Dissecting the Learning Curve of Taxi Drivers: A Data-Driven Approach

Menghai Pan[1], Yanhua Li[1], Xun Zhou[2], Zhenming Liu[3], Rui Song[4], Hui Lu[5], and Jun Luo[6]

[1]Worcester Polytechnic Institute, [2]University of Iowa, [3]College of William & Mary, [4]North Carolina State University, [5]Guangzhou University, [6]Lenovo Group Limited.

May 4, 2019

Acknowledgements

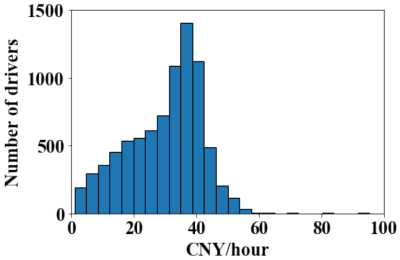
Collaborators:

Background

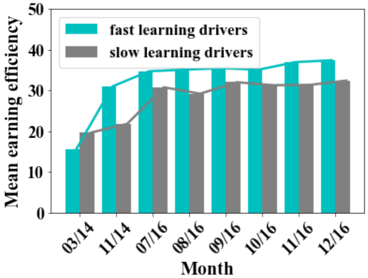
- Taxi service is a vital component in public transportation systems in both urban and suburban settings.

Observations

- The taxi operation efficiency differs significantly over different drivers.
- Some taxi drivers can learn faster than others.



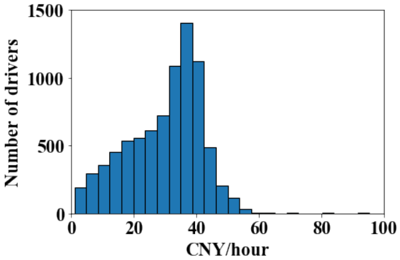
(a) The distribution of hourly earnings in July 2016. (Diversity)



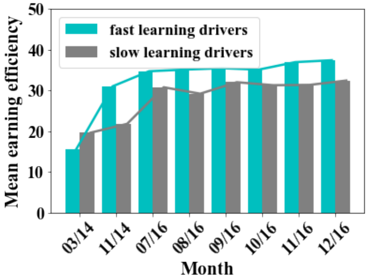
(b) Average hourly earnings over months. (Dynamics)

Observations

- The taxi operation efficiency differs significantly over different drivers.
- Some taxi drivers can learn faster than others.



(a) The distribution of hourly earnings in July 2016. (Diversity)

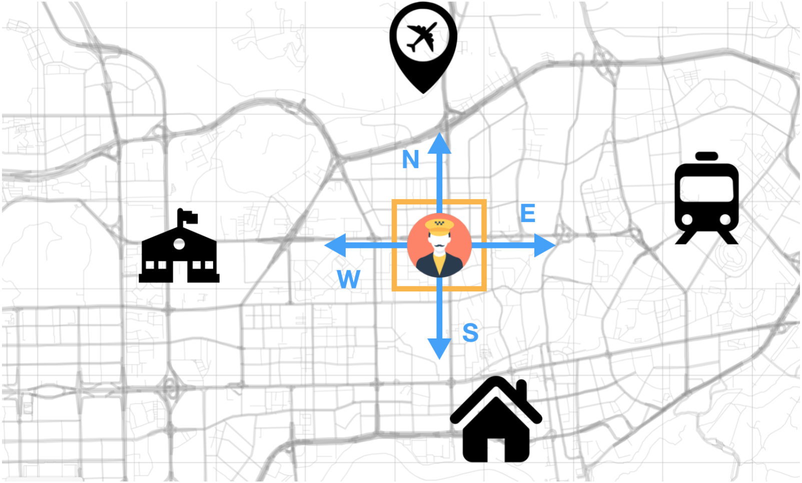


(b) Average hourly earnings over months. (Dynamics)

What did taxi drivers learn over time?

