COLUMBIA UNIVERSITY
Data Science Institute

ENGI E4800 - Data Science Capstone & Ethics

Progress Report I – DiDi Reinforcement Learning

Danyang Han (dh2990), Anita Pinto (ap3650), Elizabeth Yum (eyy2106)

Mentors: Zhiwei (Tony) Qin and Vineet Goyal

February 27, 2021

Abstract

We propose a novel method to tackle the problem of driver repositioning on ride-hailing platforms through off-policy evaluation by the creation of an environment simulator. First, we propose to create an environment simulator using historical data that stochastically generates demand, supply, passenger/driver matching and destination probability. Next, the environment simulator may be leveraged for off-policy evaluation through Monte Carlo rollout or Dynamic Programming. Finally, we propose that a generalized policy iteration process may be applied for off-policy learning. This progress report details the problem being solved, overall framework, data sources, and initial progress toward this goal.

I. Introduction

Many cities increasingly rely on for-hire vehicle (FHV) services such as DiDi, Uber, Lyft, Juno, and Via as a critical part of their transportation infrastructure. In large, highly competitive markets such as New York City, taxi drivers' incomes are highly dependent on their ability to be able to find their next passenger quickly. Taxi drivers' daily incomes can vary widely depending on their working behaviors and preferences – when and where to start a shift and their repositioning behavior between rides.

Previous partnerships with DiDi and Columbia's Data Science Institute have focused on pattern recognition of factors that lead to driver performance and reinforcement learning to minimize driver idle time and maximize driver income based on historical data. In fall 2019, a previous Capstone project aimed to identify patterns that contribute to driver succession DiDi's ride hailing platform through a random forest model that predicts a driver's Performance Ratio. In Fall 2020, another Capstone project applied reinforcement learning to New York City taxi data to aim to find the best spatiotemporal policies to maximize driver income.

In contrast to previous iterations of projects have been constrained to model learning based on observed, on-policy data, this new iteration of the project aims to use historical New York City Taxi & Limousine Company (TLC) data to create a full simulation environment to enable off-policy evaluation for new repositioning policies and generalized policy iteration. The following sections outline our proposed methodology, problems encountered and solutions being pursued, our initial progress, and our next steps.

II. Methodology and Data Sources

Figure 1 below depicts at a high level the proposed overall process. First, historical data is leveraged to help create an environment simulator that generates demand, supply, matching probability and destination probability by taxi zone (a set of 263 areas tracked by the TLC roughly based on NYC Department of City Planning's Neighborhood Tabulation Areas which are meant to approximate neighborhoods) and time.

The most challenging piece of the environment simulator is the demand generator. We are restricted to use of NYC TLC publicly available data which is confined solely to observed, captured demand. However, successful implementation of a full environment simulator hinges on the ability to estimate total demand including both captured and uncaptured, and therefore unobserved data. Section IV, Environment Simulator – Current Approach & Progress, discusses in further detail approaches currently being explored for elements of the environment simulator.

