Optimization of Taxi Revenue

*Soham Chitalia (*Author*)
Student, Department Of Computer Engineering
Dwarkadas J Sanghvi College Of Engineerng
Mumbai, India
9.soham@gmail.com

*Miloni Joshi (*Author*)
Student, Department Of Computer Engineering
Dwarkadas j Sanghvi College Of Engineering
Mumbai, India
joshimiloni@gmail.com

*Shraddha Shaligram (*Author*)
Student, Department Of Computer Engineering
Dwarkadas J Sanghvi College Of Engineering
Mumbai, India
shraddhatps@gmail.com

Sindhu Nair (*Guide*)
Assistant Professor, Department Of Computer Engineering
Dwarkadas J Sanghvi College Of Engineering
Mumbai, India
sindhu.nair@djsce.ac.in

Abstract— **Taxi services play a pivotal role in the urban transportation system to serve the large population in Mumbai. Expanding city limits and the inability of transportation services such as trains, metros, and buses to reach the city outskirts has made taxis an integral part of the transportation system. Improving the taxi business is an important societal concern since it will not only increase the revenue of taxi drivers but also serve the increasing customer demand with more efficiency. By analyzing the past trajectory data and guiding taxi drivers to appropriate locations, we intend to reduce the cruising time between trips and thereby reduce fuel wastage. We propose the Q-learning technique of Reinforcement Learning to suggest optimal driving strategies and consequently ensure revenue maximization for the taxi drivers. The main objective is to yield higher revenues which will bring more drivers into the system facilitating better availability for the customers in dense urban cities.**

Keywords—Markov Decision Process, Optimization, Q-learning, Reinforcement Learning, Taxi driver.

I. INTRODUCTION

In Mumbai, the increase in the use of fleet taxi services such as OLA and Uber has reduced the revenues earned by local taxi drivers. In pursuit of ensuring a fair competition between these technologically supported systems and the Indian cab drivers, our research focuses on providing the necessary analytics, and other data sources to the Indian cab drivers as well. The main objective of our research is to increase the profits earned by local taxi drivers by reducing their passenger seeking time and guiding them to appropriate locations for passenger pick-up. This has a two-fold advantage, while it guides the driver to the appropriate locations, it also increases the ease with which someone can get a taxi. Apart from this, using the following approach also reduces pollution caused due to the burning of fuel.

One major challenge is to ensure that the optimization of revenue earned by taxi drivers takes place over a longer period of time and not just for one ride. Hence it is important to lead the taxi driver to a location where it is easier to find a passenger and the destination of the passenger should be such that the taxi driver can get the next passenger easily. To overcome this challenge, we have planned on using Reinforcement Learning. The taxi driver acts as an agent who goes through various states and receives rewards and punishments for correct and wrong actions respectively. In order to come up with the optimal reinforcement learning model, we have analyzed the research work performed for similar optimization problems.

II. RELATED WORK

A. *Optimization of the Revenue of the New York City Taxi Service using Markov Decision Processes.*

This paper [1] uses Markov Decision process (MDP) for better routing to optimize the total revenue earned by taxi drivers. This method calculates the probability of finding a customer and determines the corresponding drop-off location using historical data. This historical data used by MDP consist of a combination of location, time, current and previous actions. Data analysis is followed by providing performance indicators of the taxi drivers which will be used to optimize the routing decisions of the taxi drivers in the MDP. The states, actions and reward are supported by defining probability parameters, time parameters, reward function, state-transition function and the objective function. Open Street Map (OSM) has been used for determining the number of taxis at the nearest junctions. The major reason for using MDP is its ability to handle

*All authors have contributed equally to this paper.

uncertainty in taxi demand. Dynamic programming has been used to implement the MDP model. Different models have been used for daytime and night time as well as for weekdays and weekends. The model proposes the best action to be performed by the taxi driver at each state in order to reduce the seek time and increase revenue. Though the model used in this paper shows a 10% improvement in accuracy, the results are not clear.

B. Reinforcement Learning Based Algorithm for the Maximization of EV Charging Station Revenue.

The work presented in this paper [2] focus on reinforcement learning based application which increases the revenue of one particular electric vehicles (EV) station by taking into consideration the fluctuations in demand. The model proposed here considers different parameters such as discrete time intervals (hourly intervals) and the different levels of charge that a client require. Q-learning has been used for optimization with the main objective to strike a balance between the expenditure (buying additional power from the power grid) and the revenue earned from clients who charge their vehicles at the EV charging station. Q-learning learns the power consumption patterns so that the charging station does not have to buy excess units of power from the power grid. Using Q-learning has shown an increased the station's income in the range of 40-80%. The work done in this paper can be correlated to the taxi revenue optimization problem where we need to increase the profits made by the taxi drivers by decreasing the time wasted between rides.

