DHPA: Dynamic Human Preference Analytics Framework: A Case Study on Taxi Drivers' Learning Curve Analysis

MENGHAI PAN, WEIXIAO HUANG, and YANHUA LI, Worcester Polytechnic Institute
XUN ZHOU, University of Iowa
ZHENMING LIU, College of William & Mary
RUI SONG, North Carolina State University
HUI LU and ZHIHONG TIAN, Guangzhou University
JUN LUO, Lenovo Group Limited

Many real-world human behaviors can be modeled and characterized as sequential decision-making processes, such as a taxi driver's choices of working regions and times. Each driver possesses unique preferences on the sequential choices over time and improves the driver's working efficiency. Understanding the dynamics of such preferences helps accelerate the learning process of taxi drivers. Prior works on taxi operation management mostly focus on finding optimal driving strategies or routes, lacking in-depth analysis on what the drivers learned during the process and how they affect the performance of the driver. In this work, we make the first attempt to establish Dynamic Human Preference Analytics. We inversely learn the taxi drivers' preferences from data and characterize the dynamics of such preferences over time. We extract two types of features (i.e., profile features and habit features) to model the decision space of drivers. Then through inverse reinforcement learning, we learn the preferences of drivers with respect to these features. The results illustrate that self-improving drivers tend to keep adjusting their preferences to habit features to increase their earning efficiency while keeping the preferences to profile features invariant. However, experienced drivers have stable preferences over time. The exploring drivers tend to randomly adjust the preferences over time.

CCS Concepts: • **Computing methodologies → Spatial and physical reasoning**;

Additional Key Words and Phrases: Urban computing, inverse reinforcement learning, preference dynamics

ACM Reference format:
Menghai Pan, Weixiao Huang, Yanhua Li, Xun Zhou, Zhenming Liu, Rui Song, Hui Lu, Zhihong Tian, and Jun Luo. 2020. DHPA: Dynamic Human Preference Analytics Framework: A Case Study on Taxi Drivers' Learning Curve Analysis. *ACM Trans. Intell. Syst. Technol.* 11, 1, Article 8 (January 2020), 19 pages.
https://doi.org/10.1145/3360312

A preliminary version of this work was presented at the SIAM Conference on Data Mining in 2019 (SDM '19) [16] and won the best applied data science paper award. The project source code and dataset are available on the project website [2]. Y. Li, M. Pan, and W. Huang were supported in part by NSF grants CNS-1657350 and CMMI-1831140, and a research grant from DiDi Chuxing Inc. X. Zhou was partially supported by NSF grant IIS-1566386. Z. Liu was supported by NSF grant IIS-1755769. R. Song's research was partially supported by NSF grant DMS-1555244.
Authors' addresses: M. Pan, W. Huang, and Y. Li, Worcester Polytechnic Institute; emails: {mpan, whuang2, yli15}@wpi.edu; X. Zhou, University of Iowa; email: xun-zhou@uiowa.edu; Z. Liu, College of William & Mary; email: zliu@cs.wm.edu; R. Song, North Carolina State University; email: rsong@ncsu.edu; H. Lu (corresponding author) and Z. Tian, Guangzhou University; emails: {luhui, tianzhihong}@gzhu.edu.cn; J. Luo, Lenovo Group Limited; email: jluo1@lenovo.com.
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2020 Association for Computing Machinery.
2157-6904/2020/01-ART8 $15.00
https://doi.org/10.1145/3360312

1 INTRODUCTION

Taxi service is a vital part of transportation systems in large cities. Improving taxi operation efficiency is a crucial urban management problem, as it helps improve the transportation efficiency of the city and at the same time improves the income of taxi drivers. In the same city, taxi operation efficiency might differ significantly. Figure 1(a) shows the earning efficiency (total amount earned normalized by total working time) of different taxi drivers in Shenzhen, China. The top drivers earn three to four times more money than the bottom drivers.

A major cause of such difference is the difference in working experiences. Figure 1(b) shows the growth of earning efficiency of new drivers over years.[1] From March 2014 to December 2016, the new drivers became more experienced and had much higher earning efficiency. During the same time as shown in Figure 1(c), there is no obvious change to the local economy or market, as the average earning efficiency of all drivers is pretty stable. This shows that drivers are trying to improve their own strategies of looking for passengers based on their increasing knowledge of the city.

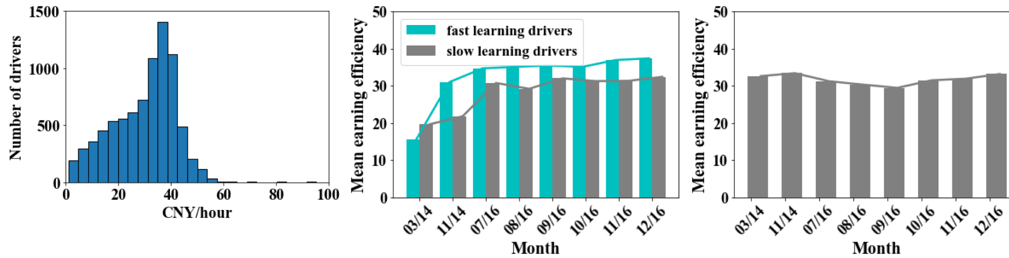
However, each driver might have gained different knowledge during the learning process, which in turn developed different preferences when making decisions. For instance, some drivers tend to look for passengers around regions near their homes, whereas others might prefer to take passengers from city hubs (e.g., train station, airport). These preferences might be unique to individual drivers and ultimately lead to differences in earning efficiency. Figure 1(b) shows that the "smart" drivers (in blue) improve their earning efficiency faster than "average" drivers and reach a higher level of earning efficiency eventually. Finding what adaptation strategies these smart drivers carry could help us understand the learning process of successful drivers and therefore help new drivers become more successful.

The passenger-seeking behavior of taxi drivers can be modeled as a Markov decision process (MDP). Prior works on taxi operation management focused on recommending the optimal policy or routes to maximize the chance of finding passengers or making profit [14, 18, 24, 25]. However, these works only studied how to find the "best" strategies based on data rather than fundamentally understanding how the drivers learned these strategies over time.

In this work, we make the first attempt to establish Dynamic Human Preference Analytics (DHPA). We inversely learn the taxi drivers' decision-making preferences, which lead to their choices while looking for passengers. We also study how these preferences evolve over time and how they help improve earning efficiency. The results shed lights on "how" the successful drivers became successful and suggests "smarter" actionable strategies to improve taxi drivers' performances. Our main contributions are as follows:

(1) We are the first to employ inverse reinforcement learning (IRL) to infer the taxi drivers' preferences based on an MDP model.
(2) We extract various kinds of interpretable features to represent the potential factors that affect the decisions of taxi drivers.
(3) We infer and analyze the preference dynamics of three groups of taxi drivers: self-improving drivers, stabilized drivers, and exploring drivers.
(4) We analyze the preference trend of different groups of taxi drivers.
(5) We conduct experiments with taxi trajectories from more than 17k drivers over different time spans. The results verify that each driver has unique preferences to various profile and habit features. The preferences to profile features tend to be stable over time, and the preferences on habit features change over time, which leads to higher earning efficiency.

[1]The dataset we have contains the records from March and November of 2014 and July through December of 2016.



(a) The distribution of earning efficiency in July 2016

(b) The mean earning efficiency of new drivers over months

(c) The mean earning efficiency of all drivers over months

Fig. 1. Dynamics of taxi drivers' earning efficiency.

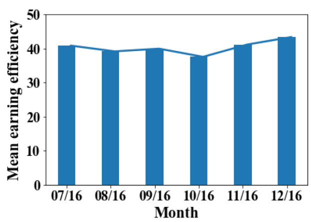
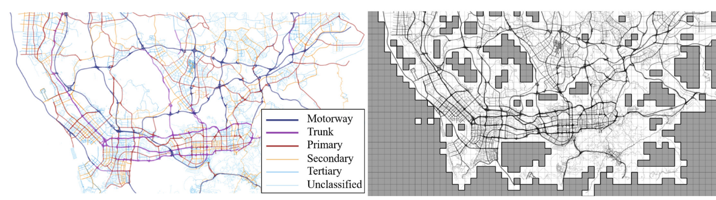


Fig. 2. Mean earning efficiency of experienced drivers over months in 2016.



