Taxi-for-all: Incentivized Taxi Actuation System for Balanced Area-wide Service

PI: Carlee Joe-Wong

Co-PIs: Haeyoung Noh and Pei Zhang

FINAL RESEARCH REPORT

Contract # 69A3551747111

Problem Description

In modern smart cities, understanding of the city-wide area is paramount to improve transportation throughout the city. City-wide sensing (such as air-pollution monitoring, traffic monitoring, road-surface sensing, etc.) and city-side transportation (taxis, buses, etc.) all require wide coverage over the entire city. However, vehicles (such as taxis) are not evenly spread and tend to be centered around the central business district. Thus, their geographical distribution is often a poor match for transportation or sensing needs during times where vehicles are especially needed. This distribution particularly threatens disadvantaged communities, e.g., if taxis are used as a city's primary means to sense air pollution, they may concentrate in wealthy business districts during morning rush hour. They then cannot sense the air pollution in poorer neighborhoods at this time, making it more difficult to assess air pollution's health impacts in those communities.

We aim to derive a method to incentivize taxis to travel to different parts of a city, with the goal of ensuring that they sense data evenly throughout the city. We assume that taxis continuously sense data at a fixed frequency regardless of their location. Thus, controlling the distribution of sensed data is equivalent to controlling the movement of the taxis. In general, taxis' natural movement will lead to them concentrating in busy locations without many potential passengers, which is undesirable for sensing or transportation. By incentivizing taxis to move towards locations with more passengers and higher needs, we can also make the distribution of sensed data more spatiotemporally uniform and thus useful for city operators.

Solving this problem, however, comes with several challenges. First, it is difficult to know in advance how taxi drivers will react to the incentives offered, making it difficult to determine which incentives should be offered. Indeed, even determining how incentives should be shown to taxi drivers may be challenging, since drivers may react differently to different types of incentives (e.g., fewer drivers may follow the incentives offered if they are required to follow specific routes to gain the incentive payments, compared to simply detouring from their planned routes). Modeling individual drivers' reactions to incentives then requires tracking significant information for individual drivers, such as their willingness to drive to specific locations, that may not be available to us. Second, even if we had a model for taxi drivers' reactions to incentives, relating these actions to the resulting distribution of data requires tracking individual drivers' movements, which may not scale well with multiple drivers or a large city. Our work aims to solve these problems by abstracting the problem into a reinforcement learning framework and evaluating the resulting pricing algorithms on real taxi and vehicle mobility datasets.

Approach and Methodology

To solve the problems outlined above, we consider an incentives framework in which drivers are offered a location-based incentive for reporting measurements from different locations around a city. Since the incentives earned by each driver are then solely tied to the driver's current location, drivers can easily understand the actions they must take to earn incentive payments. However, this framework also complicates our modeling of drivers' decision-making. For example, in order to know which routes a rational driver seeking to maximize the earned incentive will take, we must evaluate the potential incentives earned on each possible route, which is computationally infeasible for a large city. Most drivers will not even act in a perfectly rational manner, instead incorporating their intrinsic preferences for different routes into their decision making. Thus, we circumvent this challenge by abstracting driver decisions into a reinforcement learning framework that allows us to learn how they move through the city as a function of the incentives that we offer at each location.

Our reinforcement learning framework takes as possible *actions* the prices offered at each location, and defines the *reward* from these choices as the difference between the distribution of data collected and a spatiotemporally uniform distribution. As is typical in reinforcement learning frameworks, we then choose the prices according to *state* variables; our goal is to learn the action that should be taken when each state is observed (sometimes called learning the *policy*), so as to maximize the reward over time. Our state variables encode the current distribution of active (i.e., currently carrying passengers and thus unresponsive to incentives) and idle vehicles across different locations in the city, as well as the number of passengers available at each location. Taxi drivers would naturally be inclined to travel towards those locations with multiple passengers available, and intuitively we would offer fewer incentives there.

Given this reinforcement learning framework, our next step is to design the algorithm that can learn an optimal policy. One option is to simply utilize existing reinforcement learning algorithms, e.g., deep deterministic policy gradient (DDPG) approaches. A key challenge to doing so, however, has been the scalability of our algorithms to realistic time durations and city sizes. In particular, the number of action variables scales linearly with the number of locations, as does the number of state variables. We have begun investigating the use of clustering algorithms to reduce the size of our state space and learn policies with a manageable number of free action variables. Clustering allows us to identify locations that should offer the same prices given some state variables (e.g., due to similarities in their states), reducing the number of actions we must consider.

