Reinforcement Learning for Taxi Driver Re-positioning Problem in NYC

Tian Wang, Yingyu Cao, Bo Jumrustanasan, Tianyi Wang, Xue Xia

December 10, 2020

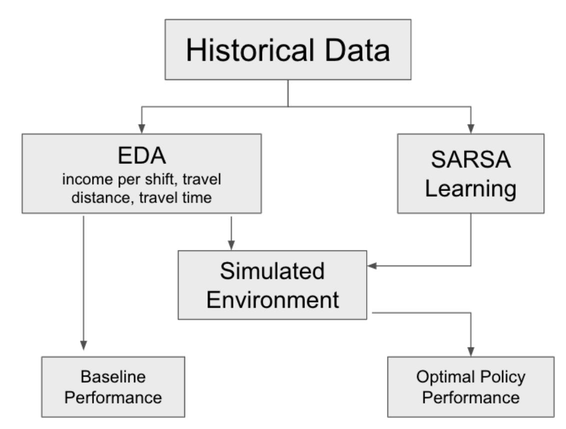
Data Science Institute @ Columbia University

Motivation

Goal
Utilizing Data Science to optimize drivers' decision making

Problem: Reposition Strategy

Project Overview:



Demand for improving traffic efficiency is increasing

- **Platform**: Supply-demand balance
    - Rise of for-hire vehicle (FHV) companies such as Uber and Lyft
- **City**: Traffic congestion
    - From 6.1 mph in 2010 to 4.3 mph in 2018
- **Driver**: Income
    - 41% time seeking passengers

Background
Drivers' decision making is affecting their earning

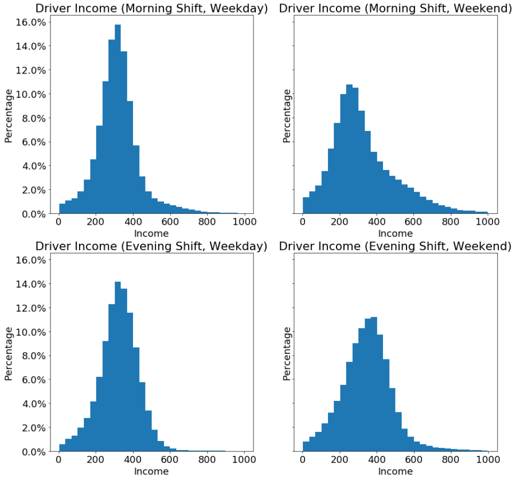


Figure 1: Driver Income Distribution

- Working shifts
- Starting time
- Re-positioning
- ...

Data

Data Schema

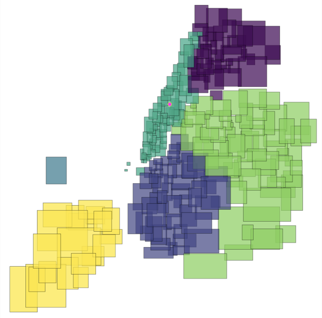
- Dataset: New York City Taxi Trip Data (2010-2013)
- Features used:
  - hack license (Driver ID)
  - pickup datetime
  - dropoff datetime
  - trip time in secs
  - trip distance
  - trip fare
  - pickup longitude & pickup latitude
  - dropoff longitude & dropoff latitude

Data Preprocessing

- Sampling: Yellow Taxi trip records in June 2013
- Removal of abnormal data: trips with
  - traveling time < 1 min or > 3h
  - average speed > 50 mph
  - distance > 30 miles
  - trip fare $> \$150$
  - pick-up or drop-off location not in NYC

Data Preprocessing

- Each position coordinate is assigned to a taxi zone.
  - Speed up the process using bounding box and R-tree algorithm.
- Driver trajectory is recovered based on the shortest path weighted on distance between taxi zones.



Methodology

Key Assumptions

1. Coordinates are discretized into taxi zones.
2. Time is discretized into time intervals Δt.
3. There is only one driver following the optimized policy the model derives, i.e. one agent.
4. Drivers can cruise to other taxi zones without taking any orders in current taxi zone (drivers can refuse rides).