(a) Shenzhen road map

(b) Map gridding

Fig. 3. Shenzhen map data.

The rest of the article is organized as follows. Section 2 motivates and defines the problem. Section 3 details the methodology. Section 4 presents evaluation results. Related works are discussed in Section 5, and the article concludes in Section 6.

2 OVERVIEW

In this section, we introduce the motivation of the proposed study and formally define the problem.

2.1 Motivation

It is a common perception that new drivers gradually learn how to make smart choices with regard to time and can improve their working efficiency over time. We verify this perception through data analysis. In Figure 1(b), the average earning efficiency of new drivers who joined in March 2014 increased by up to 100% in 2 years, whereas in Figure 2, the same measure of experienced drivers in 2016 did not change much. This can be explained by the fact that experienced drivers learned enough knowledge to make nearly optimal decisions.

We further noticed that drivers have very different learning curves, which ultimately affects earning improvements they can achieve. As mentioned previously, in Figure 1(b), the two colors represent two subgroups of new drivers who joined in March 2014. One group (in blue) includes those who became "top" drivers after 2 years with higher earning efficiency, and the other (gray) includes the rest of the drivers. Apparently, the former gained more useful knowledge, which contributed to their earning improvement.

Little is known about what specific knowledge the drivers gained and which pieces are contributing the most to the earning improvement. Answering these questions would potentially guide and train new drivers to become quick learners.

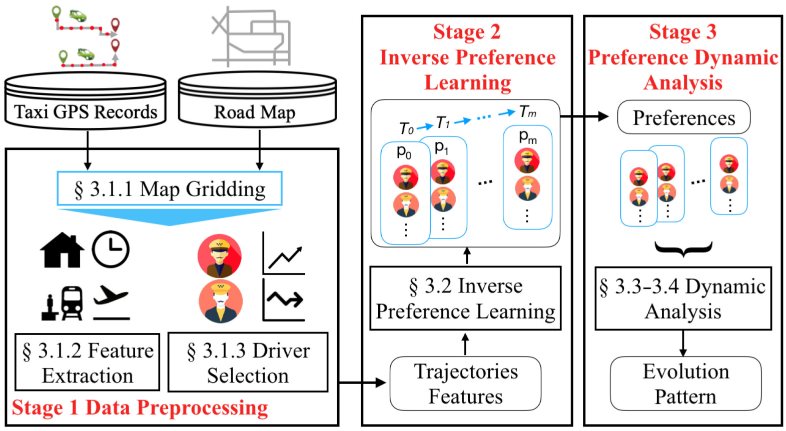


Fig. 4. DHPA framework.

We consider such "knowledge" as a series of preferences of a driver when making each decision, such as "how frequent to visit the train station" and "how far away from home to go when seeking passengers." Specifically, we extract features from the data to represent such decisions a taxi driver might face while working. To achieve the aforementioned goal, in this study we aim to answer two questions: (1) how to recover the preferences of taxi drivers when making these choices and (2) how these preferences change over time for different groups of drivers.

Problem definition. In a time interval T_0 (i.e., 1 month), given a taxi driver's trajectory data $\tilde{\mathcal{T}}$, and k environmental features $[f_0, f_1, \ldots, f_k]$, which influence drivers' decision-making process over time, we aim to learn the driver's preference $\theta = [\theta_0, \theta_1, \ldots, \theta_k]$ (i.e., weights to features when the driver makes decisions). Second, for a long time horizon, with multiple time intervals $[T_0, T_1, \ldots, T_m]$, we analyze the evolution pattern of the driver's preferences over time.

2.2 Data Description

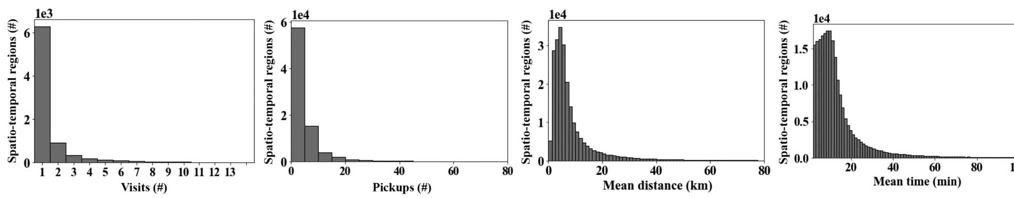
Our analytical framework takes two urban data sources as input, including (1) taxi trajectory data and (2) road map data. For consistency, both datasets were collected in Shenzhen, China, in 2014 and 2016.

The *taxi trajectory data* contain GPS records collected from taxis in Shenzhen, China, during March and November in 2014, and July to December in 2016. In total, 17,877 taxis were equipped with GPS sets, where each GPS set generated a GPS point every 40 seconds on average. Overall, a total of 51,485,760 GPS records were collected on each day, and each record contained five key data fields, including taxi ID, timestamp, passenger indicator, and latitude and longitude. The passenger indicator field is a binary value, indicating if a passenger is aboard or not.

The *road map data* of Shenzhen covers the area defined between 22.44° to 22.87° in latitude and 113.75° to 114.63° in longitude. The data is from OpenStreetMap [1] and has 21,000 roads of six levels. Figure 3(a) shows the road map in Shenzhen.

3 METHODOLOGY

Figure 4 outlines our DHPA framework, which takes two sources of urban data as inputs and contains three key analytical stages: (1) data preprocessing, (2) inverse preference learning, and (3) preference dynamic analysis.



(a) *P1:* Number of visits (b) *H1:* Number of pickups (c) *H2:* Mean trip distance (d) *H3:* Mean trip time

Fig. 5. Statistical distributions of $P1$, $H1$, $H2$, and $H3$.

3.1 Data Preprocessing

3.1.1 Map and Time Quantization. We use a standard quantization trick to reduce the size of the location space. Specifically, we divide the study area into equally sized grid cells with a given side-length s in latitude and longitude. Our method has two advantages: (1) we have the flexibility to adjust the side length to achieve different granularities, and (2) it is easy to implement and highly scalable in practice [10, 11]. Figure 3(b) shows the actual grid in Shenzhen, China, with a side-length $l = 0.01^\circ$ in latitude and longitude. Eliminating cells in the ocean, those unreachable from the city, and other irrelevant cells gives a total of 1,158 valid cells.

We divide each day into 5-minute intervals for a total of 288 intervals per day. A spatio-temporal region r is a pair of a grid cell s and a time interval t. The trajectories of drivers then can be mapped to sequences of spatio-temporal regions.

3.1.2 Feature Extraction. Taxi drivers make hundreds of decisions throughout their work shifts (e.g., where to find the next passenger, and when to start and finish working in a day). When making a decision, they instinctively evaluate multiple factors (i.e., features) related to their current states and the environment (i.e., the current *spatio-temporal region*). For example, after dropping off a passenger, a driver may choose to go back to an area that she is more familiar with or a nearby transport station (e.g., airport, train station). Here, we extract key features the drivers use to make their decisions.

Note in our framework that each feature is defined as a numeric characteristic of a specific spatio-temporal region, which may or may not change from driver to driver. For example, let f_r represent the average number of taxi pickups in history in location s during time slot t. Apparently, the value of feature f_r is the same for every driver. However, another feature g_r at r could be the distance from s to the home of the driver. The value of this feature varies from driver to driver, depending on their home locations. However, it does not change over time. The features we extract can be roughly categorized by *profile features* and *habit features*, as detailed next.