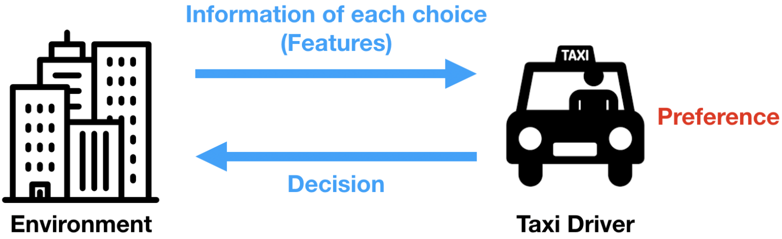
Background

Decisions that taxi drivers make when idle:



Taxi driver decision making process

- Taxi drivers make decisions based on their preferences to the features.

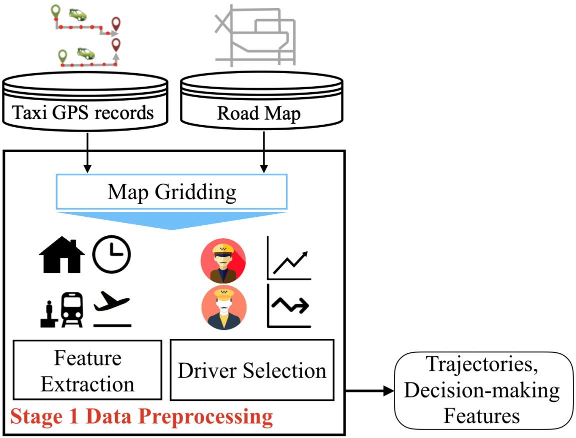


- Single step decision example:
  Features: [distance to home, time to finish, traffic condition, request likelihood]
    - $Preference_1$ = [high, high, low, low]
    - $Preference_2$ = [low, high, high, high]

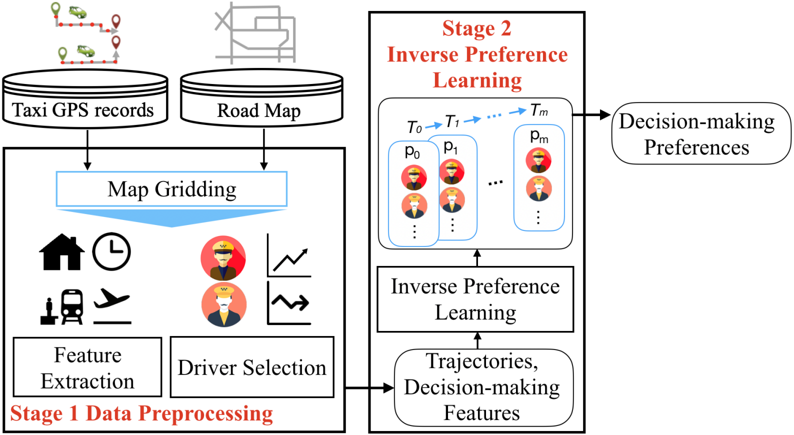
Questions to answer

- **Recover decision-making preferences of taxi drivers.**
- **Evaluate the preference dynamics over time.**

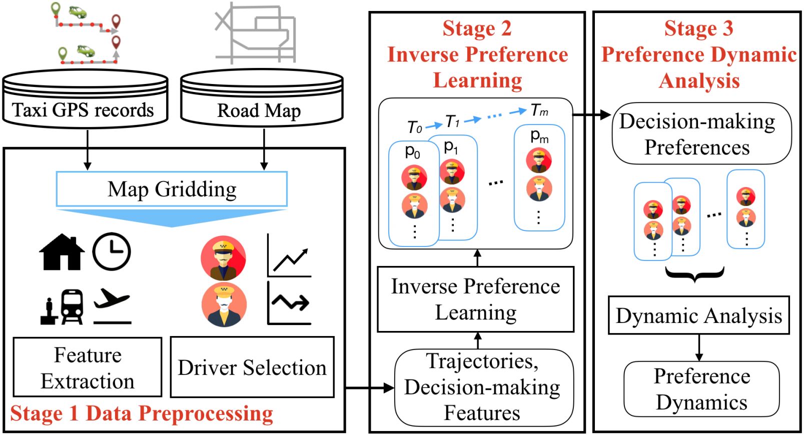
Solution framework



Solution framework



Solution framework



Stage 1 Data Preprocessing

Feature Extraction: each feature is defined as a numeric characteristic of a specific spatial-temporal region.

e.g., Distance to home and distance to train station.

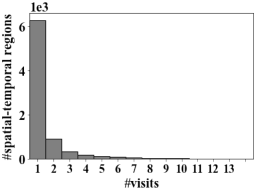
Stage 1 Data Preprocessing

Profile features

Related to the unique personal profile information.