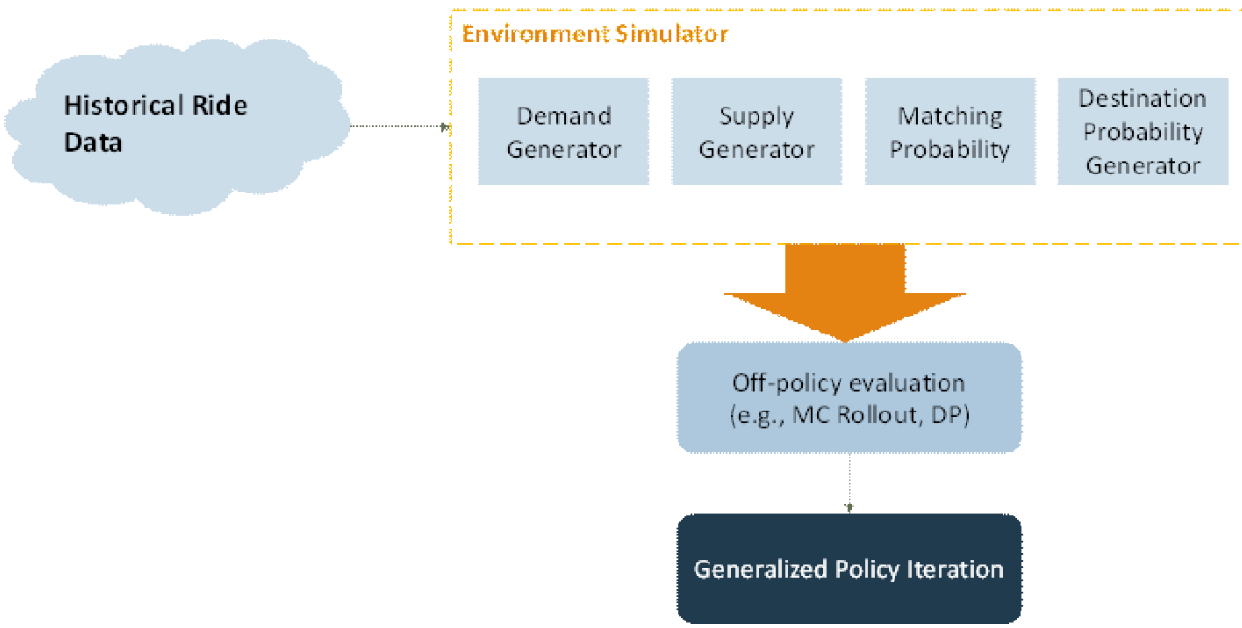

Figure 1: Project Methodology

For historical ride data, we are leveraging data from the NYC TLC, the agency responsible for regulating New York City's medallion (yellow) taxis, street hail livery (green) taxis, for-hire vehicles (FHVs), commuter vans, and paratransit vehicles. The TLC has collects and publishes trip data for each taxi and for-hire vehicle trip completed by licensed drivers and vehicles on a monthly basis.

For the purposes of this study, we are solely leveraging Yellow Taxi trip data from June 2013, due primarily to barriers to using broader or more recent data. We are constraining the data solely to Yellow Taxi data as Yellow taxis are the only vehicles permitted to respond to a street hail from a passenger in all five boroughs, and were the dominant mode of taxi & limousine transportation in 2013. As such, the observable behavior in Yellow Taxi data is unconstrained compared to that of Green Taxis (which may only respond to street hails in select locations), and captures nearly all of the passenger ride services trips in 2013. We are limiting ourselves to data from 2013 as historical Yellow Taxi data included individual Taxi identifiers which enables identification of individual driver behavior, movements, and outcomes, which more recent Yellow Taxi data lacks. Additionally, the rise and increasing popularity of ride hailing apps means that a significant proportion of ridership in recent years is through FHV services which are subject to lower data reporting standards and missing critical data fields compared to Yellow Taxi and Green Taxi services.

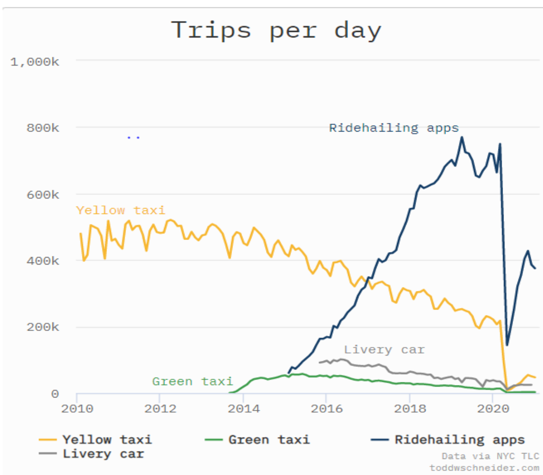

Figure 2: NYC Taxi Ridership from 2010-2020
Source: Schneider, Todd. "Taxi and Ridehailing Usage in New York City", Accessed: Feb. 26, 2021, https://toddwschneider.com/dashboards/nyc-taxi-ridehailing-uber-lyft-data/

The key features of the dataset and a short description of the transformations of the dataset undertaken are listed below.