Profile features. Each driver has unique personal (or profile) characteristics, such as home location, daily working schedule (time duration), and preferred geographic area. For each spatio-temporal region, we build the profile features. Here, we extract four profile features:

- *P1: Visitation Frequency.* This group of features represents the numbers of daily visits to different regions of a driver as extracted from the historical data. Figure 5(a) shows the distribution of visitation frequency to different regions of an arbitrarily chosen driver. Here, visitation frequencies vary significantly across regions.
- *P2: Distance to Home.* Each taxi driver has a home location, which can be extracted from the driver's GPS records. This feature characterizes the distance (in miles on the road network) from the current location to the driver's home location. Different drivers may have different preferences in regard to working close to their homes or not.

P3 & P4: Time from Start & Time to Finish. Taxi drivers typically work according to consistent starting and finishing times. We construct two features to characterize the differences of the current time from the regular starting and finishing time.

Habit features. The habit features represent the habits of the drivers, which typically are governed by experience (e.g., remaining near the train station instead of traveling around to find passengers). We extract six habit features:

H1: Number of Pickups. This feature characterizes the demands in a cell during a time interval, and is extracted and estimated using the historical trajectories from all drivers. The distribution on the numbers of pickups is shown in Figure 5(b).

H2 & H3: Average Trip Distance & Time. These features represent the average distance and the travel time of passenger trips starting from a particular spatio-temporal region. Different spatio-temporal regions can have different expected trip distances and travel time. For example, the passengers picked up near the airport probably have longer trip distances than the passengers picked up near the train station since the airport is farther away from the downtown area than the train station. A driver's preferences to these features characterize to what extent the driver prefers long versus short distance trips and how well the driver gains knowledge on the lengths of trips over the spatio-temporal regions. The distribution of these features across spatio-temporal regions are showed in Figure 5(c) and Figure 5(d), respectively:

H4: Traffic Condition. This feature captures the average traffic condition based on the time spent by a driver in each spatio-temporal region. A long travel time implies traffic congestion. The preference of drivers over this feature represents how much drivers would like avoid the traffic.

H5 & H6: Distance to Train Station & Airport. These features reflect the distances from the current cell to Shenzhen train station and airport, respectively.

3.1.3 Driver Selection. Different drivers have different earning efficiencies, as shown in Figure 1(a). In the following, we describe the criteria we use to select drivers.

We estimate the earning efficiency of each driver in different time periods from their historical data. The estimated earnings E of a driver in the whole sampling span (e.g., per month) is calculated from the distance d_o that the taxi is occupied with passenger. The factors we take into consideration include the taxi fare in Shenzhen in 2014 and 2016, and the expense for gas. The taxi fare is 11 CNY for the first 2 km, and the charge for each additional kilometer is 2.4 CNY. The estimated expense for gas is 0.5 CNY per kilometer. The calculation of E is as follows:

$$E = \begin{cases} 11 - 0.5 * d_o & if\, d_o < 2 \\ 11 + (d_o - 2) * 2.4 - 0.5 * d_o & else. \end{cases} \tag{1}$$

Note that our model is easy to extend to other definitions of "earnings." Given the data we have, and without losing much accuracy in regard to calculated earnings in terms of representing the driver's profits, we employ Equation (1) to estimate the earnings of each driver.

The earning efficiency r_e is defined as the average per-hour income (i.e., in Equation (2)).

$$r_e = \frac{E}{t_w}, \tag{2}$$

where E is the income in the whole sampling time span, span (e.g., per month), and t_w represents the driver's working time.

Driver selection criterion. We select drivers with the highest earnings because the preference learning algorithms require the input data to be generated by the converged policy (see more details in Section 3.2). We note that drivers with high earning efficiencies are likely the most experienced (i.e., they use converged policies to make decisions).

3.2 Inverse Preference Learning

This section explains our inverse learning algorithm for extracting drivers' decision-making process. We use an MDP to model drivers' sequential decision making and relative entropy IRL to learn their decision-making preferences.

3.2.1 Markov Decision Process. A *Markov decision process* (MDP) [4] is defined by a 5-tuple $\langle S, A, T, \gamma, \mu_0, R\rangle$ so that

- S is a finite set of states and A is a finite set of actions,
- T is the probabilistic transition function with $T(s'|s,a)$ as the probability of arriving at state s' by executing action a at state s,
- $\gamma \in (0,1]$ is the discount factor,[2]
- $\mu_0 : S \to [0,1]$ is the initial distribution, and
- $R : S \times A \to$ is the reward function.

A randomized, memoryless policy is a function that specifies a probability distribution on the action to be executed in each state, defined as $\pi : S \times A \to [0,1]$.

We use $\tau = [(s_0, a_0), (s_1, a_1), \ldots, (s_L, a_L)]$ to denote a trajectory generated by the MDP. Here, L is the length of trajectory. We model the decision-making process of taxi drivers with the MDP as follows:

- *State*: A spatio-temporal region specified by a geographical cell and a time slot.
- *Action*: Traveling from the current cell to one of the eight neighboring cells, or staying in the same cell.
- *Reward*: The inner product of the preference function (as a vector) θ and the feature vector $\mathbf{f}$ on each state-action pair.

Note that in the MDP settings, the reward of each state-action pair is the inner product of the preference function and the feature vector. The preferences are the weights of each feature. The driver aims to maximize the accumulated reward when making decisions. To interpret each of the preferences recovered, we can consider two factors: the sign and the magnitude of the preference (weights of each feature). A positive preference means that the driver prefers to go to the regions where the corresponding feature value is higher than other locations, and larger magnitude can imply that the driver pays more attention to the corresponding feature, and vice versa for negative sign and smaller magnitude. In this work, we design two categories of features: profile features and habit features. The intuition is that the preferences for some features can be time invariant because these features are closely related to the drivers' profiles, such as home location and working schedule, which usually do not change.

The preferences to some other features can be time variant because these features are related to the habits of drivers, such as the number of pickups and the distance to train stations. The preferences for the habit features are considered as the habits of the drivers. The drivers can learn to change their habits (i.e., preferences to habit features) to improve their earning efficiencies.

Figure 6 shows an example of trajectory in the MDP: a driver starts in state s_0 with the taxi idle, and takes the action a_0 to travel to the neighboring cell S_1. After two steps, the driver reaches

[2]Without loss of generality, we assume $\gamma = 1$ in this work, and it is straightforward to generalize our results to $\gamma \neq 1$.

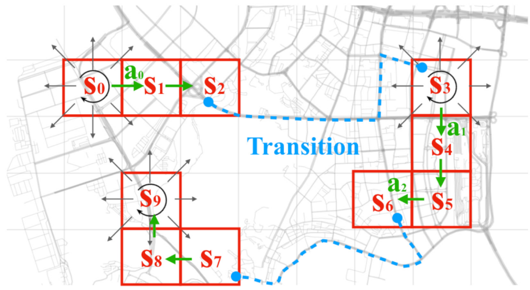


Fig. 6. MDP of a taxi driver's decision-making process.

state S_2, where she meets a passenger. The destination of the new trip is cell S_3. The trip with the passenger is a transition in the MDP from S_2 to S_3.

3.2.2 Inverse Preference Learning. Given the observed trajectory set $\tilde{\mathcal{T}}$ of a driver and the features extracted on each state-action pair (s, a), the inverse preference learning stage aims to recover a reward function (i.e., preference vector θ) under which the observed trajectories have the highest likelihood to be generated [15]. Various IRL approaches, such as apprenticeship learning [3], maximum entropy IRL [31], Bayesian IRL [17], and relative entropy IRL [5], have been proposed in the literature.

Our problem possesses two salient characteristics. First, the state space is large. We have 1,158 cells and 288 time intervals. Therefore, the total number of states is $1{,}158 \times 288 \approx 330$k. Second, the transition probability is hard to measure, partly because of the large state space issue.

Therefore, we adopt a model-free IRL approach, namely relative entropy IRL [5], that does not require estimating transition probabilities and is more scalable than other alternatives.

