}$.

The updating process is summarized in Algorithm 1. In the MDP, states $S$, actions $A$, and $\gamma$ are given, whereas the transition probability matrix $P$ is estimated using historical data.

---

**Algorithm 1:** Estimate the parameter set $\eta^q$ for each individual $q$.

**Input:** $MDP(S, A, \gamma, P)$, trajectory $\tau$, features $\Phi(s)$

**Output:** $\eta^q = (w^q, \theta^q)$

1. **Initialize** $w$ and $\theta$ to get $Q(s,a) = V(s) = w^T \Phi(s)$ and $\frac{\partial V}{\partial w} = \Phi(s)$;
2. **repeat**
3. Calculate $Q(s,a)$ using Equation 8;
4. Calculate $\pi(s,a)$ using Equation 12;
5. Calculate $\frac{\partial L}{\partial w}$ using Equation 14 and $\frac{\partial L}{\partial \theta}$ using Equation 15;
6. Update $\eta = \eta + \alpha \frac{\partial L}{\partial \eta}$; If $\eta_k < 0$, assign $\eta_k = \epsilon$;
7. Calculate L using Equation 11;
8. **until** $L$ converges or maximum iterations are reached;
Results

We performed an empirical study using real-world data. In this section, we first introduce the dataset and extract some features. Then we will compare the performance of our proposed method with two benchmarks. Finally, we provide some insights into users’ behaviors.

Data Description

The dataset used in this study was provided by an anonymous company that adopted the ridesharing app from September 30th to December 24th in 2017. Each record contained the user ID, route ID, start place, end place, scheduled start time, and actual meeting time. After data cleaning, there were 866 drivers and 2,304 passengers, among whom 31,653 ridesharing records occurred. All passengers have taken 12.4 rides on average (median=4), and all drivers have offered 22.9 rides on average (median=11). Among the 3170 unique users, 649 (20.5%) only served as a driver and 2304 (72.7%) only served as a passenger. The left users (n=217) elected to be drivers and passengers at different times. Among these 217 users, most of them are more likely to be a driver than a passenger.

We also have each individual’s demographic information, including gender, age, and level in the company. Such information is vital in determining user’s choice of being a driver or passenger. In the passenger group, 56.2% are males, the mean age is 26.7, and 6.8% are in the high level. In the driver group, 77.1% are males, the mean age is 29.2, and 28.3% are in the high level.

Features

In this subsection, we summarize the features of passengers and drivers. These features are the incentives or concerns users take into consideration when they evolve in the ridesharing. In particular, a feature is assigned a positive value if it serves as an incentive and a negative if it is a concern. For instance, the financial incentive will be given a positive value, while the extra driving time will be assigned a negative value.

- **TimeAdv & TimeLate** (a cost for both drivers and passengers) These two features measure the difference between the scheduled start time and the actual meeting time. If the actual meeting time is earlier than the scheduled start time, the user suffers from “TimeAdv” as people would usually prefer a start time right after their planned work-finishing time. Similarly, “TimeLate” also generate utility cost but it happens if a user starts the trip later than scheduled. Note all the time-related features used in this work are measured in minutes.

- **TimeDiff** (cost for drivers but benefit for passengers) For drivers, this feature measures the time cost as they need to drive via a transit point during the trip when taking a passenger. Suppose the trip is from A to B via C, the “TimeDiff” is calculated using the time(A,C) plus time(C,B) minus time(A,B). The time between two places is also calculated using Baidu Map API. Passengers, however, can save time since the driving time is typically much less than the public transit time.

- **DistDiff** (cost for drivers only) The “DistDiff,” extra driving distance, is also an important cost as drivers need to pay more gas fee and car maintenance fee. Similar to the calculation of “TimeDiff”, we use Baidu Map API but replace time with distance. Passengers, however, do not have this feature as they do not pay these fees.

- **Financial Incentive** (benefit for both drivers and passengers) This feature represents the monetary incentive for all the users. Passengers could save public transit ticket fees by taking a ride, whereas drivers could get the allowance for each passenger they take.

- **Relation** Another common incentive for the rideshare is that each user would like to get to know new people and maintain casual interactions with ride sharers, especially colleagues. There are two types of “Relation”, namely “New Relation” and “Existing Relation.” The “New Relation” is measured using the number of users on a trip who have not shared rides before. To measure “Existing Relation” in a trip, we first detect the user pairs which already exist in previous trips, noted by $U$. For each pair $k$, we count the number of occurrence in previous trips, noted by $c_k$. Then the “Existing Relation” is $\sum_{k \in U} \log\left(\frac{c_{k+1}}{c_k}\right)$. The rationale is that the marginal utility of taking the ride together is decreasing with the number of occurrences.
Experimental Results and Analysis

When performing the experiment, we split the state-action pair sequence of each individual by 70% (training data) and 30% (test data). We assume each individual’s state-action pair sequence begins with the first day starting using the ride-sharing app as he or she may not know the app well before. Since each user may have different lengths of sequence, we evaluate the model performance using the average log-likelihood $L_{ave}$ computed in the test data. We further compare the performance of our proposed HEV-IRL model with two benchmarks. One is the MLIRL model proposed by Babes-Vroman et al. 2011. Another is a native model (NM), which computes the $L_{ave}$ using the $\pi(s,a)$ estimated from the training data.