C. Augmenting Decisions of Taxi Drivers through Reinforcement Learning for Improving Revenues..

The work presented in this paper [3] focuses on improving performance from a taxi driver's perspective. The decisions taken by the taxi drivers are annotated to provide the taxi trajectory data followed by a Reinforcement learning model that learns from the annotated data. Annotating the data derives a driver's activity graph which is converted to a sequence of state-action pairs. The activity graph is divided into episodes for the purpose of application of Reinforcement learning. Each node of the activity graph denotes a state and the action to be taken is denoted by the subsequent nodes. Monte Carlo methods are used to estimate the state-action pair and also compute the best action in each state. It performs calculation of state abstraction by using clustering and learn from activity graphs by using Monte Carlo RL method. Dynamic state abstraction approach is used to improve the performance. The performance is evaluated by comparing with the revenues already earned by taxi drivers. On evaluation it is found that the RL agent performs better than top 10 percentile of the drivers.

D. The Rich and the Poor: A Markov Decision Process Approach to Optimizing Taxi Driver Revenue Efficiency.

The work performed in this paper [4] models the impact of current and past passenger destinations on future passenger-seeking as a Markov Decision Process(MDP). A detailed analysis is performed to identify the strategic differences between the most and least successful drivers followed by modeling the passenger-seeking process as a MDP. The study area is partitioned into grids where each state in MDP includes the current grid, time and driving direction. The best action for a vacant taxi in a particular time slot is determined to maximize the total revenue in that time slot. This includes the model learning a different input set for MDP from the historical data for each time slot. The MDP is solved using the dynamic programming approach. The output contains the best action to be taken for each state. Using this model, the revenue efficiency of inexperienced drivers increases by up to 15%.

E. Resource Adaptations for Revenue Optimization in Cognitive Mesh Network Using Reinforcement Learning.

In this paper [5], Cognitive Mesh Network's ability to offer quality of service (QoS) for real time applications is explored. The approach considered includes licensed users renting surplus spectrum to unlicensed users to get some monetary reward. However, when a licensed user does this, its quality of service is degraded due to reduction of the spectrum. This requirement is embedded in the Reinforcement Learning Model in this paper. Reinforcement Learning is proposed as a resource management scheme to obtain an optimal policy that supports primary user adaptation to changing network conditions and maximization of the revenue earned by the primary user over time in the radio environment. Requirements such as rewards and Quality Of Service for licensed users are integrated in the proposed scheme. Reinforcement guarantees the required output as it is shown that Cognitive Mesh Networks can support additional unlicensed users traffic while still ensuring licensed users quality of service and profit maximization.

F. Optimize taxi driving strategies based on Reinforcement Learning.

Taxi trajectories being long sequences cannot provide the global optimum solution through single step optimization. Hence this paper [6] focuses on presenting Markov Decision processes as an optimization for the taxi trajectories which guarantees global profit maximization. Based on historical data from trajectory logs, assistance is provided to the passengers for finding a cab and to cab drivers for finding the next passenger. When cab drivers receive multiple requests from different locations at the same time, they use their past experience to cater to one of the request. This past experience is often subjective and may not be accurate enough to improve the revenue. Hence, a Reinforcement Learning (RL) methodology has been proposed. Q-learning algorithm is used to solve this problem since it provides the optimal choice for cab drivers at any location with maximum cumulative rewards. The advantage of using this model is that there is no requirement of prior knowledge and it is globally optimal.

G. Improving Taxi Revenue with Reinforcement Learning.

The research performed in this paper [7] explores the use of Reinforcement learning to address the controversy between

Car aggregation systems (Uber, Lyft, etc.) and the New York taxi drivers. A Linear Regression model is used to determine the relationship between the taxi demand and factors such as weather, geolocation, date and time followed by a Reinforcement Learning model to maximize the total revenue generated by an individual Taxi Driver. Dynamic Programming is used to back test the records of drivers to differentiate their performance in terms of revenue with and without the optimal policy used. Different states of taxi drivers are based on the zip code in which they are located. The action includes travelling from one zip code to another. The optimal policy ensured ~20% improvement in the taxi revenue.