The optimization problem. Let $\mathcal{T}$ denote the set of all possible trajectories of the driver decision-making MDP, outlined in Section 3.2.1. For any $\tau \in \mathcal{T}$, denote $P(\tau)$ as the trajectory distribution induced by the taxi driver's ground-truth policy and $Q(\tau)$ as the trajectory distribution induced by a base policy. The relative entropy between $P(\tau)$ and $Q(\tau)$ (in Equation (3)) characterizes the distribution difference between $P(\tau)$ and $Q(\tau)$:

$$H(P\|Q) = \sum_{\tau \in \mathcal{T}} P(\tau) \ln \frac{P(\tau)}{Q(\tau)}. \tag{3}$$

The driver's trajectory distribution is governed by the driver's preference θ and thus is a function of θ (i.e., $P(\tau|\theta)$). The relative entropy IRL aims to find a reward function θ that minimizes the relative entropy in Equation (3) and matches the trajectory distribution to the observed trajectory data.

***P1: Relative Entropy IRL Problem.* The relative entropy IRL problem is presented as follows:**

$$\min_{\theta}: \quad H(P(\theta)\|Q) = \sum_{\tau \in \mathcal{T}} P(\tau|\theta) \ln \frac{P(\tau|\theta)}{Q(\tau)}, \tag{4}$$

$$\text{s.t.:} \quad \left| \sum_{\tau \in \mathcal{T}} P(\tau|\theta) f_i^{\tau} - \hat{f}_i \right| \leq \epsilon_i, \forall i \in \{1, \ldots, k\}, \tag{5}$$

$$\sum_{\tau \in \mathcal{T}} P(\tau|\theta) = 1, \tag{6}$$

$$P(\tau|\theta) \geq 0, \qquad \forall \tau \in \mathcal{T}, \tag{7}$$

where i is the feature index, f_i^τ is the i's feature count in trajectory τ, and $\hat{f}_i = \sum_{\tau \in \tilde{\mathcal{T}}} f_i^\tau / |\tilde{\mathcal{T}}|$ is the feature expectation over all observed trajectories in $\tilde{\mathcal{T}}$. ϵ_i is a confidence interval parameter, which can be determined by the sample complexity (the number of trajectories) via applying a Hoeffding's bound. The constraint Equation (5) ensures that the recovered policy matches the observed data. The constraints (Equations (6) and (7)) guarantee that the $P(\tau|\theta)$'s are non-negative probabilities and thus sum up to 1.

Solving P1. The function $Q(\tau)$ and $P(\tau|\theta)$ can be decomposed as

$$Q(\tau) = T(\tau)U(\tau) \text{ and } P(\tau|\theta) = T(\tau)V(\tau|\theta),$$

where $T(\tau) = \mu_0(s_0) \prod_{t=1}^{K} T(s_t|s_{t-1}, a_{t-1})$ is the joint probability of the state transitions in τ, for $\tau = [(s_0, a_0), (s_1, a_1), \ldots, (s_K, a_K)]$, with $\mu_0(s_0)$ as the initial state distribution. $U(\tau)$ (resp.ectively, $V(\tau|\theta)$) is the joint probability of the actions conditioned on the states in τ under driver's policy π_θ (respectively, a base policy π_q). As a result, Equation (4) can be written as follows:

$$H(P(\theta)\|Q) = \sum_{\tau \in \mathcal{T}} P(\tau|\theta) \ln \frac{V(\tau|\theta)}{U(\tau)}. \tag{8}$$

Moreover, when $\pi_q(a|s)$ at each state s is uniform distribution (e.g., $\pi_q(a|s) = 1/|A_s|$, with A_s as the set of actions at state s), the problem **P1** is equivalent to maximizing the causal entropy of $P(\tau|\theta)$ (i.e., $\sum_{\tau \in \mathcal{T}} P(\tau|\theta) \ln V(\tau|\theta)$) while matching $P(\tau|\theta)$ to the observed data [30]. Following the similar process outlined in Boularias et al. [5], **P1** can be solved by a gradient descent approach, with the stepwise updating gradient as follows:

$$\nabla g(\theta) = \hat{f}_i - \frac{\sum_{\tau \in \mathcal{T}^\pi} \frac{U(\tau)}{\pi(\tau)} \exp(\sum_{j=1}^{k} \theta_i f_j^\tau)}{\sum_{\tau \in \mathcal{T}^\pi} \frac{U(\tau)}{\pi(\tau)} \exp(\sum_{j=1}^{k} \theta_i)} - \alpha_i \epsilon_i, \tag{9}$$

where $\alpha_i = 1$ if $\theta_i \leq 0$ and $\alpha_i = -1$ otherwise. $\mathcal{T}^\pi$ is a set of trajectories sampled from $\tilde{\mathcal{T}}$ by an executing a given policy π. $U(\tau)$ is the joint probability of taking actions conditioned on the states in a observed trajectory τ, induced by uniform policy $\pi_q(a|s) = 1/|A_s|$.

See Algorithm 1 for our IRL algorithm.

ALGORITHM 1: Relative Entropy IRL

Input: Demonstrated trajectories $\tilde{\mathcal{T}}$, feature matrix F, threshold vector ϵ, learning rate α, and executing policy π.
Output: Preference vector θ.

```
Randomly initialize preference vector θ.
Sample a set of trajectories. 𝒯^π using π.
Calculate feature expectation vector f̂.
repeat
    Calculate each feature count f_i^τ.
    Calculate gradient ∇g(θ) using Equation (9).
    Update θ ← θ + α∇g(θ).
until ∇g(θ) < ϵ.
```

3.3 Preference Dynamic Analysis

Using Algorithm 1, we can inversely learn the preference θ for each driver, during each time interval (e.g., a month) over time, and obtain a sequence of preference vectors $\{\theta_1, \ldots, \theta_N\}$. For each driver, we can conduct hypothesis testing to examine if the change of the preference vectors over months is significant or not. We denote the preference vector learned for taxi driver p in period T_i as θ_i^p and that in period T_j as θ_j^p. Then, we can obtain two preference vector sample sets in i-th and j-th months as S_i and S_j over a group of n drivers as follows:

$$S_i = \left\{\theta_i^1, \theta_i^2, \ldots, \theta_i^n\right\}, \tag{10}$$

$$S_j = \left\{\theta_j^1, \theta_j^2, \ldots, \theta_j^n\right\}. \tag{11}$$

With S_i and S_j, we will examine if the entries in preference vectors changed significantly or not from the i-th to j-th month using a paired sample t-test [22]. For each feature f_m, the null hypothesis is that the difference between the m-th entry of each θ_i^p in S_i and θ_j^p in S_j equals 0, which means drivers' preference to feature f_m does not change significantly from the i-th month to the j-th month. Otherwise, the alternative hypothesis indicates a significant change. Taking the difference between S_i and S_j as $\Delta S_{ij} = \{\Delta\theta_{ij}^1, \Delta\theta_{ij}^2, \ldots, \Delta\theta_{ij}^n\} = \{\theta_i^1 - \theta_j^1, \theta_i^2 - \theta_j^2, \ldots, \theta_i^n - \theta_j^n\}$.

The t-test statistics of the m-th entry is as follows:

$$t_{ij}(m) = \frac{Z}{s} = \frac{\Delta\overline{\theta}_{ij}(m) - \mu}{\delta/\sqrt{n}}, \tag{12}$$

where μ is the sample mean, n is the sample size, and δ is the sample square error. The t-distribution for the test can be determined given the degree of freedom $n-1$. Given a significance value $0 < \alpha < 1$, we can get a threshold of the t-value t_α in the t-distribution. Then, if $t_{ij}(k) > t_\alpha$, the null hypothesis should be rejected with significance α; otherwise, we can accept the null hypothesis with significance α. Usually, we set $\alpha = 0.05$, which also means that the confidence of the test is $1 - \alpha = 0.95$.

