On Demand Transportation Scheduling

Stanford University

Weini Yu, Zhouchangwan Yu, Yutian Li

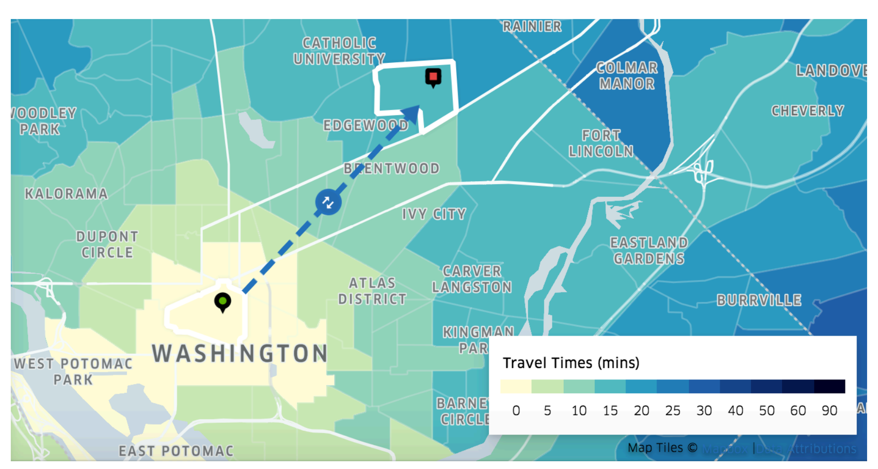
Introduction

Motivation: better matching needed for riders (demand) and drivers (supply) in on demand transportation service platforms like Uber and Lyft.

Goal: develop an intelligent model that effectively match the customer requests with the drivers in the single vehicle pickup and delivery problem (SVPDP).

Data Acquisition

Map of the city of Washington D.C. discretized into zones, from Uber Movement



Our model relies on the following types of data, all gathered from Uber APIs.

- **Location Coordinates**: arithmetic mean of the zone coordinates.
- **Travel Time**: average travel time from one zone to another
- **Fares**: estimated fare from one zone to another