H. Optimizing Airline Seat Allocation using Reinforcement Programming.

This paper [8] focuses on optimizing the airline seat allocation to maximize the total revenue generated during low season. The basic model subsequently considers multiple fares on each route, time-dependent demands, and booking control on an extended network. It uses a new algorithm which uses the basic concept from reinforcement learning called reinforcement programming. The results obtained reinforcement programming are compared with genetic algorithm, particle swarm optimization, simulated annealing and ant colony optimization. Reinforcement programming has the best performance.

III. DATA

The dataset we referred was from 2013 NYC Taxi records [9] [10]. We analysed the dataset to understand the major components of the data given in Table 1. Apart from the trajectory logs, the data includes other important features such as taxi status: passenger on-board or vacant. We also intend to use the weather data as they directly affect the traffic conditions on the road.

TABLE 1

Taxi_ID
Pickup_latitude
Pickup_longitude
Pickup_timestamp
Dropoff_latitude
Dropoff_longitude
Dropoff_timestamp
Fare_amount

IV. PROPOSED METHODOLOGY

Reinforcement learning is defined as a technique to optimize the objective function by allowing an agent to take actions and interact with an environment.

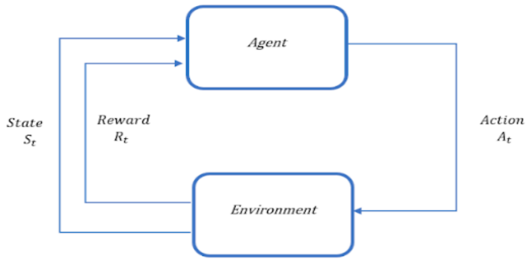

Figure 1

As shown in Figure 1, Agent receives state S_0 from the Environment. It takes action A_0 based on state S_0. According to the Environment then transitions to a new state S_1 and gives reward R_1 to the agent. The goal of the agent is maximization of the expected cumulative reward.

Q-learning is a reinforcement learning technique commonly used in machine learning. The goal of *Q*-Learning is to learn a policy, it prompts an agent to take an action under the given circumstances. For any finite Markov decision process (FMDP), *Q*-learning finds a policy that is optimal in the sense that it maximizes the expected value of the total reward over all successive steps, starting from the current state to the goal state.

Q learning will iterate through the state action pairs stored in the Q table (Q[s,a] arrays). Updates to the Q table are made at the end of each iteration (referred to as an episode) through the 'Bellman equation':

$$Q(s,a) = r + \gamma(\max(Q(s',a'))$$

'S' is the current state and 'a' is the action taken in state 's'. R is the reward received for the action 'a'. 'S'' is the next state that the agent goes to and 'a'' is the action taken in the state 's''. $\boldsymbol{\gamma}$ is the discount rate which adds the discounted rewards of the actions taken in the future states. $\boldsymbol{\gamma}$ is important since it determines the amount of importance to be given to future rewards based on the distance from the goal state.

In our case study, The agent (taxi driver) operates in an environment of which it has either a complete or partial knowledge. Here the city as a whole acts as the environment. In this case, since the agent is not familiar with the fluctuating demands, it has partial knowledge of the environment. The environment can be defined in terms of states and actions as follows:

- State : <Pickup_Latitude, Pickup_Longitude, TimeStamp, Dropoff_Latitude, Dropoff_Longitude, Current_Status>
- Action : <Pickup_Customer, Dropoff_Customer, Waiting,Seeking>

Based on the actions performed by the agent, a reward is given. Rewards and punishments are defined as follows:

- Reward : Right Drop-off : +100; Wrong Drop-off: -100; Waiting Time Penalty : -20;

With the environment, states, actions and rewards as described above, we propose to use Q learning to solve the problem since the number of states is less. Since Q learning is an on-line learning process, it converges to an optimal solution. We will be using ϵ-greedy as the behavior policy, i.e. with probability ϵ we either take a random action or choose the action with the highest Q value.

By using the Q learning approach, the taxi driver is asked to take an action depending on the current state. If the action is beneficial (eg. Right drop off) then the agent will receive a reward. But if the action leads to loss of revenue (either by increase in waiting time leading to loss of fuel or by a incorrect drop off), a punishment will be inflicted on the agent. Rewards are then calculated at each step and updates are made to the Q table. Based on the rewards obtained for past actions, suitable actions are proposed by the Q learning algorithm. The proposed system architecture flowchart is shown in Figure 2.