3.4 Preference Trend Analysis

In Section 3.3, we employ the hypothesis test to examine if the change of the preference over months for each driver is significant or not. In this section, we investigate the trends of the preferences for different groups of drivers regarding the significantly changed preferences. For the significantly changed preference to feature f_s, we can obtain the preference of driver k in the i-th and j-th ($i < j$) months as $\theta_i^k(s)$ and $\theta_j^k(s)$. Then, the set of preferences to feature f_s over a group of n drivers in the i-th month can be denoted as follows:

$$S_i(s) = \left\{\theta_i^1(s), \theta_i^2(s), \ldots, \theta_i^n(s)\right\}, \tag{13}$$

$$S_j(s) = \left\{\theta_j^1(s), \theta_j^2(s), \ldots, \theta_j^n(s)\right\}. \tag{14}$$

We want to investigate in detail how the preferences change (i.e., some preferences trend up over time, whereas others trend down). Here, we define the positive trend rate r_p in Equation (15) to characterize the increasing trend of the each preference for each group of drivers, and the negative trend rate r_n to characterize the decreasing trend, which equals $1 - r_p$:

$$r_p = \frac{\sum_{k=1}^{n} I_{ij}^s(k)}{n}, \tag{15}$$

where n is the number of drivers in the group and

$$I_{ij}^s(k) = \begin{cases} 1 \ if \ \theta_i^k(s) < \theta_j^k(s) \\ 0 \ if \ \theta_i^k(s) > \theta_j^k(s). \end{cases} \tag{16}$$

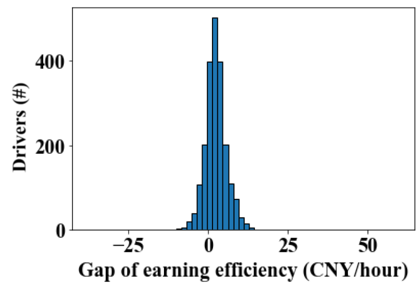


Fig. 7. Earning efficiency gap distribution.

4 EXPERIMENTS

In this section, we conduct experiments with real-world taxi trajectory data to learn the preferences of different groups of taxi drivers and analyze the preference evolution patterns for each group.

4.1 Experiment Settings

When analyzing the temporal dynamics of the drivers' decision-making preferences, the null hypothesis is that the difference between the preferences in two time periods is not significant. The alternative hypothesis is that the temporal preference difference is significant. We choose the t-test significance value $\alpha = 0.05$.

Driver group selection. We aim to analyze how taxi drivers' decision making preferences evolve over time. For each month, we select 3,000 drivers with the highest earning efficiency. The reason we select these drivers is that they are likely more experienced drivers, thus with near-optimal policies, which is required by the maximum entropy principle [31] to ensure precise preferences recovered by IRL from the demonstrations. To evaluate the preference change across 2 months (i.e., the i-th and j-th months), we find those drivers from those experience drivers, who also show up in both months for our study. For example, in July and December of 2016, there are 2,151 experienced drivers in common.

Then, we calculate the difference of earning efficiency of each driver in the 2 months. Figure 7 shows the gap distribution in July and December of 2016. We will choose three groups of drivers for preference dynamics analytics based on the drivers' earning efficiency gaps:

- *Group #1 (self-improving drivers):* 200 drivers whose earning efficiencies increase the most.
- *Group #2 (stabilized drivers):* 200 drivers whose earning efficiency gaps are small (i.e., close to 0).
- *Group #3 (exploring drivers):* 200 drivers whose earning efficiencies decrease the most.

The self-improving drivers are more likely to have learned a lot during the time span from July to December of 2016. By analyzing their preference dynamics, we can get a sense of how they learned over time. The stabilized drivers are those whose earning efficiencies did not change much from July to December of 2016, and we want to validate if their preferences were also stable during the time span to cross validate how taxi drivers gain knowledge over time. The exploring drivers are those whose earning efficiencies decrease the most from July to December of 2016, and we want to figure out why this happened to these drivers by analyzing their learning curve via our DHPA framework. As the first attempt to analyze the learning curve of taxi drivers, in this work we do not explore deeper to individual taxi drivers. In our future work, we will explore the learning curve of individual taxi drivers.

Group #1 (Self-Improving Drivers)

	P1	P2	P3	P4	H1	H2	H3	H4	H5	H6
Aug	-1.09	0.82	-0.27	0.61	0.15	-0.44	-0.03	0.93	1.23	-1.33
Sep	0.70	0.79	0.26	-0.36	-1.88	1.11	0.66	0.70	-0.33	-0.43
Oct	-0.18	0.08	-0.48	0.51	0.02	-1.66	0.89	-1.30	0.19	1.12
Nov	1.75	-0.96	-0.10	-1.63	-2.20	0.58	1.39	2.80	1.30	-0.34
Dec	-0.43	0.43	-0.32	-0.10	-2.51	-0.28	2.22	2.11	0.01	-0.34

Fig. 8. Preference dynamics between July and each of the following 5 months of Group #1 drivers.

Group #2 (Stabilized Drivers)

	P1	P2	P3	P4	H1	H2	H3	H4	H5	H6
Aug	-1.05	-1.23	1.30	-1.28	0.71	-0.66	-0.28	-1.94	-0.47	1.21
Sep	-0.08	-0.23	1.44	-0.85	0.09	0.83	-0.99	-0.98	-0.63	0.24
Oct	0.04	-1.74	1.87	-1.20	0.84	0.25	-0.63	-1.48	-0.19	-0.40
Nov	-0.62	0.77	-0.11	-0.95	1.05	-0.25	-1.50	-1.06	0.44	0.05
Dec	0.88	-1.13	1.89	0.36	-0.21	-0.36	0.57	-0.76	-1.27	-0.11

Fig. 9. Preference dynamics between July and each of the following 5 months of Group #2 drivers.

Experiment plan. We use 12 months' trajectory data across 3 years for our study (i.e., July–December 2014 and July–December 2016). We evaluate the preference dynamics across month pairs. First, we set up the month July of 2016 as the base month and compare the preferences of drivers in Group #1 and Group #2 with that of 5 subsequent months (August–December 2016), respectively, to examine the dynamics of potential habits' preferences in a short period and a long period. We define the short period as July and August of 2016 and July through September of 2016, and the long period as July through November of 2016 and July through December of 2016. We apply familywise t-tests with Bonferroni correction to avoid an inflation of false positives.

4.2 Preference Dynamics Analysis

Now we present the results of the analysis of the preference dynamics of two driver groups over time.

4.2.1 Results for Group #1. The table in Figure 8 shows the t-values obtained for comparing preferences (with respect to each feature) in July 2016 to that of August through December of 2016, respectively. For these self-improving drivers, the boxes of failed tests are marked with red color and the corresponding t-values. Note that these tests are conducted individually without comparisons among them because we want to examine whether there exists a significant preference change for individual features in a specific month compared with July. First, with a time span of less than 3 months, the preferences do not show any significant change. However, when the time space is larger than 3 months, preferences to some habit features change significantly if viewed as individual tests, including *H1: Number of Pickups*, *H3: Average Trip Time*, and *H4: Traffic Condition*. This makes sense, as over time the self-improving drivers tend to gain the knowledge of where the demands, low traffic, and long trip orders are. However, the preferences to all four profile features and other habit features stay unchanged over half a year.

According to the results of the preceding individual tests, we notice that the preferences to three habit features (H1, H3, H5) might change significantly over a long period (i.e., after 3 months). To validate these preference dynamics over a long period, we conduct familywise hypothesis tests to examine if the preferences to these features change significantly after a long period. We consider July to August and July to September as the short period, and July to November and July to December as the long period. Since only the preferences to the three habit features potentially change significantly, we have six tests in total. After Bonferroni correction, the results are presented in Figure 10. We observe that, after a long period, the preferences to *H1: Number of Pickups* and *H4: Traffic Condition* change significantly, whereas the preference to *H3: Average Trip Time* does not show a significant change. In addition, preferences to these three habit features do not change after a short period.