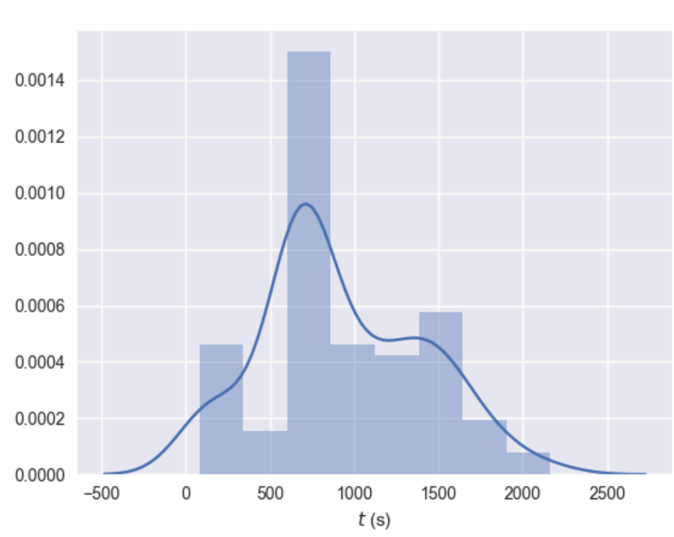


Fig1. distribution of traveling time

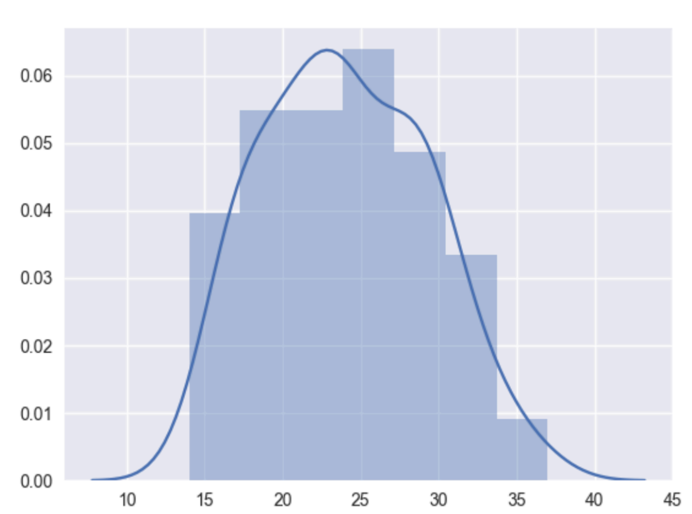


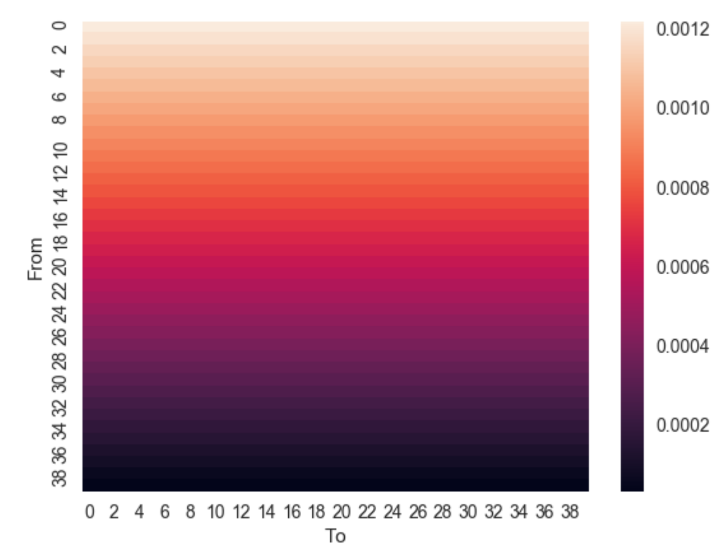
Fig 2. distribution of estimated fare

Method

Model: Markov Decision Process

State: Driver's current location
Action: (source, destination) tuple where source is the up zone index and destination is the drop off zone inde
Reward: fare − α · travel time
Generation of requests:

- Source: lower indices, higher probability.
- Destination: uniformly randomly
- Favor locations with smaller indices and create an imbalance between states, making some inherently better.



Algorithm: Q-learning

Features

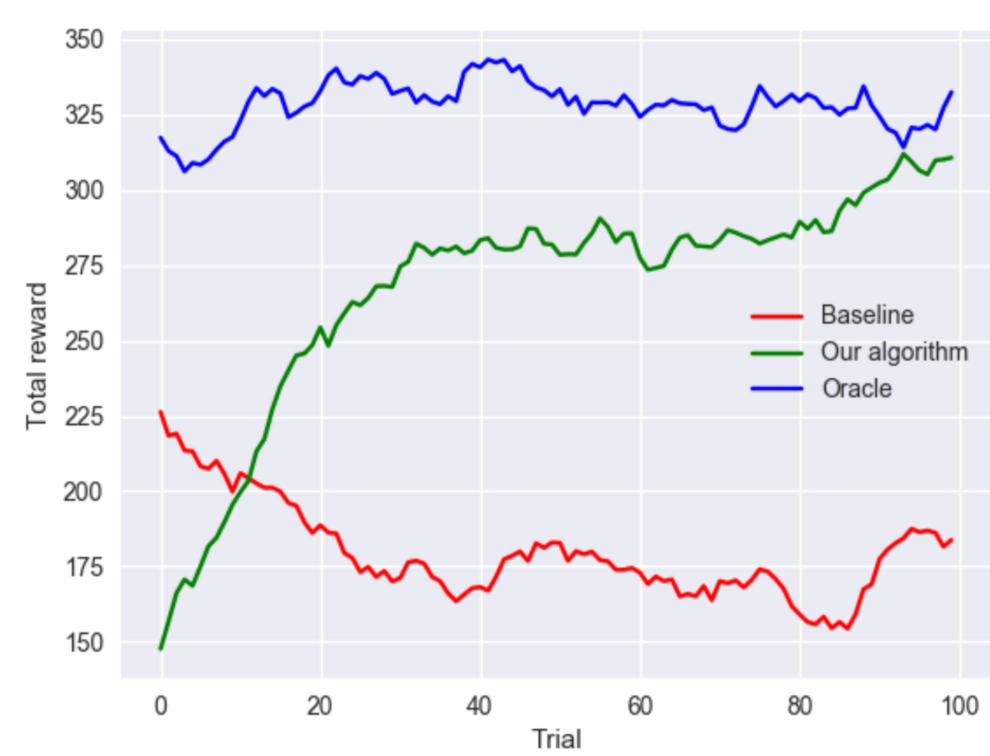
- Travel distance: from current location to source, and from source location to destination
- Zone: indicator features of current location, source, and destination

Training

- S.G.D. using exponential weighted moving average
  $w_i \leftarrow w_i - \eta(Q(s,a) - r)\phi_i(s,a)$
- discount γ = 0, because driver unaware of legal actions at future states.

Results and Analysis

Baseline: driver randomly picks one request at each step.
Oracle: driver knows true reward function for all requests, and picks greedily
Our Results:



Total rewards:
Baseline: converges to 180
Oracle: converges to 330
Our algorithm: 315

As training goes on, our algorithm is slowly approaching the oracle results, **75%** higher than the baseline reward and only **5.5%** lower than the oracle reward.

Achieved good results, almost converging to the oracle.

- Q-learning is well suited for the task
  - state does not matter much because the driver may return to the same spot with drastically different accumulated rewards.
  - state-action pair conveys much more information.
- Relative distance as a feature captures the correlation of distance, fare, and travel time well.

Future Work

- Rather than selecting the first 40 zones in the data set, randomly sample 40 zones to reduce similarity between zones
- Define more advanced features
- Consider other non-zero values for discount γ
- Explore Neural Network instead of handcrafting linear features
- Expand the model to accommodate carpool situation
- Compare global optimum strategy with Nash Equilibrium

References

[1] M. Hosny, C. Mumford, *The single vehicle pickup and delivery problem with time windows: intelligent operators for heuristic and metaheuristic algorithms*. Journal of Heuristics, 2008.
[2] J. Mrkos, et.al., Liftago On-Demand Transport Dataset and Market Formation Algorithm Based on Machine Learning, 2016.
[3] M.Fogleman. Ridesharing Algorithms in TransLoc OnDemand.