

# Seeking for Passenger under Dynamic Prices: A Markov Decision Process Approach

Qianrong Shen

Jinan University, Guangzhou, China

Email: 18135181939@163.com

**How to cite this paper:** Shen, Q.R. (2021) Seeking for Passenger under Dynamic Prices: A Markov Decision Process Approach. *Journal of Computer and Communications*, 9, 80-97.  
<https://doi.org/10.4236/jcc.2021.912006>

**Received:** November 29, 2021

**Accepted:** December 28, 2021

**Published:** December 31, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In recent years, ride-on-demand (RoD) services such as Uber and Didi are becoming increasingly popular. Different from traditional taxi services, RoD services adopt dynamic pricing mechanisms to manipulate the supply and demand on the road, and such mechanisms improve service capacity and quality. Seeking route recommendation has been widely studied in taxi service. In RoD services, the dynamic price is a new and accurate indicator that represents the supply and demand condition, but it is yet rarely studied in providing clues for drivers to seek for passengers. In this paper, we proposed to incorporate the impacts of dynamic prices as a key factor in recommending seeking routes to drivers. We first showed the importance and need to do that by analyzing real service data. We then designed a Markov Decision Process (MDP) model based on passenger order and car GPS trajectories datasets, and took into account dynamic prices in designing rewards. Results show that our model not only guides drivers to locations with higher prices, but also significantly improves driver revenue. Compared with things with the drivers before using the model, the maximum yield after using it can be increased to 28%.

## Keywords

Ride-on-Demand Service, Markov Decision Process, Dynamic Pricing, Taxi Services, Route Recommendation

---

## 1. Introduction

Dynamic pricing is one of the key features that makes RoD attractive to both passengers and drivers—an effort to manipulate supply (the number of cars on the road) and demand (the number of passenger requests). Specifically, higher prices would attract more drivers and delay requests from passengers who are

not in a hurry; Lower prices are the opposite [1].

Dynamic pricing can help to improve service quality, but it also poses new problems for researchers. In RoD services, the price multiplier is a new indicator that drivers choose to look for strategies, and it more accurately describes local supply and demand conditions. But effective strategies remain to be explored. For example, if all drivers flock to a particular area with a high price multiplier, there will be an oversupply in that area, causing the price multiplier to fall sharply. This not only created unstable prices, but also headaches for drivers seeking high prices. As a matter of fact, many news reports and research papers have discussed this intuitive “chase the wave” strategy; however, they sometimes give contradictory advice [1]. Therefore, how to recommend drivers to areas with high dynamic prices is a pressing issue rather than verbal advice and discussion.

Recommending seeking routes [2] to drivers has been a frequently-studied problem in traditional taxi service [3] [4] such as mining the pattern of customer seeking strategy in the taxi GPS track, establishing MDP model so as to evaluate some strategies, etc. But in RoD service, dynamic pricing is a new and accurate indicator of supply and demand condition, and should be considered in seeking route recommendation. Nevertheless, dynamic pricing is rarely used in the research of RoD to recommend destinations to drivers. This is because most suggestions that come from news reports or blogs are not rigorous enough, and some existing studies are mostly based on theoretical models that require a lot of assumptions and approximations. As the matter of fact, the lack of real data in RoD service has hampered research based on data analysis methods [1].

In this paper, we design a Markov Decision Process [5] [6] [7] [8] (MDP) model to answer “how to use dynamic prices to help drivers in seeking for passengers”. We first illustrate the need to consider the impacts of dynamic prices by analyzing passenger order and car GPS trajectories data [9] [10]. We then establish an MDP model based on above datasets. In the model, first of all, the study area is meshed. Secondly, the passenger travel data and driver travel data are matched to the grid. Then we calculate the pickup probability of each grid to get the destination probability of passengers. Finally, we consider dynamic price as reward in our MDP model. To sum up, we adopt dynamic programming algorithm to solve MDP and obtain the optimal dynamics of each grid, thus recommending it to drivers.

The main contributions of this paper are as follows: 1) Introduce dynamic pricing into the seeking route recommendation problem in RoD service. 2) We design an MDP model to answer “how to use dynamic prices to help drivers in seeking for passengers”. We also design a dynamic programming algorithm to solve the MDP. Finally, compared with real drivers, the maximum yield of our model can be increased to 28%. 3) We conduct extensive experiment on real service datasets (including: passenger data, driver GPS trajectories data [11], etc.) and then we get our evaluation results.