4.2.2 Results for Group #2. The table in Figure 9 shows the t-values obtained for preference comparison of drivers in Group #2.

	H1	H3	H4
08 & 09	-1.18	0.45	1.20
11 & 12	-2.83	2.09	3.19

Fig. 10. Validation on the long-term preference dynamics.

Group #3 (Exploring Drivers)

	P1	P2	P3	P4	H1	H2	H3	H4	H5	H6
Aug	0.94	-0.35	0.76	-0.68	0.35	-1.23	1.64	-1.33	0.45	-0.96
Sep	-2.32	-1.22	1.43	2.43	2.45	-3.22	-2.09	1.16	-1.35	-2.33
Oct	-2.58	1.78	-3.43	-4.24	-3.35	-4.92	4.11	-3.43	3.22	-5.25
Nov	-3.54	-2.45	2.54	2.23	2.97	3.21	-2.25	-2.63	2.23	2.44
Dec	3.11	-2.31	2.37	2.55	-3.18	-2.54	3.77	-2.75	-3.47	-2.87

Fig. 11. Preference dynamics between July and each of the following 5 months of Group #3 drivers.

Group #3 (Exploring Drivers)

	P1	P2	P3	P4	H1	H2	H3	H4	H5	H6
Aug	0.51	0.47	0.42	0.43	0.49	0.44	0.52	**0.63**	0.44	0.45
Sep	0.49	0.47	0.50	**0.38**	**0.39**	0.55	**0.37**	0.47	0.53	**0.39**
Oct	0.55	**0.60**	0.46	0.45	**0.38**	**0.62**	**0.63**	0.51	**0.63**	0.53
Nov	**0.64**	0.45	**0.39**	0.48	0.44	**0.37**	0.55	0.44	0.55	**0.65**
Dec	0.52	0.51	0.59	0.50	**0.61**	0.45	**0.36**	0.57	**0.37**	0.47

Fig. 12. Positive rates of each preference in Group #3.

	P1	P2	P3	P4	H1	H2	H3	H4	H5	H6
Group #1	0.48	0.49	0.47	0.51	**0.38**	0.46	**0.63**	**0.65**	0.43	0.57
Group #2	0.57	0.47	0.58	0.50	0.46	0.45	0.49	0.51	0.48	0.51

Fig. 13. Positive rate of each preference in Groups #1 and #2.

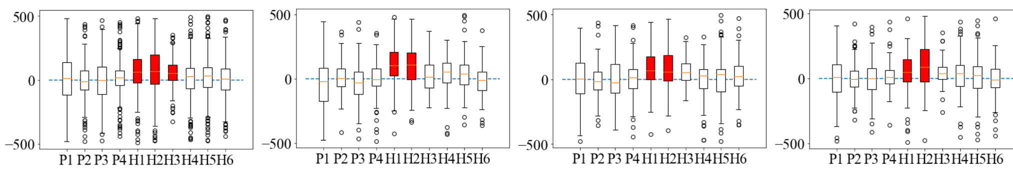
Clearly, the preferences to all profile and habit features stay unchanged over half a year, which means that these stabilized drivers have kept the same strategy of finding passengers over half a year. This is consistent with their unchanged earning efficiencies over time. The reason the stabilized drivers maintained stable preferences is either because they were very experienced and already obtained the optimal strategies given the profiles they had, which means that they had reached the upper bound of earning efficiency for the drivers who have similar profiles, or they had found a near optimal strategy and still have potential space for improvement but were not motivated to find a better one, or some other potential reasons. As the first attempt to analyze the preference dynamics in this work, we do not dig that deep to figure out the exact reasons these drivers maintained stabilized preferences, but we will investigate this problem in future work.

4.2.3 Results for Group #3. The table in Figure 11 shows the t-values obtained for preference comparison of drivers in Group #3. The preference of this group of drivers changed a lot over months, most habit features changed, and the preferences to one or two profile features changed in November and December. Group #3 drivers are the exploring drivers, whose earning efficiencies drop over these months, and the results reveal that they try changing their preferences significantly over time to explore new strategies, but the attempts do not work for the growth of their earning efficiencies.

4.3 Preference Trend Analysis

In this section, we present the results of the preference trend analysis over three driver groups.

Results for Group #1. The results of the preference trend analysis for Group #1 are shown in Figure 13. The values in the table are r_p's for each preference between July and December. The values in red indicate that most drivers have a positive trend on the preference, which is consistent with the result of the preference dynamics analysis, since the most significantly changed preferences are



(a) All drivers (b) Group 1: Self-improving drivers (c) Group 2: Stabilized drivers (d) Groups 3: Exploring drivers

Fig. 14. Box plots of the preferences in December 2016.

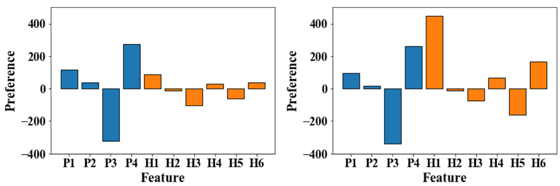
the same (i.e., the preferences to *H1: Number of Pickups*, *H3: Average Trip Time*, and *H4: Traffic Condition*). For example, the preference to feature H1 has $r_n = 0.62$, which is prominently larger than $r_p = 0.38$. This indicates that most of the self-improving drivers tend to reduce their preference to feature H1 (i.e., the number of pickups). This is reasonable because the regions with a large number of pickups are usually crowded (e.g., downtown area). To complete a service trip is time consuming, which can damage the earning efficiency of taxi drivers. As for the significantly changed preferences to features *H3: Average Trip Time* and *H4: Traffic Condition*, the r_p's are prominently larger than the r_n's, respectively, which indicates that the number of self-improving drivers who increase their preferences on feature H3 and H4 is prominently greater than those who decrease.

Results for Group #2. The results of the preference trend analysis for Group #2 are shown in Figure 13. We notice that the r_p's are close to 0.50 regarding the preference to each of the features, which is consistent with the results of the preference dynamics analysis in Section 4.2.

Results for Group #3. The results of the preference trend analysis for Group #3 are shown in Figure 12. The values in the table are the r_p's calculated between July and each of the following months. The values in red indicate that most drivers have a positive trend on the preference compared with July, and the blue ones indicate that most drivers have a negative trend on the preference compared with July. We find that the exploring drivers change their preferences randomly. Taking the preference to feature *H5: Distance to Train Station* as an example, a significant positive trend is found in August and October, whereas in December it switches to a negative trend. Similar patterns can be found in the preferences to features H1, H3, and P3.

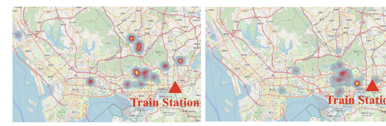
4.4 Preference Distribution Analysis

To explore the different features that different groups of drivers pay attention to, we analyze the distribution of the preferences in December for self-improving (Figure 14(b)), stabilized (Figure 14(c)), and exploring (Figure 14(d)) groups, as well as the entire taxi driver population (Figure 14(a)). We find that in Figure 14(a), the preference median of profile features are all close to 0, and the preference medians to *H1: Number of Pickups*, *H2: Average Trip Distance*, and *H3: Average Trip Time* are relatively higher than the other three habit features, which implies that overall, taxi drivers prefer a higher number of pickups and longer trips. From Figure 14(b) through (d), we find that the self-improving drivers and the stabilized drivers have higher preferences to *H1: Number of Pickups*, which indicates that they gained sufficient knowledge on which regions have high demands. In contrast, the exploring drivers have lower preferences to *H1: Number of Pickups*. This reveals that they are still learning the distribution pattern of travel demands. Moreover, the preference to *H2: Average Trip Distance* of the exploring drivers is higher than that of other groups, which implies that the exploring drivers paid excessive attention to the long distance trips. This may be one of the negative effects leading to a decreasing earning efficiency trend, because the longer the trip distance, the more time it costs.