Finally, the last part of our methodology is to evaluate the algorithms we propose to learn the optimal pricing policy. We have built a simulation platform for taxi movement around a city as a function of the incentives offered to taxi drivers, which will allow us to test our formulation and algorithms. We can specify how taxi drivers make their decisions (e.g., randomly, greedily maximizing the incentives earned, etc.) in order to model different types of driver actions. As part of the simulator construction, we have implemented existing reinforcement learning techniques that can take as input generic state and reward variables and attempt to learn the optimal actions. Currently, the simulator integrates data on taxi movement from Beijing.

Findings

Our first finding identifies the challenges in applying reinforcement learning to problems with a large number of spatial state variables, in particular scalability challenges as the size of the state space grows. We believe that this challenge will arise in general spatiotemporal learning scenarios beyond taxi incentivization, where an action variable must be optimized at each location over a large geographical space. The algorithms we are currently developing, which intelligently cluster state variables to learn a reduced policy that is nonetheless almost optimal, may thus be useful for a more general class of reinforcement learning problems.

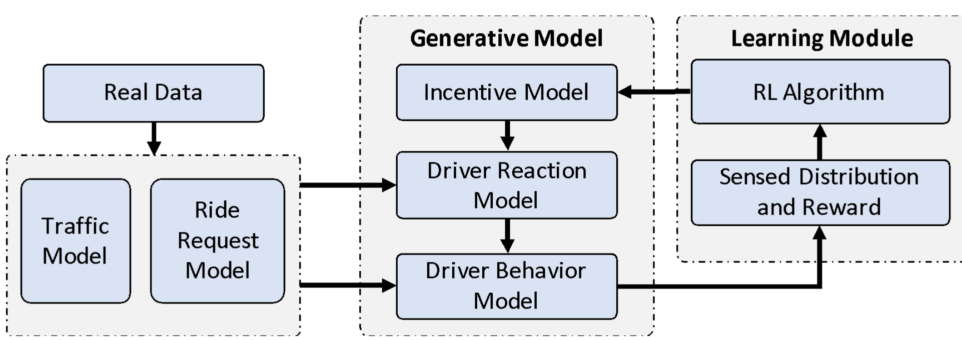

Figure 1: Architecture of our taxi simulator with incentives.

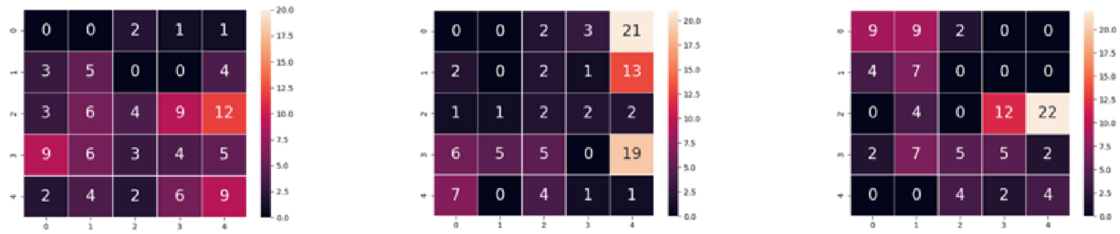

Figure 2: Heat map of the distribution of vehicles in a five-by-five grid with no (left), uniform (center), and random (right) incentives offered. We observe that the distribution changes dramatically depending on the incentives offered.

Our second finding is that driver reactions can significantly impact the distribution of data collected. Figure 1 shows the architecture of the taxi simulator that we developed to test these findings. By changing the incentive, driver reaction, and driver behavior models, we can find that we can change the distribution of data collected over the city. Figure 2 illustrates heat maps of the spatial distribution of collected data under no, uniform, and random incentives in a simplified five-by-five city grid. Numbers in each grid square represent the number of measurements collected over one day. We observe significantly different distributions for the three different incentive types, indicating the importance of these incentives to changing driver behavior.