The remainder of this paper is organized as follows. Section 2 analyzes the patterns of dynamic prices. In Section 3 we design a Markov Decision Process (MDP) model to answer “how to use dynamic prices to help drivers in seeking for passengers”. In Section 4 we evaluated our model. Finally, in Section 5 concludes the paper.

## 2. Dynamic Price Analysis

In this section, we will introduce what dynamic pricing is, and analyze the passenger data and driver GPS trajectory data of Shenzhou UCar [12] in November, and discover the potential impact of dynamic pricing on passenger and driver travel.

### 2.1. Dynamic Price Definition

Dynamic pricing can manipulate supply (the number of cars on the road) and demand (the number of passenger requests). Specifically, a higher price would attract more drivers and delay requests from passengers who are not in a hurry; and a lower price are the opposite. In most cases, dynamic prices are represented by price multipliers, so the fare of a trip is the product of a dynamic price multiplier (depending on conditions of supply and demand) and a fixed normal price (based on travel time and distance), hence, dynamic pricing can be written as:

$$D_y = p_k \times (15 + 2.8 \times \text{distance}) \quad (1)$$

In Equation (1), we utilize  $p_k$  to denote dynamic price coefficient and  $p_k \in \{1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6\}$ . Meanwhile from Equation (1), it can be seen that the dynamic pricing has an effect on driver’s revenue that the higher the dynamic pricing multiplier, the more the driver earns.

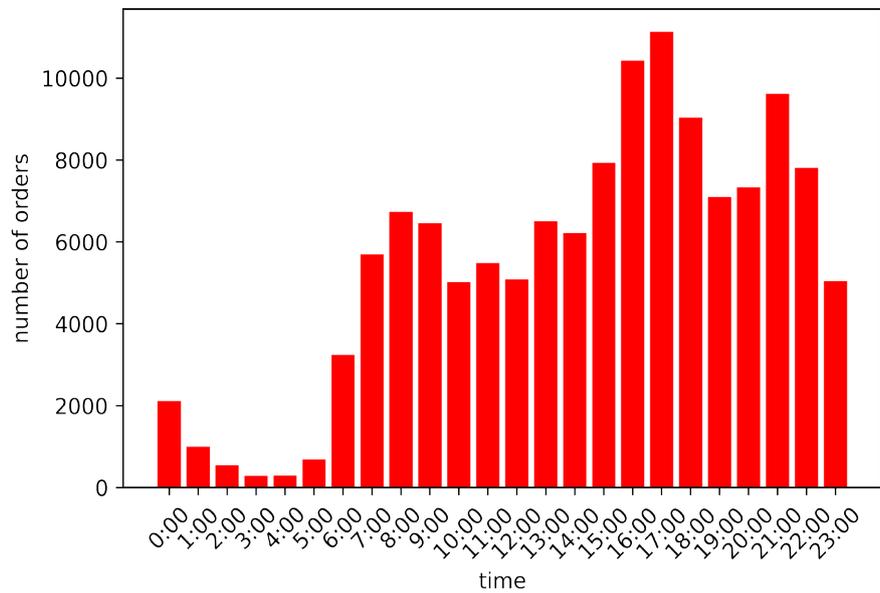
### 2.2. Data Analysis

In this paper, we randomly selected passenger data of one day for data analysis, as shown in **Figure 1**.