Features	Description
Medallion	Taxi medallion, also known as a CPNC (Certificate of Public Necessity and Convenience). This can be treated as the unique ID for the vehicle.
hack_license	New York hack license to drive a yellow taxi. This can be treated as the unique ID for the driver.
vendor_id	A code indicating the TPEP provider that provided the record. 1= Creative Mobile Technologies, LLC 2= VeriFone Inc.
rate_code	The final rate code in effect at the end of the trip. 1= Standard rate, 2=JFK, 3=Newark, 4=Nassau or Westchester, 5=Negotiated fare, 6=Group ride
store_and_fwd_flag	This flag indicates whether the trip record was held in vehicle memory before sending to the vendor, aka "store and forward," because the vehicle did not have a connection to the server. Y= store and forward trip

	N= not a store and forward trip
pickup_datetime	The date and time when the meter was engaged.
dropoff_datetime	The date and time when the meter was disengaged.
passenger_count	The number of passengers in the vehicle. This is a driver-entered value.
trip_time_in_secs	The length of trip time in seconds.
trip_distance	The elapsed trip distance in miles.
pickup_longitude	Longitude where the meter was engaged.
pickup_latitude	Latitude where the meter was engaged.
dropoff_longitude	Longitude where the meter was disengaged.
dropoff_latitude	Latitude where the meter was disengaged.

Table 1: Dataset Features

Similar to work done in the prior iteration of the capstone, the following data cleaning steps were applied to remove abnormal data:

- Trip time: the serving time of a single trip. We drop trips with traveling time that is less than 1 minute or longer than 3 hours as most of these records have extremely short or long travelling time.
- Trip distance: the travelling distance of a single trip. We calculated the average speed of each trip based on trip distance and trip time, and drop records that have speed over 50 mph, which is the highest speed limit in New York[1]. Trips that are longer than 30 miles are also removed because they are about twice as long as the length of Manhattan[2].
- Trip fare: the total money a driver earns from a single trip. We drop records with trip fare that is higher than $150, which is the trip fare of travelling 60 miles.
- Pick-up and drop-off locations. We drop the trips whose pick-up or drop-off location were not in NYC.

The total amount of data removed from the data cleaning steps applied above was 3% of all trip data (499,717 of 14,385,456 trips were removed leaving 13,885,739 total trips available for analysis), and we do not believe the data cleaning steps above introduces any undue bias. In addition, we also similarly applied a transformation to map pick-up and drop-off coordinates to the TLC's 263 taxi zones.

In addition, for simulating demand, we also applied the following transformations.

- For calculating arrival rate by time zone, we assumed 24 hour days were split by an initial time interval of 30 minutes. Each day was divided into 48 30-minute time intervals, and for each taxi-zone, the average pick-up was applied to approximate the λ parameter for the Poisson distribution.
- For calculating inter-pick up times and the frequency of drop-offs within each taxi zone, computed the inter-pickup time as the difference between the prior trip's drop off time and the next trip's pick up time and the number of drop offs within each 15 minute interval.

[1] https://www1.nyc.gov/html/dot/downloads/pdf/current-pre-vision-zerospeed-limit-maps.pdf
[2] https://www.nycgo.com/plan-your-trip/basic-information/

Please see the following section for more information and preliminary data visualization for these transformations.

III. Existing Literature

To our knowledge, no equivalent research has been published that aims to create an off-policy evaluation and generalized policy iteration process through the creation of an environment simulator.

Research on optimal routing of a vacant taxi includes work from Yu et al. (2018) proposing a process of expressing the routing problem as a Markov Decision Process. States were defined by the node at which a vacant taxi is located and actions were taken as the link to take out of the node. State transition probabilities were expressed as passenger matching probabilities and passenger destination probabilities, and the MDP problem was solved through value iteration. In contrast to work done by Yu et al., our process proposes to create a sophisticated method of demand estimation that captures not just observed demand, but seeks to capture total demand.

As such, the most critical piece of the environment simulator is an effective way of generating demand solely on historical observations which are effectively censored as they cannot reflect unfulfilled demand. Research on censored demand that tackles the newsvendor problem may have some parallels. In the newsvendor problem, a newsvendor seeks to optimize stock to maximize profit under fluctuating demand. Research from Mersereau (2015) approached the problem of estimating sales when sales are a censored representation of the underlying demand process and the impact of inventory record inaccuracy in an environment with censored observations.