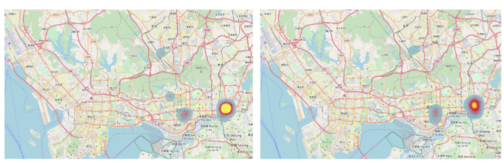
(a) Preference in 07/16 (b) Preference in 12/16

Fig. 15. The decision-making preferences of John.



(a) July, 2016 (b) December, 2016

Fig. 16. Heatmap of trajectory of John.



(a) July, 2016 (b) December, 2016

Fig. 17. Heatmap of pickups of John.

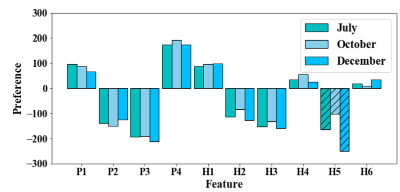


Fig. 18. The decision-making preferences of Mike in July, October, and December.

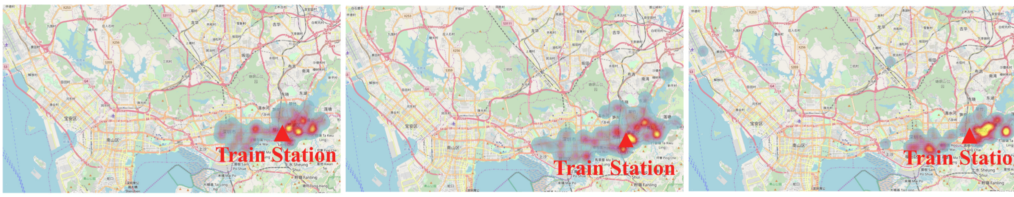
4.5 Case Study

4.5.1 Case of Preference Dynamics. To further understand the preference dynamics, we look into individual drivers to showcase how the preference and working behaviors evolve over time. Here we show one randomly selected driver from Group #1. Let us call him "John." John's earning efficiency grew from 41.84 CNY/hour to 52.24 CNY/hour from July to December of 2016. His preferences in the months of July and December are listed in Figure 15. Clearly, the preferences to the profile features remain unchanged, whereas the preferences to some habit features, such as *H1: Number of Pickups* and *H5 & H6: Distance to Train Station & Airport*, changed. When we look into John's driving behaviors, it matches the preference change perfectly.

Preference change to H1. The preference change to feature H1 indicates that John increased his preference to areas with a high volume of pickup demands. Figure 16(a) and (b) show the distribution of trajectories when the taxi was idle in the morning rush hours in July 2016 and December 2016, respectively. Figure 17(a) and (b) show the all-taxi pickup demand distributions in the morning rush hours in July and December of 2016, respectively. The citywide demand distribution does not change. However, during the morning rush hours, John changed his strategy from July to December of 2016 (i.e., to look for passengers from the high-demand areas). This is consistent to the preference change to feature H1 (number of pickups).

Preference change to H5. The preference change to feature H5 (i.e., distance to train station) is also significant. The negative preference indicates that John prefers to be closer to the train station to look for passengers. Over time, this preference became stronger. To explain this phenomenon, we highlighted the train station in Figure 16. The statistics we obtained from John's trajectory data showed that the percentage of orders received near the train station increased from 11.93% in July 2016 to 14.21% in December 2016, which is consistent with the preference change.

The results of preference dynamics analysis in Section 4.2.1 show the overall pattern of Group #1, which illustrate that the self-improving drivers showed significant dynamics on their preferences to features *H1: Number of Pickups* and *H4: Traffic Condition*, and the preference to H1 tends to



(a) July, 2016 (b) October, 2016 (c) December, 2016

Fig. 19. Heatmap of trajectory of Mike.

decrease in the group. However, in the case study, "John" is an individual driver randomly selected from the group, who happened to have a reversed trend regarding the preference to feature H1. This is not a contradiction since the results in Section 4.2.1 are for the overall group, whereas in the case study, the randomly selected driver is an individual in the group, who might not maintain the same preference dynamics as the whole group.

4.5.2 Case of Preference Trend. To further understand the preference trend, we also look into an individual driver's showcase to investigate how exactly the preference changed over time from July to December. We randomly select a driver from Group #3. Let's call him "Mike." Mike's earning efficiency dropped from 46.15 CNY/hour to 41.34 CNY/hour from July to December in 2016. His preferences in July, October, and December of 2016 are shown in Figure 18. The preferences to some features changed randomly (e.g., H5). H5 is the habit feature: distance to the train station. The negative values indicate that Mike prefers to be closer to the train station. Compared with the preference to H5 in July, the preference becomes weaker in October and becomes stronger in December. To explain this, we visualize the distributions of the trajectories of Mike in Figure 19. The distribution near the train station becomes more scattered in October and more concentrated in December compared with that in July.

4.6 Takeaways and Discussions

From our studies on a large amount of taxi trajectory data spanning for 3 years, we made the first-ever report on how real-world taxi drivers make decisions when looking for passengers and how their preferences evolve over time. Overall, three key takeaways are summarized as follows:

1. Each driver has unique preferences to the driver's profile features, which tend to be stable over time.
2. While learning the environments, drivers may change their preferences to habit features.
3. While exploring the environments, drivers may change their preferences to profile and habit features randomly.

Our findings can potentially be utilized to assist and guide taxi drivers to improving their earning efficiencies. For example, for those slow-learning drivers, by learning their preferences, especially the preferences to habit features, we can diagnose which knowledge in terms of the features they are lacking (e.g., not familiar with the high-demand regions). As a result, some guiding messages may be sent directly to the drivers about such information to assist the drivers to improve a better policy faster. In addition, our proposed DHPA framework can easily adapt to different time interval analyses (e.g. over months, over days, over time in a day). One only needs to change the trajectory extraction in Stage 1 according to the different settings of time intervals. Due to space limitations, we may not present the results of all possible settings of time intervals in this article.

5 RELATED WORKS

Taxi operating strategies (e.g., dispatching, passenger seeking) and driver behavior analysis have been studied extensively in recent years due to the emergence of the ride-sharing business model and urban intelligence. However, to the best of our knowledge, we make the first attempt to employ IRL to analyze the preference dynamics of taxi drivers. Works related to our study are summarized next.

Urban computing, transportation, and geo-informatics are general research areas that integrate urban sensing, data management, and data analytics together as a unified process to explore, analyze, and solve crucial problems related to people's everyday life [6, 7, 9, 10, 12–14, 20, 21, 23, 26, 27, 29]. In particular, several works study taxi operation management, such as dispatching [8, 19] and passenger seeking [24, 25, 28], aiming at finding an optimal actionable solution to improve the performance/revenue of individual taxi drivers or the entire fleet.

Rong et al. [18] solved the passenger seeking problem by giving direction recommendations to drivers. However, all of these works focus on finding "what" are the best driving strategies (as an optimization problem) rather than finding "why" and "how" good drivers make these decisions. By contrast, our work focuses on analyzing the evolving preferences of good drivers that helped them make better and more profitable decisions.

IRL aims to recover the reward function under which the expert's policy is optimal from the observed trajectories of an expert. There are various IRL methods. For example, Ng and Russell [15] found that there is a class of reward functions that can lead to the same optimal policy, and it proposed a strategy to select a reward function. However, this method is not proper for analyzing human behaviors because it uses the deterministic policy in the MDP, whereas human decisions tend to be non-deterministic. Ziebart et al. [31] proposed an IRL method by maximizing the entropy of the distribution on state actions under the learned policy. Although this method can employ stochastic policy, the computation efficiency is not friendly to large-scale state space, and it requires the information of the model. In this article, we employ relative entropy IRL [5], which is model free and employs softmax policy. Our work, compared with the preceding related work, is the first to apply IRL to study the evolving driving preferences of taxi drivers.

