Finding an Optimal Route: Minimizing Police Stops

Priya Chatwani, Amanda Jacquez, Daniel Salz
Stanford University
CS 238: Decision Making Under Uncertainty
priyac@stanford.edu, ajacquez@stanford.edu, dasalz@stanford.edu

Abstract

The goal of our project is to create an algorithm that will enable user to find an optimal route to a destination in San Francisco. We determine a route as optimal if one can avoid police stops and traffic as much as possible during the course of a trip. We utilized a OSMnx dataset to capture the streets network of San Francisco, giving us a graph in which the nodes were represented as intersections and the edges were streets. We then implemented SARSA to compute the Q-values of our model. Historic traffic and police stop data informed the calculations of the rewards.

1 Introduction

Traveling from one point to another within a city is an activity done by hundreds of thousands of people every day. As residents of the Bay Area, it is clear that this activity is especially difficult to navigate in a densely populated city like San Francisco. Whether commuting to work or going for a meal across the city, there is a constant movement of people transporting themselves from one location to the next. Since a busy city implies a more active police force, we were inspired to create an algorithm that would enable citizens of a city to take the route that minimizes traffic and police activity.

We began this project with the intention of using value iteration to calculate the Q-Value of all states. We defined our states as a tuple, with the current node and the destination node. However, after calculating the size of our state space, value iteration was no longer feasible, so we settled upon SARSA and used the python random library to generate sample data. With Q-Learning, we defined a policy that determines which action to take based on the given state and the desired destination. Given the graph representing San Francisco's street network, we were able to add a higher cost or more negative reward to edges that had lots of traffic and were historically frequented by police. This weighting allowed us to determine an optimal route.

2 Motivation

Police are a source of fear for many people. With this in mind, we decided to use patterns in traffic and police stops to optimize one's route from their starting location to their desired destination. Currently, routing applications only take distance, traffic, and time of day into account. While Waze does provide live updates from other users when a police officer is seen, it does not allow users to pick a route that is historically less frequented by police.

We wanted to implement our model in a large city that is congested both in terms of vehicle traffic and population density. San Francisco seemed like a clear choice because it is the nearest large city to campus. Having all lived in San Francisco, we felt that these results could impact both our lives and the lives of those around us. As a relatively diverse city, San Francisco also enabled us to make predictions on varying races and genders.

Submitted to 30th Conference on Neural Information Processing Systems (NIPS 2016). Do not distribute.

Figure 1: San Francisco Drivable Streets

3 Literature Review

For this project, we wanted to incorporate police activity with reinforcement learning because there are little to no studies that involve the criminal justice system and this particular type of machine learning. Integrating police and criminal justice data with machine learning is still a relatively new topic and a growing field, and it mostly deals with the likelihood of the police targeting and condemning particular groups of people. Most of these existing machine learning models are biased against Black and Brown people and do not help uplift the communities being targeted at large. [2]

For our project, we wanted to implement a machine learning model that would enable those who are typically targeted by the police to drive with a stronger sense of security. A Stanford Computational Policy Lab report produced by Pierson et al. (2019) shows that police officers are more likely to pull over Black and Latinx drivers compared to white drivers. [3, 4] The project includes 100 million records of traffic stop and search data nationwide, and determines that, compared to white drivers, Black drivers are 20% more likely to be ticketed and Latinx drivers are 30% more likely to be ticketed. Furthermore, reputable newspapers like the New York Times and San Francisco Chronicle have accused San Francisco's Police Department in particular of being racially biased in their traffic searches and for stopping African-Americans in disproportionately high numbers compared to San Francisco's black population. [5, 6]

Victims of police harassment may be familiar with their own hometowns, but are often caught off guard when visiting or driving in a new city. Marginalized groups deserve to be aware of routes that would allow them to avoid the police and, as a result, avoid the harassment that could accompany these interactions with the police.

Since we are dealing with driving routes, we also wanted to integrate traffic data to allow for routes that have less traffic. We found that previous research exists with regards to traffic stops and reinforcement learning.