Table 2 summarizes the evaluation metric $L_{ave}$ using our proposed model and the two benchmarks. Outliers were excluded before the mean was calculated (same for the remaining analysis in the section). Then we performed the paired two-sample t-test. The results show that our proposed model (HEV-IRL) perform better than the two benchmarks. That is, our proposed approach can best predict users’ future state-action pairs among the three methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Passenger</th>
<th>Driver</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n  mean</td>
<td>P-value</td>
</tr>
<tr>
<td>HEV-IRL ($L_{ave}$)</td>
<td>2120 -0.397</td>
<td></td>
</tr>
<tr>
<td>MLIRL ($L_{ave}^{(1)}$)</td>
<td>2120 -0.409</td>
<td>(&lt; 0.001)***</td>
</tr>
<tr>
<td>NM ($L_{ave}^{(2)}$)</td>
<td>2120 -0.433</td>
<td>(&lt; 0.001)***</td>
</tr>
</tbody>
</table>

Table 2. Comparison of $L_{ave}$ using the three methods.

Next, we compare the risk of users’ different choices. A larger $\theta$ indicates a larger variance of the utility, i.e., risk, since the variance of Gumbel distribution is $\frac{1}{\theta^2}$. Figure 2 shows the distributions of $\frac{\theta_p}{\theta_o}$ in the passenger group and $\frac{\theta_d}{\theta_o}$ and $\frac{\theta_p}{\theta_o}$ in the driver group. As we can see, most users has a larger scale parameter choosing to submit an order of driver or passenger than choosing other transit means. This result is aligned with existing literature that the ridesharing has more risk than solo driving or public transit. Interestingly, submitting an order as a passenger brings more risk than submitting an order as a driver (the median in Figure 2(a) is larger than that in Figure 2(b)). This may due to the fact that drivers have the right to select passengers while passengers can only wait for the selection.

![Figure 2. The distribution of scale parameters. The vertical line has x-axis value of 1.](image)

**Feature Analysis**

We compare the weights of different features for drivers and passengers. The weights represent the strength of incentives or concerns. As shown in Figure 3, the strongest incentives of passengers are to save money and keep good relation with either new or existing colleagues. The features “TimeAdv” and “TimeLate” share almost equal weights. For the drivers, the strongest incentive is to keep a good relationship with colleagues rather than earning the allowance, whereas the primary concern is the extra distance, which will cost their extra vehicle maintenance fees. Interestingly, building new relationships is more important than keeping existing relationships for both drivers and passengers (P-value < 0.001 using the paired two-sample t-test).
With the weights of different features, we performed persona profiling based on user segmentation. We used k-means clustering with three clusters. The three clusters of drivers have sample size 187, 284, and 103. Table 3 shows the cluster means of weights. We find that cluster 2 and 3 have more weights on the financial incentive, relation, time difference, and distance difference than cluster 1. We also compared the demographics, including gender, age, and level in the company with different clusters. Age is found to be significantly lower in cluster 2 and 3 than cluster 1 (The age means in cluster 1, 2, 3 are 30.6, 29.0, and 29.1, respectively, P-value < 0.001 in the paired two-sample t-test). This result is reasonable as younger drivers usually have more financial constraints and care more about the income and cost.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>A</th>
<th>NR</th>
<th>ER</th>
<th>TimeDiff</th>
<th>DistDiff</th>
<th>TimeAdv</th>
<th>TimeLate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.07</td>
<td>0.77</td>
<td>0.08</td>
<td>-0.24</td>
<td>-0.49</td>
<td>-0.46</td>
<td>-0.53</td>
</tr>
<tr>
<td>2</td>
<td>0.13</td>
<td>1.48</td>
<td>1.90</td>
<td>-0.36</td>
<td>-0.92</td>
<td>-0.37</td>
<td>-0.40</td>
</tr>
<tr>
<td>3</td>
<td>0.19</td>
<td>3.97</td>
<td>2.07</td>
<td>-0.46</td>
<td>-1.30</td>
<td>-0.30</td>
<td>-0.38</td>
</tr>
</tbody>
</table>

A = Allowance; NR = New Relation; ER = Existing Relation.

Table 3. Persona analysis of drivers: Each column represents the importance of a feature.

Conclusions and Discussions

Ridesharing has become much popular as it is environmental and financial friendly. However, it is also faced with challenges such as the inefficient matching system. To improve the ridesharing system, we need to understand users’ motivation when they use the service. In this work, we quantitatively study the user incentives in the employ ridesharing using an integrated utility maximization framework. The framework consists of the transit mode choice modeling and personalized reinforcement learning. The empirical evaluation based on real-world data demonstrates that our framework performs well and help explain user behaviors and associated risk in the ridesharing. In particular, though monetary incentive is critical, social relationship plays as the strongest incentives for user adoption, especially for drivers. Further, we find that users’ preferences over different incentives vary among different age groups.

The findings in our paper have several managerial implications. Most importantly, the company can create a ranking metric based on the learned user preferences over different incentives to maximize their overall utility. For instance, the ridesharing system can recommend a fresh passenger to a driver with a strong incentive to build new relationships with colleagues. Moreover, since the extra driving distance is the primary concern for drivers, dynamic pricing may be useful to mitigate this concern, especially for those with a strong financial incentive. Further, the model can also be adopted in broad ridesharing applications, including commercial systems like Uber and Lyft. The primary step is to extract application-specific features.

Our work is not without limitations. First, a unique setting of our study is that it is primarily used by em-
ployees of the same company. The off-line social connectivity may play a role in users’ motivation, but we are not able to incorporate this factor due to lack of data. Second, the sample size of a company’s employee ridesharing data is typically small as compared to that of commercial ridesharing, making a robust estimation of model parameters critical. A Bayesian approach that sets priors and updates the parameters with longitudinal data would be helpful. We hope to address these limitations in our future work.

References