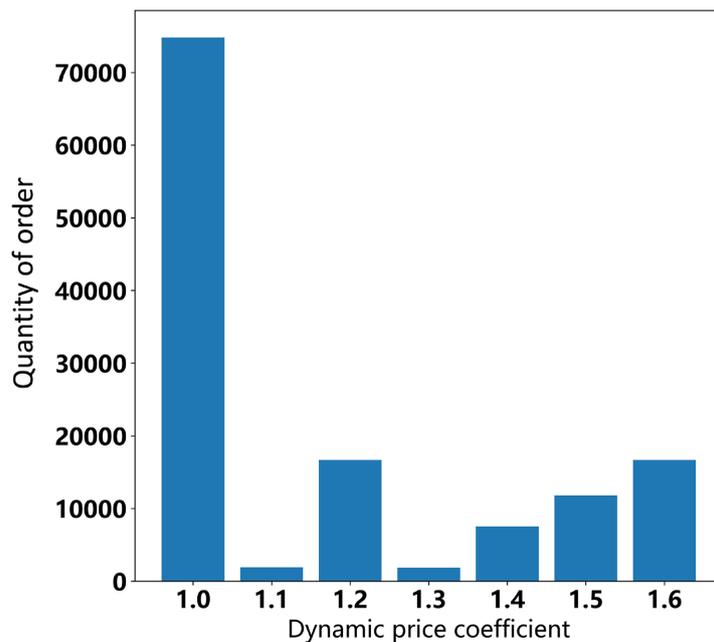
From **Figure 1**, it can be found that passenger demand continues to decrease from 12 midnight to 3 am, passenger demand reaches the lowest at 4 am, and then begins to increase slowly. Passenger demand from 6 am to 7 am it increases sharply, reaching the peak at 8 o’clock and decreasing at 9 o’clock. Therefore, we regard 6 - 9 as the morning peak time, and for the same reason, we regard 16 - 18 as the evening peak time.

According to Section 2.1 we divide orders into seven types based on dynamic price multipliers as shown in **Figure 2**.

It can be seen from **Figure 2** that orders with a dynamic pricing multiplier of 1.0 accounted for the majority, and the rest accounts for a small part. This is in line with the laws of the market, because the dynamic price coefficient will gradually stabilize with the continuous changes of supply and demand. Therefore, it is especially important for drivers to pursue orders with highly dynamic



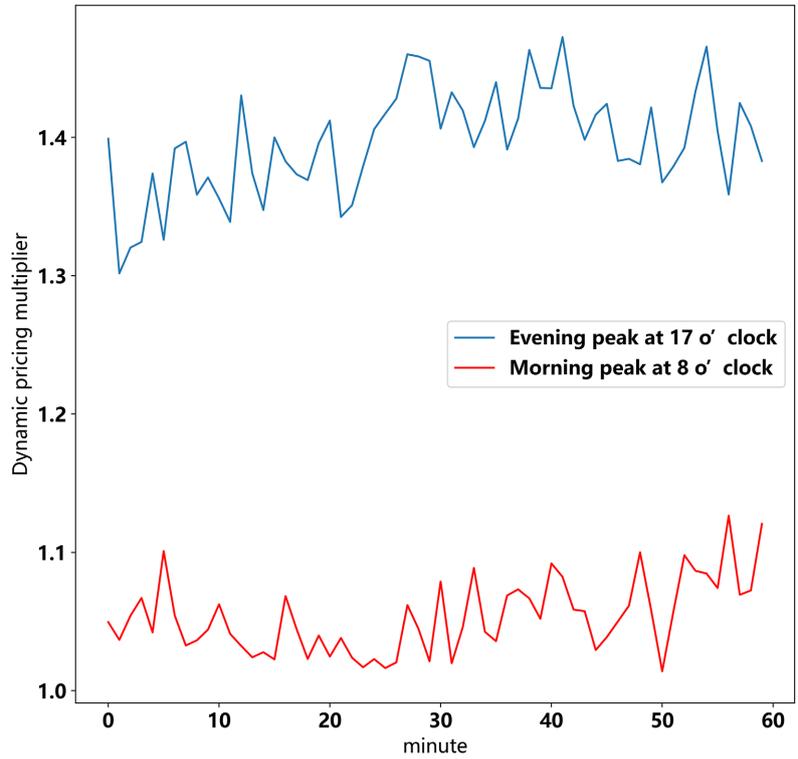
**Figure 1.** Passenger order data for one day.