Horiguchi et al. (2005) created a model to minimizes routes with traffic congestion by using reinforcement learning to optimize the control parameter of a neural network. [1] In this model, reinforcement learning helps avoid traffic by rewarding routes that are not typically congested. Almost all of the existing models in this topic area utilize this basic outline, yet they vary based on how the rewards are weighted and how states are created in a particular location.

4 Data Analysis and Visualization

For this task, our datasets were crucial to our work. To determine a state space, we chose to use the OSMnx dataset. This dataset includes a multitude of information about every city's driving, walking, and biking networks. In OSMnx, the intersections are encoded as nodes and the paths between intersections are encoded as edges, creating a network of the streets that is easy to use and manipulate. The OSMnx graph of the city of San Francisco is cited above.

In order to ingest police data, we used a dataset from the Stanford Open Policing Project. Specifically, we chose a San Francisco CSV file with 905,070 entries of police stops from December 2006 to June 2016. This particular CSV file was beneficial because it has a large array of different features analyze. These features include: date, time, location, latitude, longitude, district, subject_age, subject_race, subject_sex, arrest_made, citation_issued, warming_issued, outcome, contrabound_found, search_conducted, search_vehicle, search_basis, reason_for_stop, and raw_search_vehicle_description. We decided to use the features subject_race and subject_sex, because we predicted those two features to be the most important in determining what factors are most influenced by interactions with the police force.

Finally, we used the City and County of San Francisco's dataset on traffic calming features to understand traffic patterns among the streets of San Francisco. This dataset includes location information about traffic calming features such as speed humps, speed cushions and speed tables. Having been published by the Municipal Transportation Agency, it also uses

5 Implementation

In outlining the implementation of our project, we will describe our intuitions behind the model, the reward function, the reinforcement learning algorithm used, and the ways in which we extended our state space to include more features.

5.1 Model

In analyzing the intricate street network of San Francisco, we needed to define our model as compactly as possible.

To do so, we first thought about the state space. The OSMnx dataset provided us with every intersection in San Francisco. With this in mind, we decided to assign a number, or "id", to each intersection, or node, beginning with 1 and ending in the total number of nodes. Thus, our state is represented as a tuple, where the first value refers to the beginning node's "id", and the second value refers to the end node's "id".

With regards to the action space, our understanding of street networks informed how we defined the possible actions. The OSMnx dataset provides edges to and from every node to reflect the different pathways going through that particular node. The maximum number of edges coming out of a node, in the San Francisco dataset, is 6. For that reason, we defined our model to have 6 possible actions. The map outputted of the city of San Francisco is included below

If one of those six possible actions is impossible from a particular node, then we assigned a large negative reward to that state, action pair. The next section will go into more detail on how the rewards were calculated.

5.2 Rewards

The rewards were calculated using both the police and traffic datasets. During SARSA, each state, action pair was assigned a reward that consisted of a weighted sum of the police and traffic reward. Our process for calculating rewards is described below.

For both the police stop and traffic calmer datasets we found the node in our model that was closest to each data point in either dataset. Closeness was defined by the minimum haversine distance based on longitude and latitude coordinates. Haversine distance is defined as :

$$2rsin^{-1}(\sqrt{sin^2(\frac{lat2-lat1}{2})+cos(lat1)cos(lat2)sin^2(\frac{long2-long1}{2})})$$

Where r is the radius of the Earth.

For the police dataset once the closest node was identified, the count at that node for police stops is decremented by 1 in the simple model, where the count is initialized to 0. In the more complex model where the state consists of the (current node, gender, race, goal node) a different calculation was used. With the extra states the reward at a specific node becomes a ratio of matching features in the state to features present in the police stops at the node the action is being taken. For example if a node has three stops where all three were male, but one was white and the other two asian; if you are a white male the reward associated with going to that node would be:

$$-\frac{\frac{3}{3}+\frac{1}{3}}{2} = -\frac{\frac{\#matchingfeature1}{\#stops}+\frac{\#matchingfeature2}{\#stops}}{\#features}$$

For both the simple and complex model the reward values were normalized by dividing all police reward values by the maximum number of police stops at any node in the graph.