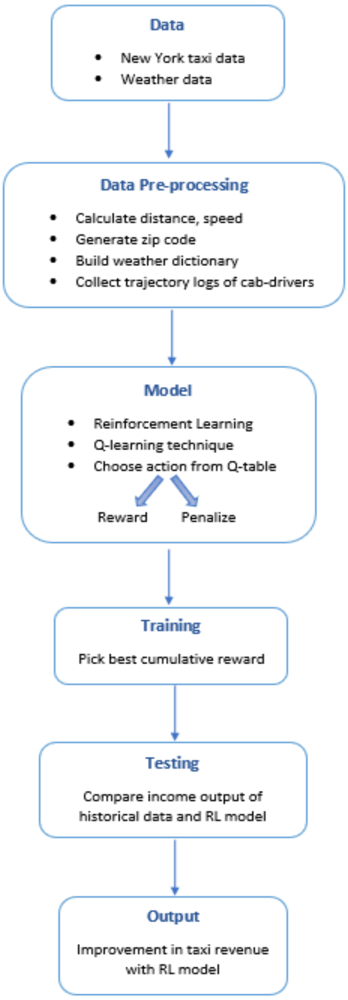

Figure 2

The use of the Reinforcement learning for this problem is justified because: (a) The problem demands maximizing the total revenue(i.e reward) generated by a taxi driver based on some actions such as staying in the same area to find a new customer, or travelling to a new location. (b) Reinforcement learning approaches can handle uncertainty well, in this case, uncertainty in quantity and destination of demand. (c) Reinforcements are well defined in terms of the fare and cost of travelling between locations of the trips. (d) Reinforcement learning adapts to fluctuations in demand quite well because it is focused on learning.

The use of Q-learning technique of Reinforcement Learning is justified because Q-learning can solve the process with stochastic transitions and rewards without requiring adaptations. The algorithm can be performed online and fully incrementally with minimum computation. These advantages work effectively for the taxi problem, because the status transition probabilities and rewards are indeterminate initially, and trajectories are accumulated incrementally.

V. Conclusion and Future Work

In this paper we have analysed and compared various techniques involving reinforcement learning for optimisation of taxi revenue. Out of all the techniques, Q-learning provides most accurate and desirable results. It is the most efficient methodology for increasing the profits gained by taxi drivers.

The dataset [9] [10] that we have currently referred contains trip data of New York taxi drivers. We intent to extend the proposed methodology and implement the model for Indian cab drivers by acquiring the appropriate datasets. In our further research, we would like to implement a similar model which can account for multiple agents at the same time and optimize the rewards of all the agents simultaneously. This would help boost the overall business of local taxi drivers. There is always a scope of improvement of a project. Therefore, better and more accurate results can be achieved by more better pre-processing and training of the dataset.

References

[1] G. Eason, B. Noble, and I.N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (*references*)

[2] S. Dimitrov and R. Lguensat, "Reinforcement Learning Based Algorithm for the Maximization of EV Charging Station Revenue," *2014 International Conference on Mathematics and Computers in Sciences and in Industry*, Varna, 2014, pp. 235-239.

[3] Verma, T., Varakantham, P., Kraus, S., and Lau, H. C. (2017). Augmenting decisions of taxi drivers through reinforcement learning for improving revenues. In ICAPS

[4] Huigui Rong , Xun Zhou , Chang Yang , Zubair Shafiq , Alex Liu, The Rich and the Poor: A Markov Decision Process Approach to Optimizing Taxi Driver Revenue Efficiency, Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, October 24-28, 2016, Indianapolis, Indiana, USA

[5] A. Alsarhan and A. Agarwal, "Resource adaptations for revenue optimization in cognitive mesh network using reinforcement learning," *2010 IEEE Globecom Workshops*, Miami, FL, 2010, pp. 1108-1112.doi: 10.1109/GLOCOMW.2010.5700107

[6] Yong Gao, Dan Jiang & Yan Xu (2018): Optimize taxi driving strategies based on reinforcement learning, International Journal of Geographical Information Science, DOI:10.1080/13658816.2018.1458984

[7] Wang, J., and Lampert, B. 2014. Improving taxi revenue with reinforcement learning.

[8] M. I. Isfahani, A. R. Barakbah and E. Martiana, "Optimizing airline seat allocation using reinforcement programming," *2015 International Electronics Symposium (IES)*, Surabaya, 2015, pp. 200-205. doi: 10.1109/ELECSYM.2015.7380841

[9] Dataset: http://blog.uber.com/uberdata/ #uberdata Uber Blog. Uber, n.d. Web.12 Dec. 2014.

[10] Whong, Chris. "FOILing NYCs Taxi Trip Data." FOILing NYCs Taxi Trip Data. N.p., 18 Mar. 2014. Web. 11 Dec. 2014.