IV. Environment Simulator – Current Approach & Progress

Demand Generator

As the most critical piece of the environment simulator, demand generation has a number of challenges we are attempting to work through:

Challenge #1: Demand highly sensitive to zone, time, and day

> As illustrated in Figure 3 below, demand varies highly by zones and by day of the week. Additionally, while not explicitly depicted below, demand is also highly sensitive to time with demand peaking for most zones within the typical "rush hours" (early AM and early evening coinciding with typical work schedules). While not explicitly a methodological challenge, any approach to demand estimation will need to be appropriately sensitive to the differences in demand across zones, times and days.

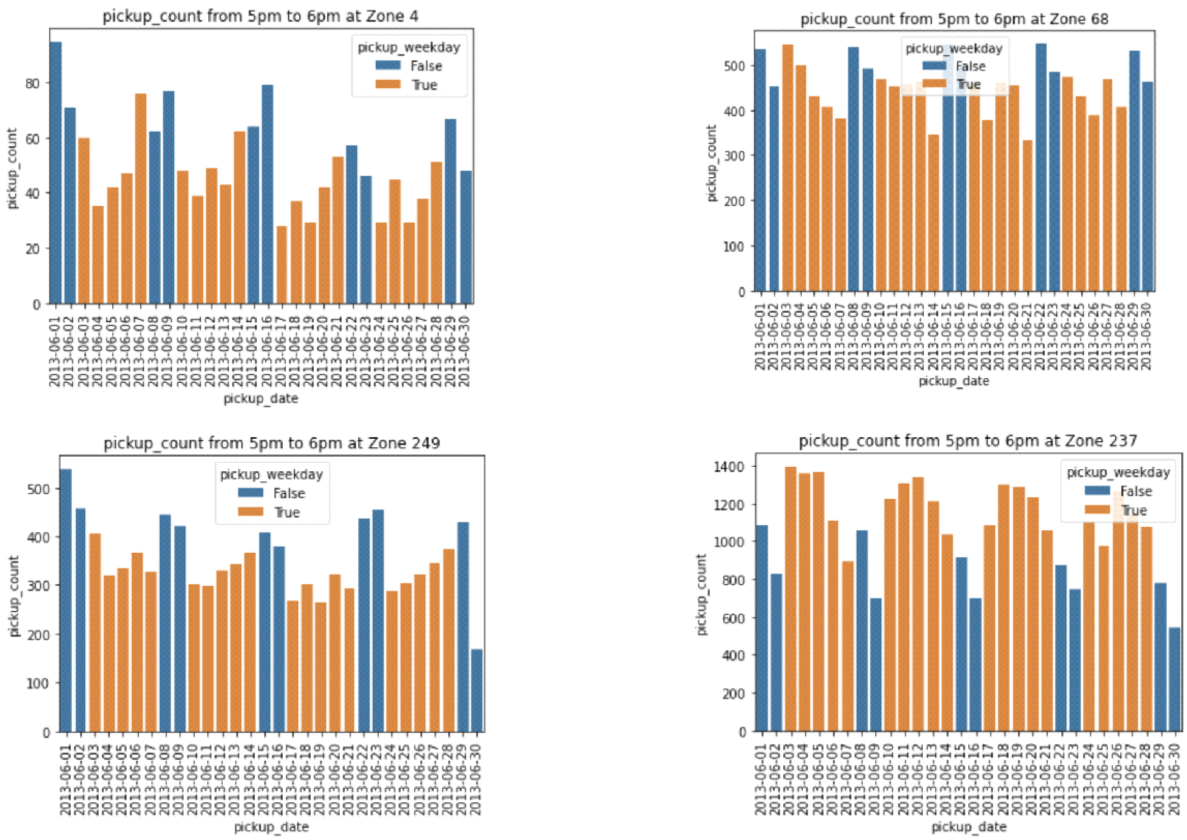

Figure 3: Pick-up Count From 5-6 PM Sampled Across 4 Zones

As an initial starting point, we have done preliminary data processing to calculate arrival rate by time zone. Each day was split into forty-eight 30 min intervals. For each of the taxi-zones, using the given clean data, the average pick up at a given time zone was calculated and taken to be λ parameter for the poisson distribution estimating the probability of arrival of passenger at a given ∆t to generate a table of 263x48 λ parameters, however, this process will need to refined and adapted based on how we approach the other challenges.