For the traffic calming dataset once the closest node was identified, the count at that node for traffic calmers is decremented by 1, where the count is initialized to 0. All traffic calmer rewards were normalized by dividing all traffic reward values by the maximum number of traffic calmers at any node in the graph.

If the current node and goal node in the state are the same, the reward is 1 regardless of the police stop or traffic data at the goal node.

5.3 Reinforcement Learning

We came into this project with the goal of using value iteration. However, by defining our state as a tuple of nodes, our state space became 9,615 x 9,615 = 92,448,225. Thus, we turned to model-free reinforcement learning, and decided to implement SARSA using epsilon-greedy as our exploration strategy.

In order to mimic the sampled states that are observed in SARSA, we used the Python's random library to randomly select the nodes for each state with every iteration. Then, we determined which of the six possible actions were possible from a given state. Using the epsilon-greedy strategy, we randomly select an action 0.01% of the time, and the rest of the time, we chose the action that maximized the pre-computed rewards. We used a learning rate of 0.01, and a discount factor of 0.95. Having pre-computed the node that is reached after taking a particular action from the current node, we easily computed the next state to be used in the modified Bellman Equation.

We also needed to find the nearest neighbor of every state that was not sampled. Given an unvisited state, we referenced the latitude and longitude of its current and destination nodes using the OSMnx dataset. We then compared those latitude and longitude values to the latitude and longitude values of all seen states, allowing us to find the closest match.

5.4 Adding Driver Features to State

As is evident from our motivation and literature review, we wanted to build a project that would be useful to groups historically targeted by police. Thus, we decided to add race and sex to our state space, allowing the route to be optimized differently based on the race and sex of the individual, in addition to their starting and ending nodes. Our expanded state looked as follows: ((starting node, sex, race), ending node).

To implement this, we modified our rewards to account for the race and sex documented in a police stop. If the driver's race and/or sex matches the race and/or sex of the driver who was previously stopped, then the reward was decreased to reflect the higher risk of being stopped.

With regards to the SARSA Q-Learning, to account for the larger state space, the modified model also sampled from the possible race and sex classifications. In finding the nearest neighbor, we had to be thoughtful about calculating the closest match to a state. Given an unseen state, we decided to traverse every edge of the starting node, in the fashion of Breadth First Search, and check whether any of those neighbor nodes had been seen with the same race, gender, and destination. If none had been seen, then we changed the gender and checked if any of those states had been sampled. Next, since we did not want to discard or ignore race, we checked the neighbors of all the neighbors, moving one degree outwards. This process continued until a seen sample was found.

6 Results

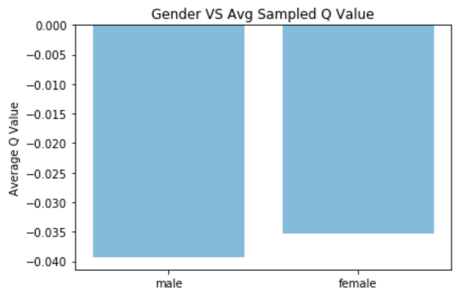

Figure 2: Q Values Gender Breakdown

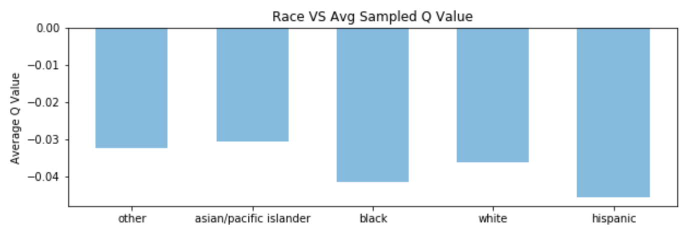

Figure 3: Q Values Race Breakdown

6.1 Quantitative Analysis

As seen in the bar graphs (Figures 2 and 3), the Q Values from any start and end node can be broken down by gender and race. Thus, when the traffic and police rewards are weighted equally, the average Q Values by gender are as follows:

$$\text{Male} : -0.0394047701512522 66$$

$$\text{Female} : -0.03537874319378656$$

Broken down by race, the average Q Values are:

$$\text{Black} : -0.041722440002357 74$$

$$\text{Hispanic} : -0.04572235983446658$$

$$\text{Asian/Pacific Islander} : -0.030605217538222023$$

$$\text{White} : -0.036306118889654015$$

$$\text{Other} : -0.032474203552906436$$

Due to the police data, the state-action value is lowest when the race "Hispanic" is included in the state. The second lowest average state-action value appears when the race "Black" is in the state.

In contrast, when the reward weights the police data by 0.75 and the traffic data by 0.25, the average Q Values are:

$$\text{Male} : -0.023323540427791455$$

$$\text{Female} : -0.019268010887294348$$

$$\text{Black} : -0.021964136512749085$$
$$\text{Hispanic} : -0.022377988795041274$$
$$\text{Asian/Pacific Islander} : -0.020099872327573355$$
$$\text{White} : -0.02082697220514877$$
$$\text{Other} : -0.021207243541240667$$

Once again, Black and Hispanic drivers have more negative rewards, on average, than other races, but the gender breakdown looks different. This time, men are seen to be at higher risk than women.

This confirms our knowledge of police officers dis-proportionally targeting black and brown drivers, which causes the rewards for black and brown drivers to be generally lower because those groups are at higher risk of encountering the police.

6.2 Next Steps

Right now, the model only accounts for police and traffic data along any edge. Our next step is to also include distance from the destination or some sort of indication that the driver is getting closer to their goal. Right now, we reward states that have the same current and end node, but we need to implement eligibility traces to spread that reward closer to the start state.

7 Group Contributions

Daniel: Created and programmed the entire structure for how rewards were created and accessed. This included processing all the data to find minimum distances, creating the methodology for weighting rewards, storing them in an efficient manner, etc.
Priya: Coded and implemented Value Iteration and then SARSA for original and expanded states, including how to combine rewards and find nearest neighbors. Quantitatively analyzed the results of the Q Values. Research of datasets and existing evidence of racial bias in policing.
Amanda: Also worked on value iteration and the SARSA implementation. Conducted research for the datasets, the literature review on past work, and the plan of action and milestones.

8 References

[1] Tsuyoshi Horiguchi, Keisuke Hayashi, Alexei Tretiakov, Reinforcement learning for congestion-avoidance in packet flow, Physica A: Statistical Mechanics and its Applications, Volume 349, Issues 1–2, 2005, Pages 329-348, ISSN 0378-4371

[2] Hao, Karen. “AI Is Sending People to Jail-and Getting It Wrong.” MIT Technology Review, MIT Technology Review, 21 Jan. 2019, https://www.technologyreview.com/s/612775/algorithms-criminal-justice-ai/.

[3] Emma Pierson, Camelia Simoiu, Jan Overgoor, et al. "A large-scale analysis of racial disparities in police stops across the United States." Stanford Computational Policy Lab, 13, Mar. 2019, https://5harad.com/papers/100M-stops.pdf/

[4] Saxon, Shani. "New Report Analyzes Racial Bias in Police Traffic Stops." Colorlines, 14, Mar. 2019, https://www.colorlines.com/articles/new-report-analyzes-racial-bias-police-traffic-stops.

[5] Palomino, Joaquin. "Racial disparities in SF traffic searches raise concerns of bias." San Francisco Chronicle, 8, Apr. 2016, https://www.sfchronicle.com/crime/article/Racial-disparities-in-SF-traffic-searches-raise-7235690.php.

[6] Williams, Timothy. "San Francisco Police Disproportionately Search African-Americans, Report Says." New York Times, 11, Jul. 2016, https://www.nytimes.com/2016/07/12/us/san-francisco-police-disproportionately-search-african-americans-report-says.html.