- P1 Personal visitation frequency.
- P2 **Distance to home.**
- P3 Time from start.
- P4 Time to finish.

The distribution of visitation frequency to different regions of a driver:

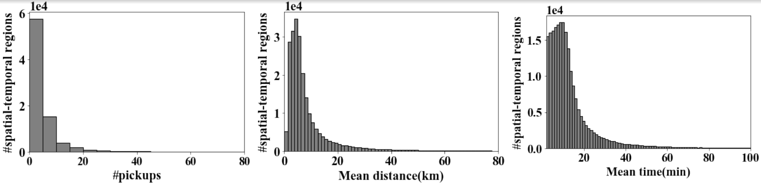


Stage 1 Data Preprocessing

Habit features

The objective characteristics which are the same for different drivers.

- H1 Number of pickups.
- H2 Average trip distance.
- H3 Average trip time.
- H4 Traffic condition.
- H5 **Distance to Shenzhen train station.**
- H6 Distance to Shenzhen Airport.



(a) H1: #pickups (b) H2: mean distance (c) H3: mean time

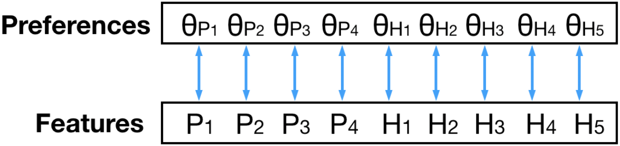
Stage 2: Inverse Preference Learning

Input

- Observed trajectories of a taxi driver.
- Decision-making Features.

Output

- The preference function of the driver regarding to the features.

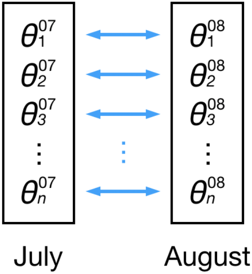


Method: Relative Entropy Inverse Reinforcement Learning.[1]

[1] A. Boularias, J. Kober, and J. Peters. Relative entropy inverse reinforcement learning. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pages 182189, 2011.

Stage 3: Preference Dynamic Analysis

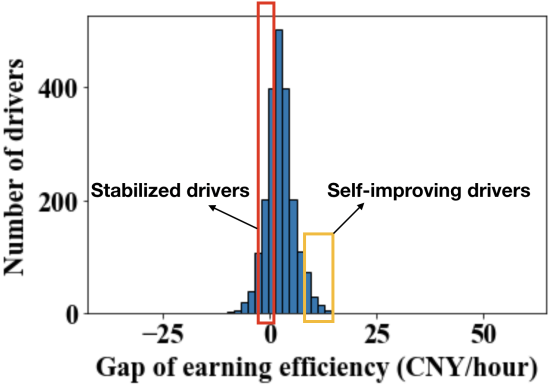
- Goal: examine if the change in each dimension of the preference vectors over time is significant or not.



- Method: Two-sample t-test.

Driver Group Selection

- 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.



Key Results for Group #1

t-values for Group #1:

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

$$\alpha = 0.05, \text{ fail: } |t| > 1.96$$

The significantly changed preferences comparing 12/16 with 07/16:

H1: Number of pickups

H3: average trip time

H4: traffic condition

Key Results for Group #2

t-values for Group #2:

Group #2 (Stabilized Drivers)

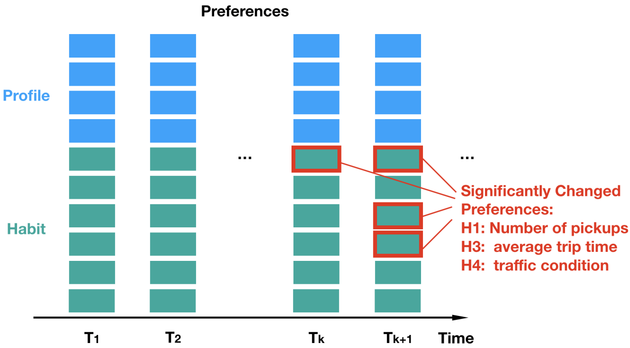
	P1	P2	P3	P4	H1	H2	H3	H4	H5	H6
Aug	-1.05	-1.23	1.30	-1.28	0.71	-0.66	-0.28	-1.94	-0.47	1.21
Sep	-0.08	-0.23	1.44	-0.85	0.09	0.83	-0.99	-0.98	-0.63	0.24
Oct	0.04	-1.74	1.87	-1.20	0.84	0.25	-0.63	-1.48	-0.19	-0.40
Nov	-0.62	0.77	-0.11	-0.95	1.05	-0.25	-1.50	-1.06	0.44	0.05
Dec	0.88	-1.13	1.89	0.36	-0.21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