Environment Setup

1. Agent: One single driver
2. State space: $\text{state} = (\text{location}, \text{time})$
3. Action space: $a \in \{0, 1, \ldots, n_{\text{taxi\_zones}}\}$

Environment Setup

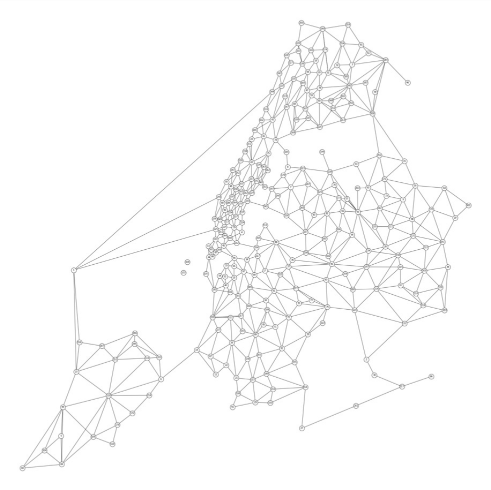


Figure 2: Environment

Training Algorithm: SARSA

Initialize $Q(s, a)$ arbitrarily.
for Each observed episode **do**
    for Each step in the episode **do**
        Observe S, A, R, S', A'
        $\delta = R + \gamma * Q(S', A') - Q(S, A)$
        $Q(S, A) = Q(S, A) + \alpha * \delta$
    end for
end for

Experiment

Evaluation - Model Convergence

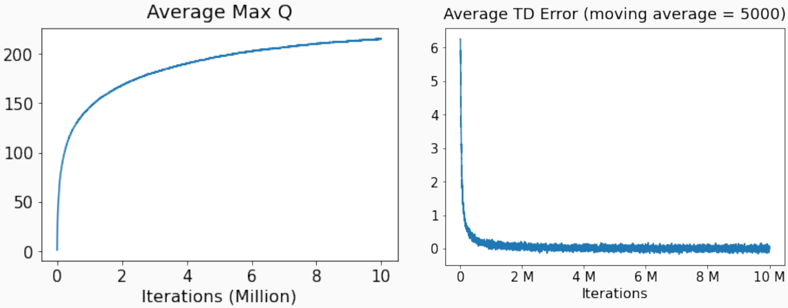


Figure 3: Day Shift Model Convergence

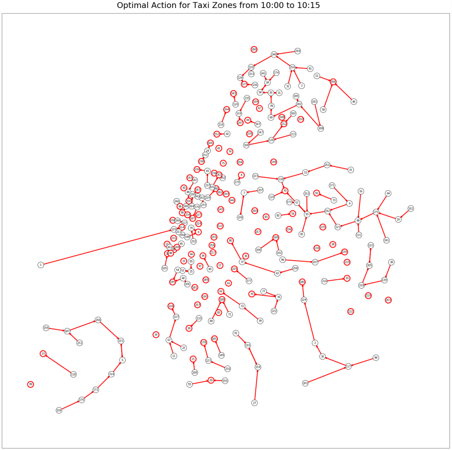
Evaluation - Strategy Performance

Earning and utilization rate from the derived policies will be compared against the historical data below.

	Day Shift	Night Shift
Earning per Shift	$321	$332
Utilization Rate	54.0%	53.8%

Table 1: Empirical Performance

Illustration of optimized policy



- Percentage of zones that travel within the same zone is much higher during rush hours, especially in Manhattan and Brooklyn areas.
- For zones that travel to adjacent zone, we can detect flow of traffic in the same direction and some popular destinations.

Conclusion & Future Work

Conclusion

- Simulated environment
    - Distance, traveling time, cruising time, etc.
- Optimized policy
    - Identify heat spots for different regions

Future Work

- Simulated Environment
- Learning Algorithm
- Data Sampling

Thanks!

For more information, you can visit our Github

Appendix

Appendix

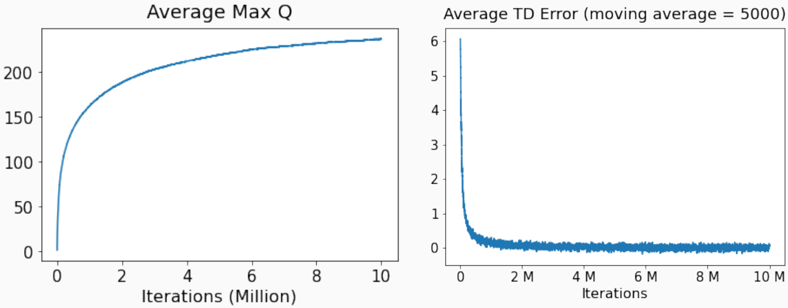


Figure 4: Night Shift Model Convergence