6 CONCLUSION

In this article, we made the first attempt to employ IRL to analyze the preferences of taxi drivers when making sequences of decisions to look for passengers. We further studied how the drivers' preferences evolve over time, during the learning processes. This problem is critical to helping new drivers improve performance quickly. We extracted different types of interpretable features to represent the potential factors that affect the decisions of taxi drivers and inversely learned the preferences of different groups of drivers. We conducted experiments using large-scale taxi trajectory datasets, and the results demonstrated that drivers tend to improve their preferences to habit features to gain more knowledge in the learning phase and keep the preferences to profile features stable over time.

REFERENCES

[1] OpenStreetMap. [n.d.]. Home Page. Retrieved November 23, 2019 from http://www.openstreetmap.org/
[2] [n.d.]. Project Website: Dissecting the Learning Curve of Taxi Drivers. Retrieved November 23, 2019 from http://urban.cs.wpi.edu/DLCTD/.
[3] Pieter Abbeel and Andrew Y. Ng. 2004. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the 21st International Conference on Machine Learning*. ACM, New York, NY, 1.
[4] Richard Bellman. 1957. A Markovian decision process. *Journal of Mathematics and Mechanics* 6, 5 (1957), 679–684. IUMJAB

[5] Abdeslam Boularias, Jens Kober, and Jan Peters. 2011. Relative entropy inverse reinforcement learning. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*. 182–189.

[6] Ye Ding, Yanhua Li, Ke Deng, Haoyu Tan, Mingxuan Yuan, and Lionel M. Ni. 2016. Detecting and analyzing urban regions with high impact of weather change on transport. *IEEE Transactions on Big Data* 3, 2 (2016), 126–139.

[7] Tianfu He, Jie Bao, Ruiyuan Li, Sijie Ruan, Yanhua Li, Chao Tian, and Yu Zheng. 2018. Detecting vehicle illegal parking events using sharing bikes' trajectories. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'18)*. 340–349.

[8] J. Yuan, Y. Zheng, L. Zhang, and X. Xie. 2013. T-Finder: A recommender system for finding passengers and vacant taxis. *IEEE Transactions on Knowledge and Data Engineering* 25, 10 (2013), 2390–2403.

[9] Amin Vahedian Khezerlou, Xun Zhou, Lufan Li, Zubair Shafiq, Alex X. Liu, and Fan Zhang. 2017. A traffic flow approach to early detection of gathering events: Comprehensive results. *ACM Transactions on Intelligent Systems and Technology* 8, 6 (2017), 74.

[10] Yanhua Li, Jun Luo, Chi-Yin Chow, Kam-Lam Chan, Ye Ding, and Fan Zhang. 2015. Growing the charging station network for electric vehicles with trajectory data analytics. In *Proceedings of the 2015 IEEE 31st International Conference on Data Engineering (ICDE'15)*.

[11] Yanhua Li, Moritz Steiner, Jie Bao, Limin Wang, and Ting Zhu. 2014. Region sampling and estimation of geosocial data with dynamic range calibration. In *Proceedings of the 2014 IEEE 30th International Conference on Data Engineering (ICDE'14)*.

[12] Chen Liu, Ke Deng, Chaojie Li, Jianxin Li, Yanhua Li, and Jun Luo. 2016. The optimal distribution of electric-vehicle chargers across a city. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'16)*. IEEE, Los Alamitos, CA.

[13] Bo Lyu, Shijian Li, Yanhua Li, Jie Fu, Andrew C. Trapp, Haiyong Xie, and Yong Liao. 2016. Scalable user assignment in power grids: A data driven approach. In *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL GIS'16)*. ACM, New York, NY.

[14] M. Qu, H. Zhu, J. Liu, G. Liu, H. Xiong. 2014. A cost-effective recommender system for taxi drivers. In *Proceedings of the 20th International Conference on Knowledge Discovery and Data Mining (SIGKDD'14)*. ACM, New York, NY, 45–54.

[15] Andrew Y. Ng and Stuart J. Russell. 2000. Algorithms for inverse reinforcement learning. In *Proceedings of the 17th International Conference on Machine Learning (ICML'00)*. 663–670.

[16] Menghai Pan, Yanhua Li, Xun Zhou, Zhenming Liu, Rui Song, Hui Lu, and Jun Luo. 2019. Dissecting the learning curve of taxi drivers: A data-driven approach. In *Proceedings of the 2019 SIAM International Conference on Data Mining*. 783–791.

[17] Deepak Ramachandran and Eyal Amir. 2007. Bayesian inverse reinforcement learning. *Urbana* 51, 61801 (2007), 1–4.

[18] Huigui Rong, Xun Zhou, Chang Yang, Zubair Shafiq, and Alex Liu. 2016. The rich and the poor: A Markov decision process approach to optimizing taxi driver revenue efficiency. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management*. ACM, New York, NY, 2329–2334.

[19] Wolfson S. Ma, Y. Zheng. 2013. A large-scale dynamic taxi ridesharing service. In *Proceedings of the 29th International Conference on Data Engineering (ICDE'13)*. IEEE, Los Alamitos, CA, 410–421.

[20] Yongxin Tong, Yuqiang Chen, Zimu Zhou, Lei Chen, Jie Wang, Qiang Yang, Jieping Ye, and Weifeng Lv. 2017. The simpler the better: A unified approach to predicting original taxi demands based on large-scale online platforms. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 1653–1662.

[21] Jie Wang, Helai Huang, Pengpeng Xu, Siqi Xie, and S. C. Wong. 2019. Random parameter probit models to analyze pedestrian red-light violations and injury severity in pedestrian-motor vehicle crashes at signalized crossings. *Journal of Transportation Safety & Security*. Published online January 24, 2019. DOI: https://doi.org/10.1080/19439962.2018.1551257

[22] R. S. Witte and J. S. Witte. 2013. *Statistics* (10th ed.). John Wiley & Sons.

[23] Tong Xu, Hengshu Zhu, Xiangyu Zhao, Qi Liu, Hao Zhong, Enhong Chen, and Hui Xiong. 2016. Taxi driving behavior analysis in latent vehicle-to-vehicle networks: A social influence perspective. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 1285–1294.

[24] Y. Ge, H. Xiong, A. Tuzhilin, K. Xiao, and M. Gruteser. 2010. An energy-efficient mobile recommender system. In *Proceedings of the 16th International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 899–908.

[25] Y. Ge, C. Liu, H. Xiong, and J. Chen. 2011. A taxi business intelligence system. In *Proceedings of the 17th International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 735–738.

[26] Zi Yang, Xinpeng Wang, Xin Pei, Shuo Feng, Dajun Wang, Jianqiang Wang, and S. C. Wong. 2018. Longitudinal safety analysis for heterogeneous platoon of automated and human vehicles. In *Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC'18)*. IEEE, Los Alamitos, CA, 3300–3305.

[27] Xiuwen Yi, Junbo Zhang, Zhaoyuan Wang, Tianrui Li, and Yu Zheng. 2018. Deep distributed fusion network for air quality prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 965–973.

[28] J. Yuan, Y. Zheng, L. Zhang, X. Xie, and G. Sun. 2011. Where to find my next passenger. In *Proceedings of the 13th International Conference on Ubiquitous Computing*. ACM, New York, NY, 109–118.

[29] Zhuoning Yuan, Xun Zhou, and Tianbao Yang. 2018. Hetero-ConvLSTM: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 984–992.

[30] Brian D. Ziebart. 2010. *Modeling Purposeful Adaptive Behavior with the Principle of Maximum Causal Entropy*. Master's Thesis. Carnegie Mellon University.

[31] Brian D. Ziebart, Andrew L. Maas, J. Andrew Bagnell, and Anind K. Dey. 2008. Maximum entropy inverse reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'08)*, Vol. 8. 1433–1438.

Received March 2019; revised August 2019; accepted September 2019