1. MINIMIZING NUMBER OF TAXIS

1.1. Overview

In this section we find the minimum number of taxis in Ithaca, subject to any number, n, of constraints of the following form: less than $x(100)\%$ of customers end up waiting for more than y minutes for their taxi. Since we have n constraints we have n values for x and n values for y. Denote these values as $x_1, ..., x_n$ and $y_1, ...y_n$. We solve this optimization problem by finding the following function

$$p(w, t) : \mathbb{R} \times \mathbb{N} \to [0, 1] \tag{1}$$

$p(w, t)$ is the proportion of customers who have to wait for more than w minutes, given that we have t taxis operating. We then look at the sequence of functions $p_{1 \leq i \leq n}(y, t)$, and find the minimum t such that $p_{1 \leq i \leq n}(y_i, t) < x_i$. Denote this value as t_i^*. Then the minimum number of taxis needed, t^*, subject to our constraints is

$$t^* = \max_i t_i^* \tag{2}$$

1.2. Graph

Our approach in answering this question is by viewing Ithaca as a complete directed graph. We split Ithaca into 20 regions, and in our model we treat each of the 20 regions as a point. The edge between any two nodes in our graph, A and B contains two values, the amount of time it takes to get from point A to point B, and the probability of having a customer need a taxi ride from node A to Node B in a given minute.

The table of nodes we treat in Ithaca is in table I. Note that the "Zone #" are the corresponding zone numbers listed on Mythical Dispatches' website.

Zone 1	Zone 6	Eastern Heights	Lansing
Zone 2 (Contains Begmen's)	**Airport**	South Hill	Cayuga Heights
Zone 3 (Contains the Dommons)	**Cayuga Mall**	**Cayuga Medical**	Allan H. Treman St. Pk
Zone 4	Ithaca College	N.E. Ithaca	Taughannock Pk
Zone 5	Buttermilk St. Pk.	Stewart Park	West Hill

TABLE I: Nodes of our graph for question 1. Nodes in blue are residential, nodes bolded are highly areas travelled frequently by taxis

1.2.1. *Time From Edges*

Denote this by the adjacency matrix A. We assume that it is symmetric, meaning the time to get from point a to point b is the same there and back. In order to determine the approximate times from node to node, we use the data provided by Google Map and measure the distance between each node.

In addition, since some of the regions are intended to be a neighborhood, there are a distance associated with traveling within the node. Certain nodes do not contain this property, such as the airport, as it doesn't make intuitive sense.

1.2.2. *Probabilities*

Deirdres Write up

TODO(me): Edit to mention difference in weekday, weekend, and airport

1.3. The Model Implementation

Since in our graph, the values are different on weekends and on weekdays, we run our to find the minimum number of taxis needed to satisfy our constraints on a weekday, on a weekend, and on an entire week. Throughout this section whenever we use the word "day" It should be understood that we use our model when "day" means "weekday", " dan on weekend", and "week"

Given the data constructed above, we simulate the a day.

We loop through 1440 minutes in a day(10080 in a week), and in each minute. In each minute we do the following:

1. Generate 400 random numbers. This way we find if we have a custom at each node i needing a taxi to node j.

2. To each customer needing a ride, we send him the taxi that will provide him with the shortest waiting time. We do this by storing the amount of time each taxi, q, has until he is free, denoted by tuf_q. (That is, whenever taxi q picks up a customer, who needs to travel from node i to node j, we increase the value in tuf_q by $A_{i,j}$). We also store the location of each taxi once he becomes free, denoted by loc_q (that is, if taxi q drops the last customer at node k, then $\text{loc}_q = k$).

   Thus to find the taxi that will provide the customer, located at an arbitrary node, C

with the smallest waiting time we find

$$q^* = \min_q \mathrm{tuf}_q + A_{\mathrm{loc}_q, C} \tag{3}$$

Thus the taxi that will pick up the customer is the taxi indexed q^*

3. Store the waiting times of each customer who needed a ride.

2. VALIDATION

let $T(d)$ be the total number of taxis operating in Ithaca, where d is the number of taxis Ithaca operating from Ithaca Dispatch. Since Ithaca Dispatch has 2/3 of the taxi rides, and the regulated taxis account for 6/7 of the total taxis we have. Since we can not have a fractional number of taxis we simply consider the floor.

$$T(d) = \lfloor \frac{7}{6}\left(\frac{3}{2}d\right) \rfloor \tag{4}$$

2.1. assumptions

1. We assume that taxis are just waiting when they are not driving passengers. If a taxi dropped off a customer at some node k, then the taxi waits at node k until he needs to pick up another customer

2. Other then airport, we have at most 1 cal/ min. This is a good assumption, since the probability of having multiple calls from one one place to another place in Ithaca, in the exact same minute is negligible.

   Deal with airport case

3. No difference between different times of year. Ithaca dispatchers Data supports this assumption(difference is surprisingly small) . The total number of taxis from Ithaca Dispatchers in a weekday during the summer is 410, and thus the total number of taxis operating during the summer is $T(410) = 478$. [1] The total number of taxis operating from Ithaca Dispatchers during the school year is 450, so the total number of taxis operating during the school year is $T(450) = 525$. [1] Thus the percent change of the number of taxis during the school year and during the summer is only 8.95%. Note that although our entire model is based on data from the school year, and results do not change significantly over the summer, our model can easily be extended(**Expain this) to deal with the summer.**

4. All taxis are identical(same speed etc.)

[1] Ithaca Dispatchers