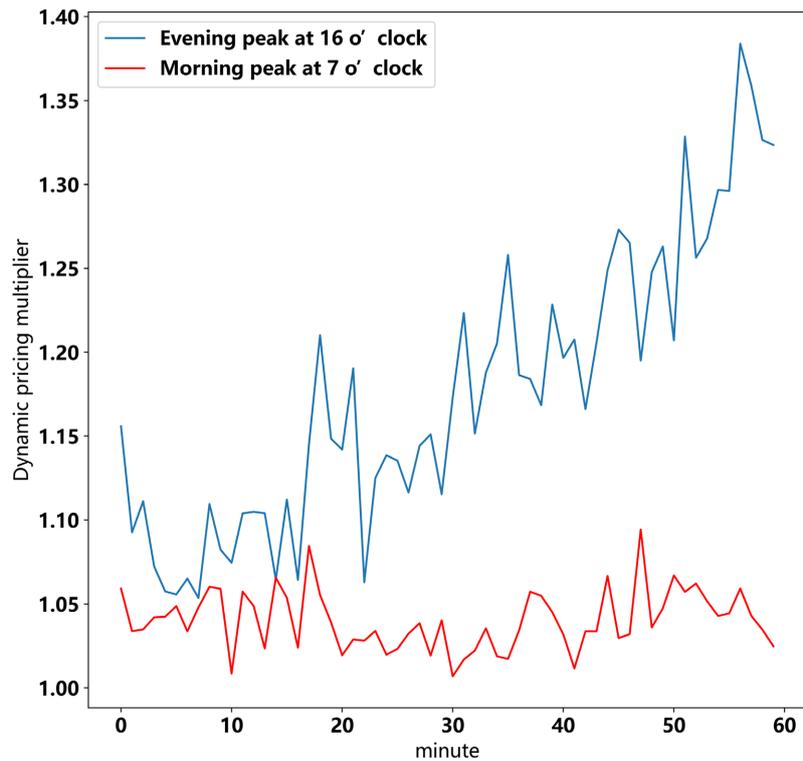
**Figure 2.** Order category.

pricing. From the above results, in order to verify the importance of dynamic pricing, we conduct a temporal and spatial analysis of dynamic pricing multiplier.

From **Figure 3(a)** and **Figure 3(b)**, it can be found that the dynamic pricing multiplier of the morning peak is generally lower than that of the evening peak. Beyond that the dynamic price coefficient at 16 o'clock in the evening peak is an upward trend and reaches dynamic equilibrium at 17 o'clock. Therefore, according to the above process, it can be found that the dynamic price multiplier



(a)



(b)

**Figure 3.** Dynamic price coefficient change graph in the morning and evening peaks. (a) Morning peak at 8 o'clock and evening peak at 17 o'clock; (b) Morning peak at 7 o'clock and evening peak at 16 o'clock.

will be affected by time. However, from **Figure 2**, orders with high dynamic price multiplier only account for a small part. As can be seen from the previous analysis, there may be areas with high dynamic price multiplier and areas with low dynamic price multiplier. Hence, we continue to do spatial analysis on order data.

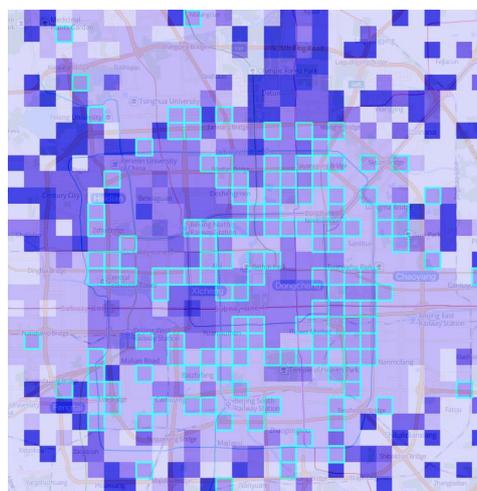
**Figures 4(a)-(c)** are price multiplier distribution maps in Beijing, which are respectively represented as low price, medium price and high price.

According to **Figure 4(b)** and **Figure 4(c)**, it can be found that in the city center, the high-price multiplier area is far less than the medium-price multiplier area. However, it can be found that there are more areas with high dynamic prices outside the urban areas than within the urban areas. The possible reason is that drivers are reluctant to go to these places because of the higher cost of picking up passengers. In this case, it is difficult for passenger to hail a car in the remote areas.

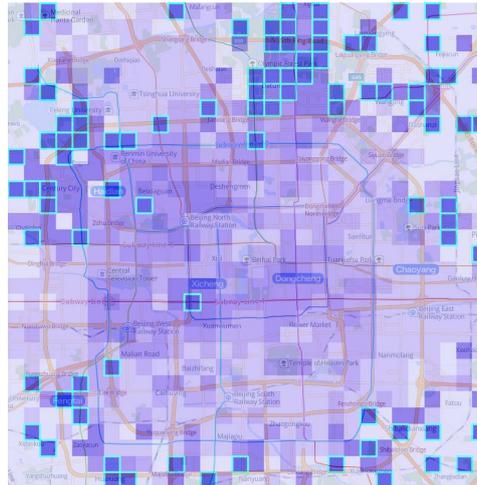
As can be seen from **Figures 5(a)-(c)**, the dynamic price coefficient fluctuates differently in different regions. In these three regions, it can be found that