$$\alpha = 0.05, \text{ fail: } |t| > 1.96$$

The preferences to all profile and habit features stay unchanged over the half a year.

Takeaways

1. Each driver has its unique preferences to their profile features, which tend to be stable over time.
2. Drivers while learning the environments, may change their preferences to habit features.



Open Source Code and We Publish Data

The code and data for inverse preference learning and preference dynamic analysis are published in the project website:

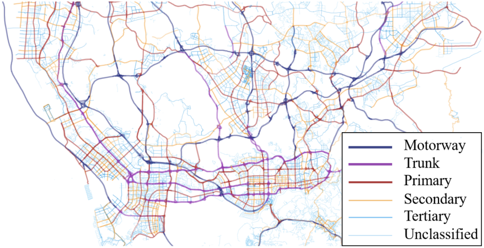
http://urban.cs.wpi.edu/DLCTD/.

Thank You!

Solution framework

Data description

- Taxi trajectory data.
  - **17877** taxis equipped with GPS sets.
  - Generate a GPS point every **40** seconds.
  - Time span: **2014-2016**.
  - Region: Shenzhen city.
- Road map data.



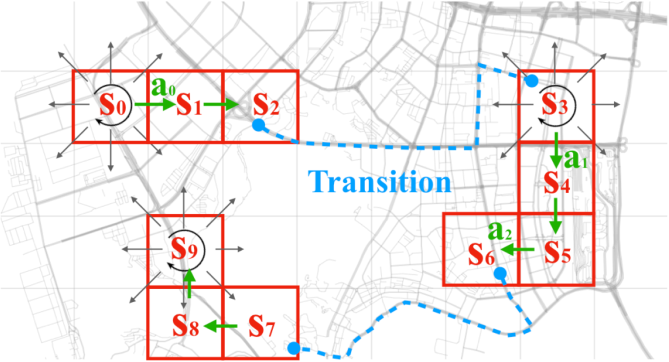
Stage 1 Data Preprocessing

Map and time quantization.

- 1158 valid cells with side-length of 0.01° in latitude and longitude.
- 288 time intervals per day.

MDP design

MDP of taxi driver's decision making process:



Stage 3: Preference Dynamic Analysis

Null hypothesis: the difference between the m-th entry of each θ_i^p in S_i and θ_j^p in S_j equals 0.

The t statistics of the m-th entry of the preferences can be calculated by:

$$t_{ij}(m) = \frac{Z}{s} = \frac{\Delta\bar{\theta}_{ij}(m) - \mu}{\delta/\sqrt{n}}. \tag{1}$$

$\Delta S_{ij} = \{\Delta\theta_{ij}^1, \Delta\theta_{ij}^2, ..., \Delta\theta_{ij}^n\} = \{\theta_i^1 - \theta_j^1, \theta_i^2 - \theta_j^2, ..., \theta_i^n - \theta_j^n\}$

REIRL problem

The REIRL aims to find a reward function θ, that

- minimizes the relative entropy,
- matches the trajectory distribution to the observed trajectory data.

$$\min_{\theta} : \quad H(P(\theta)\|Q) = \sum_{\tau \in \mathcal{T}} P(\tau|\theta) \ln \frac{P(\tau|\theta)}{Q(\tau)}, \tag{2}$$

$$\text{s.t.:} \quad |\sum_{\tau \in \mathcal{T}} P(\tau|\theta) f_i^{\tau} - \hat{f}_i| \leq \epsilon_i, \forall i \in \{1, ..., k\}, \tag{3}$$

$$\sum_{\tau \in \mathcal{T}} P(\tau|\theta) = 1, \tag{4}$$

$$P(\tau|\theta) \geq 0, \qquad \forall \tau \in \mathcal{T}, \tag{5}$$