<u>Challenge #2</u>: Historic data only contains observed demand

As noted previously, we are attempting to adapt solutions applied in the newsvendor problems to censored taxi demand. Please see below for a continued discussion of how we are approaching this issue.

<u>Challenge #3</u>: In attempting to adapt the approach for censored demand in the newsvendor problem, we must be able to identify a proxy to a true “stock-out” indicator

In contrast to the newsvendor problem, directly observable data from NYC Yellow Taxi data lacks the ability to directly observe when demand exceeds available supply. In the newsvendor problem, seller have history about sales (s_t) and stock (j_t), and also record a stockout indicator $c_t = 1$ when $s_t = j_t$, and $c_t = 0$ when $s_t < j_t$. In contrast, for Yellow Taxi data, which are continuously in supply, it is difficult to approximate when demand has been lost. Potential solutions discussed for addressing the lack of a true stockout indicator are as follows:

- "Price surge" – We explored existing data to look at fare patterns over 15 minute intervals in the figure below. However, as Yellow Taxis go by rates which do not have a directly observable price surge indicators for periods of high demand as many FHV services would have, variation was limited.

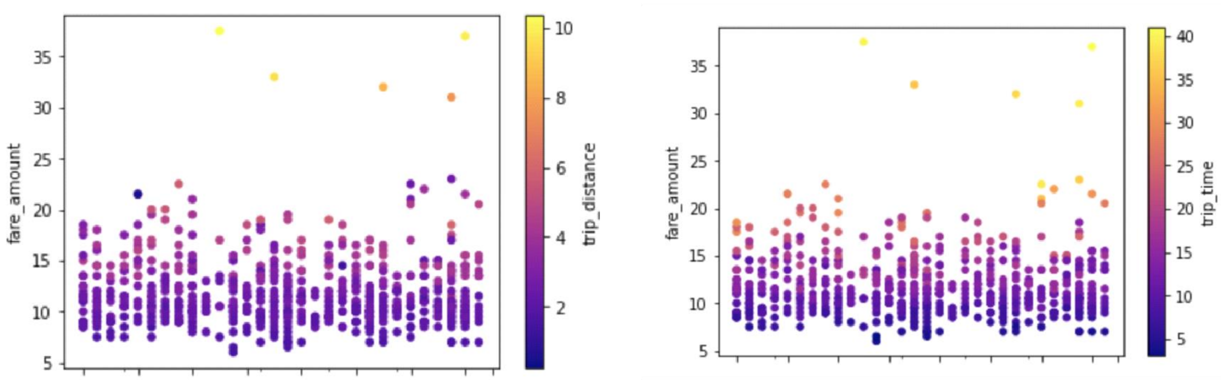

Figure 4: From taxi zone 148 to taxi zone 66 in 15 min time intervals

- Net taxis in zone – Calculated as #(taxis that dropped off in t-1) - #(taxis that picked up in t-1). This option was discussed, but it would need to customize by state and by zone, and it was unclear what the cut-off point should be
- Taxi utilization or % of available taxis in t-1 that picked up in time t. Again it was unclear what the cut-off should be, low utilization places are possibly the sources of highest lost demand (e.g., queens)
- Inter-pickup times – The approach we are currently pursuing is to detect the presence of lost demand based on inter-pickup times, or the time between when a taxi has dropped off a passenger and picks up another. Shorter inter-pickup times would signal a period of high demand when taxi utilization is high, while longer inter-pickup times would indicate a period of low demand. The figures below show a sample of this is being approached with respect to a zone that is broadly representative (zone 196, a zone that regularly experiences a moderate amount of activity).

  In addition, inter-pickup times might undergo further transformation (e.g., transforming this into the ratio of pick up interval to the number of drop offs) to normalize for periods of high activity