(a)

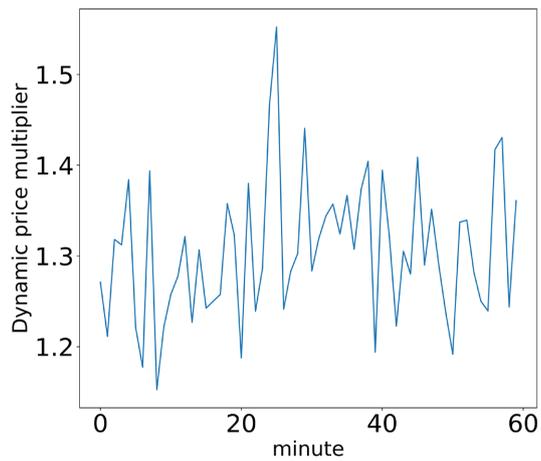


(b)

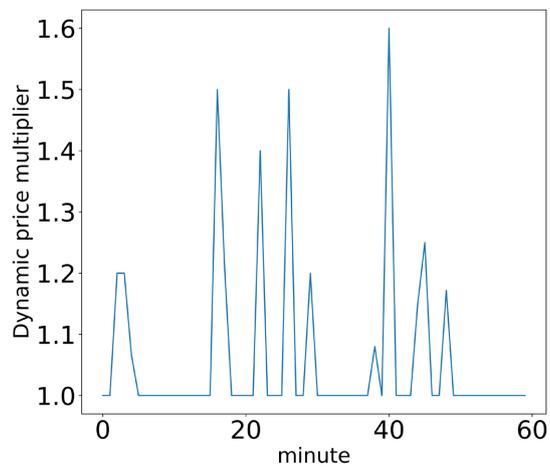


(c)

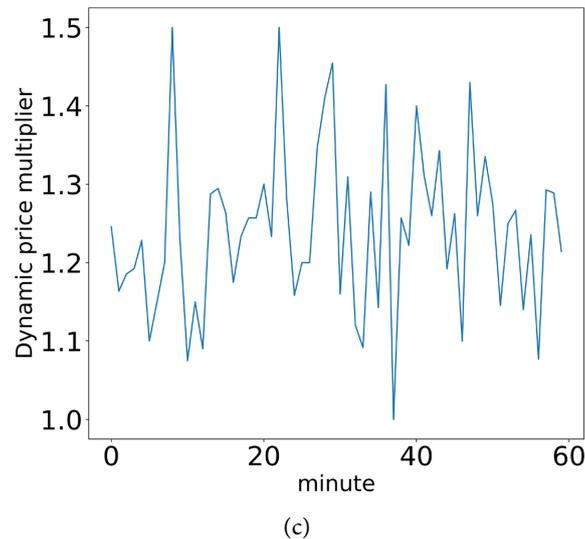
**Figure 4.** Beijing price multiplier distribution map. (a) Beijing price multiplier distribution: midprice during the evening peak; (b) Beijing price multiplier distribution: low price during the evening peak; (c) Beijing price multiplier distribution: high price during the evening peak.



(a)



(b)



**Figure 5.** Dynamic price coefficient fluctuation chart. (a) Dynamic price coefficient change graph in mid-price area; (b) Dynamic price coefficient change graph in low price area; (c) Dynamic price coefficient change graph in high price area.

the region with low dynamic price coefficient fluctuates steadily, while the region with high dynamic price fluctuates sharply. Therefore, it can be considered that not all orders in regions with low dynamic prices are low, but the probability of receiving high price multiplier orders is particularly low. In the high dynamic ratio area, not all orders are of high quality but the probability of receiving high price multiplier orders is particularly large.

In the above process, we draw three conclusions that 1) The dynamic pricing has an effect on driver's revenue that the higher the dynamic coefficient, the more the driver earns. 2) According to the division of historical data, Beijing can be divided into three regions, namely, low price area, medium price area and high price area. 3) In the grid of high price area, not all orders are orders with high dynamic price coefficient, but the probability of drivers receiving high price multiplier orders is extremely high. In the low-price grid, not all orders are low dynamic price coefficient orders, but the probability of drivers receiving low price multiplier orders is extremely high. Hence, it is necessary to introduce dynamic pricing into our model as a consideration factor.