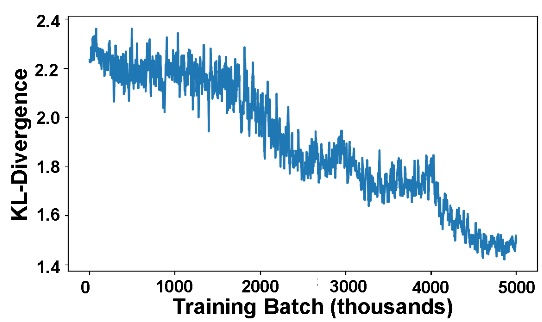

Figure 3: Our reinforcement learning approach decreases the KL-divergence between the actual distribution of collected data and a spatiotemporally uniform distribution by 32% compared to no incentivization.

Our third finding is that using reinforcement learning to optimize the prices offered is effective in making the distribution of collected data more uniform. Figure 3 shows the reduction in KL-divergence (a measure of how closely the achieved spatiotemporal distribution matches a uniform one--lower means a closer resemblance) as we train the algorithm longer. We observe a significant decrease in the KL-divergence, indicating that utilizing reinforcement learning to select prices is effective in optimizing the distribution of collected data.

Finally, training and professional development opportunities have been provided in the form of taxi datasets that have been used for projects in a data analysis course at Stanford. We plan to use these datasets in other courses in later semesters as well.

Conclusions

We report four major conclusions from this project: (i) the development and implementation of a software simulator for taxi drivers' reactions to incentives offered to them, (ii) the development of a reinforcement learning framework to optimize these incentives, (iii) novel reinforcement learning algorithms that handle the scalability challenges associated with our pricing problem and related spatiotemporal optimizations, and (iv) educational integration of our findings into relevant courses.

Our simulator, which is the basis for our workshop paper, includes modules that admit customized models for how taxi drivers react to the incentives given, e.g., some drivers may simply ignore the incentives, while others may attempt to optimize their driving destinations so as to maximize the incentives received. The simulator utilizes traces of taxi activity from Beijing, China; and models the effect of taxi driver movement on congestion throughout a city. We have released an open-source version of the simulator that would allow other researchers to use our work in their own research on incentivization in transportation networks.

Our reinforcement learning framework, as described above, specifies the prices chosen throughout the city as actions taken so as to maximize a reward function, which we define as the degree to which the taxi distribution matches a specified target distribution. These techniques will allow sensing or taxi operators to ensure that taxis spread themselves around a city according to sensing or passenger needs. These algorithms have also increased the body of knowledge pertaining to spatiotemporal reinforcement learning, and we expect to make further advances in this area over the course of the project.

Recommendations

Our first recommendation is that companies interested in controlling the distribution of taxis over a city should consider using monetary incentives as a means to influence where taxis should travel. While finding the right incentives to offer is challenging, our work shows that reinforcement learning is a promising approach to solve the resulting optimization problem. We also introduce a framework and algorithms that can help practitioners formulate and solve their problem as a reinforcement learning one.

We presented one workshop paper on our simulation platform at a workshop of ACM UbiComp in September. The platform has since been released as open-source code on Github (see below), where it is available for others to use and adapt in their own projects. We plan to continue disseminating our results by publishing additional papers on our frameworks and algorithms. In particular, we are working on a journal paper submission that outlines our reinforcement learning framework ideas.

We also anticipate working with current and future partners to put our work into practice. We have obtained taxi mobility data from DriveSally, a company that rents cars to Uber and Lyft drivers, which we will be able to use to simulate taxi movements in response to different driver incentives. In particular, Drive Sally would like to use our algorithms to incentivize their drivers to locations where there are more potential passengers and they can earn revenue by displaying effective digital advertisements. We are also discussing how our incentive framework can be useful for our original partner Roadbotics.

Publications and Products

Wu, Z., Zhang, X., Xu, S., Chen, X., Zhang, P., Noh, H. Y., & Joe-Wong, C. (2020, September). A generative simulation platform for multi-agent systems with incentives. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers* (pp. 580-587).

Simulation code: https://github.com/namabilly/iLOCuS

Project Participants

Xinlei Chen
Carlee Joe-Wong (OrcID: 0000-0003-0785-9291)
Haeyoung Noh (OrcID: 0000-0002-7998-3657)
Zhengwei (Billy) Wu
Susu Xu (OrcID: 0000-0001-7170-648X)
Pei Zhang
Xiaoxi Zhang