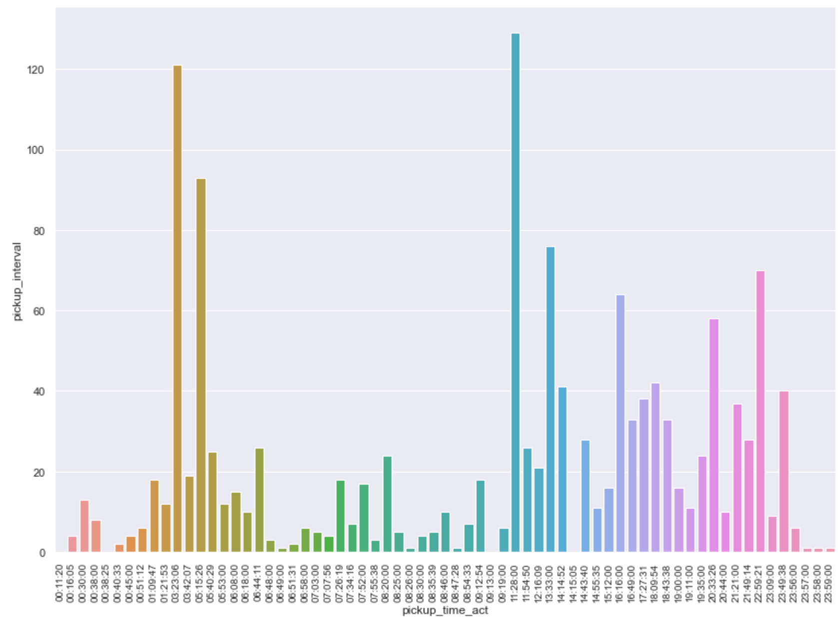

Figure 5: Interval between pickups – Taxi Zone 196

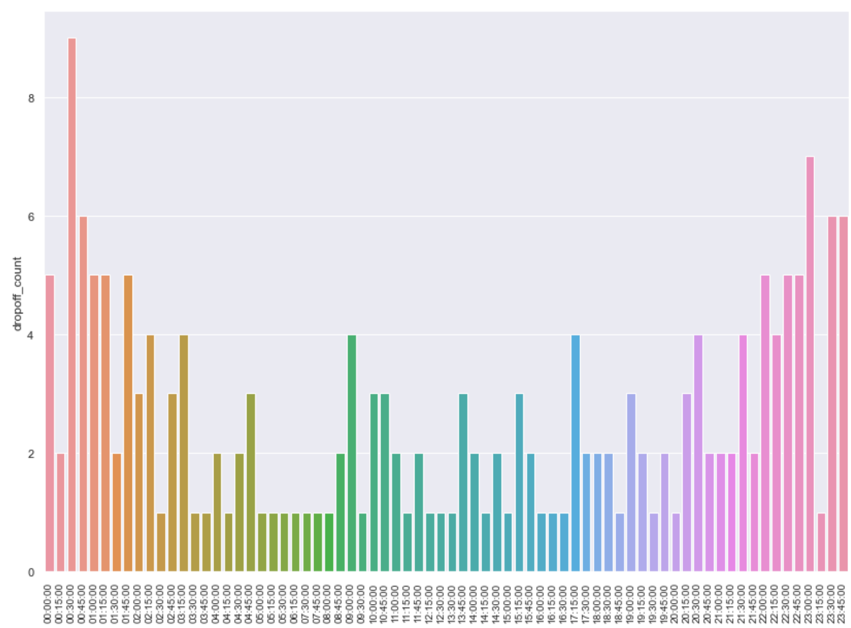

Figure 6: Number of drop-offs within 15 minute intervals – Taxi Zone 196

Work on selecting the appropriate indicator, deciding upon an appropriate cut-off value, and then adapting the indicator to our model is ongoing. We will continue to explore what the appropriate proxy value for indicating a “stockout” event should be and will continue to explore the appropriate method of then translating estimated demand from historical data to total estimated demand (e.g., scaling demand up with an average demand amount if stockout has occurred, etc.).

Supply Generator

Initial distribution of taxi’s will be generated randomly based on historical data (including information on when taxis typically start and end shifts). The movement and therefore supply of taxis at subsequent times and zones will be an outcome of the model’s environment simulator.

Matching Probability

Similarly to the approach in Yu et al., we will be implementing a matching probability that is separate from the demand probability. The following is an outline preliminary approach we propose.

- **Calculating the probability of at least one pick up at taxi zone at Δt**
  - The probability of arrival of at least one passenger at a given location is assumed to be based on a poisson distribution with parameter λ (where λ is the average rate of pick-ups at a taxi zone in that time interval Δt (30 min)
  - Pr(at least one passenger) = 1 - Prob(no passengers)
  - Pr(arrival of K passengers) = $\frac{\lambda^k e^{-\lambda}}{k!}$
  - Pr(at least one passenger) = 1 - exp(-λ) [k =0]

- **Calculating the probability of nearest taxi to a pick up taxi zone**
  - First calculate λ_d, the number of drop-offs around the taxi zone for that time interval Δt (30 min)
  - Use the adjacent zone calculator using geoshapes.
  - Calculate P_d being the probability of nearest taxi to a pick up zone, by calculating the probability of no taxis in that zone and its adjacent taxi zone.

    P_d [X(S)=0] = exp($-\lambda_d * r$)

  - So P_d would be a higher probability depending on whether the number of drop-offs around a pick-up zone are fewer
  - r is the indicator of the adjacent zone level, for this research, we only consider the vacant taxis in the immediate adjacent zone of the pick up zone. so value of r would be 1

- Ideally r should be the inter-zone distance between adjacent taxi and the pick-up zone, which would give a better indication of the nearest vacant taxi to a pick-up zone. However all our current calculations are based on taxi zones rather than using actual geo-locations.
- Alternatively, if we dont use a parametric distribution, we can also derive an indication of being the nearest taxi based on the following calculation :
- Prob(of being nearest taxi) = $1/\lambda_d$
- If λ_d is large, it means lower probability, there are other drivers who have done drop offs in nearby zones, this gives an idea of number of vacant taxis near that taxi zone and adds to the prob of matching

- **Probability of arrival at zone h earlier than all the other adjacent nodes**
  - Calculate λ_h which would be arrival rate at a pick up zone in Δt (30 min)
  - Calculate $\lambda_{N(j)}$ would be the arrival rate at respective adjacent nodes.
  - The probability that an arrival from node h is earlier than all other nodes is

    $\lambda_h / \lambda_{N(j)}$
  - Since $\lambda_{N(j)}$ would be fixed for a given time interval, this probability matching calculation would ensure to drive the policy to choose a pickup zone with a higher number of arrival rates.

The matching probability calculation is highly dependent on the demand generation. After the demand generation process is settled, the matching probability process will be further developed further as we explore the necessity of adjacent zones (and measuring proximity) and the ideal calculation of the figure r.

Destination Probability Generator

We will follow an approach similar to that of Yu et al., but may undertake modifications for how destination varies within a zone based on the day of the week and time of day. For example, the probability a passenger picked up at zone x having zone y as the destination can be approximate by the observed fraction of passengers picked up at zone x going to y at time t.

V. Next Steps

Our highest priority is refining and finalizing our demand generation process. We will continue to assess appropriate indicators for stockout, and create a process to adjust demand upward for zones and times when stockout has occurred. From there, the demand generator can be leveraged to make significant progress on other elements of the environment simulator. Our goal is to have made significant progress on the environment simulator by April 3rd for the Second Progress Report. This will then enable time for finalization of the environment simulator and off-policy evaluation and iteration for the final report. In addition, based on progress, we have the goal of submitting our process and results for consideration for potential upcoming conferences at the conclusion of this project.

Appendix

References

[1] Xinlian Yu, Song Gao, Xianbiao Hu, Hyoshin Park, A Markov decision process approach to vacant taxi routing with e-hailing, Transportation Research Part B: Methodological, Volume 121, 2019, Pages 114-134, ISSN 0191-2615, https://doi.org/10.1016/j.trb.2018.12.013.

[2] Mersereau, Adam J. "Demand estimation from censored observations with inventory record inaccuracy." Manufacturing & Service Operations Management, vol. 17, no. 3, 2015, p. 335, ISSN: 1523-4614, https://doi-org.ezproxy.cul.columbia.edu/10.1287/msom.2015.0520

Contributions

- Danyang Han: Main contributor on data transformation, processing data for arrival rates by time zone, and testing proposed models for Poisson demand estimation
- Anita Pinto: Main contributor on cleaning and preprocessing data, data visualization, formulating and testing models for probability matching calculation, inter-pickup time arrivals and lost demand estimation.
- Elizabeth Yum: Main contributor on organizing materials for meetings, setting up logistics, and managing progress. Reviewed additional research and literature of other potential models, collating and documentation of the results and reports.