### 3. MDP Model in RoD

An MDP is described by tuples  $(S, A, P, R, \delta)$ , where  $S$  stands for the state space,  $A$  denotes the allowable actions,  $R$  is the collects rewards,  $P$  defines a state transition matrix, and  $\delta$  is the discounts factor.

#### 3.1. MDP for Ride-Hailing Drivers

In this section, we will develop an MDP model for ride-hailing drivers' random passenger seeking process. Notations which will be used in the subsequent analysis are listed in **Table 1**.

**Table 1.** Major variables in MDP.

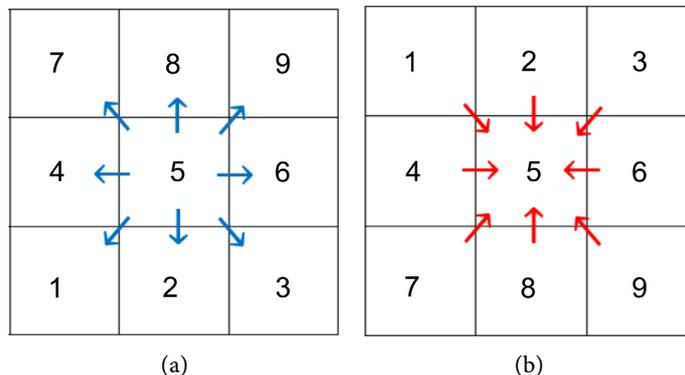
Variable	Explanation
$l$	Index of the current grid
$t$	Current time
$d$	Incoming direction to the current grid
$s$	State, $s = (l, t, d)$
$S$	State space, The collection of all the states
$a$	Action
$A$	Action space
$t_{seek}(l)$	Time spent on seeking for a passenger in grid $l$
$t_{driver}(l, k)$	Time spend on driving from grid $l$ to $k$
$d_{seek}(l)$	Distance traveled when seeking for a passenger in grid $l$
$d_{driver}(l, k)$	Distance traveled for moving from grid $l$ to grid $k$
$P_{pickup}(l)$	The probability that the driver can find out passenger in grid $l$
$P_{dest}(l, k)$	The probability a passenger picked up in grid $l$ wants to go to grid $k$
$D_y(l, k)$	The dynamic price from $l$ to $k$
$p_k$	Dynamic price coefficient
$\alpha$	Coefficient of fuel consumption

### 3.2. States and Actions

In our MDP model, the state  $s = (l, t, d)$  is composed of three components, namely,  $l$  represents the ID of the grid, current time  $t$ , and  $d$  indicates that the driver enters the current grid from different grid. Note that we will divide Beijing into 900 grids, hence  $l \in L = \{1, \dots, 900\}$ . In our model, we intend to simulate an hour of MDP, so we set  $T = 60$ .  $d$  represents the direction in which different grids enter the current grid where  $d \in D = \{\emptyset, \nearrow, \uparrow, \nwarrow, \leftarrow, \circlearrowleft, \rightarrow, \searrow, \downarrow, \swarrow\}$ . We use ten numbers (1 - 9) to index these signs, which is illustrated in **Figure 6**. Index 0 indicates that the driver dropped off a passenger, index 5 denotes that the driver does not have any arriving direction. Finally, a state in our MDP model can be written as  $s = (l, t, d)$ . The maximum number of states in our model can be calculated as:  $|L| \times |T| \times |d| = 540000$ . Nevertheless, the actual number of useful states is much less than this.

In decision-making states, each driver has nine actions to choose. Driver can choose an action from nine allowable actions to move neighbor grid or stay in current grid. Formally, it can be represented as:  $a \in A$ ,

$A = \{\searrow, \downarrow, \swarrow, \rightarrow, \circlearrowleft, \leftarrow, \nearrow, \uparrow, \nwarrow\}$ . We use nine numbers (1 - 9) to index these actions, which are shown in **Figure 6(b)**. The number five stands for staying in the current grid. From **Figure 6(a)** and **Figure 6(b)**, we can find that the index of the action and the index of  $D$  add up to ten (not including index 0), therefore, when the driver randomly chooses an action, the "incoming direction" of the next grid of  $d$  can be calculated as  $d = 10 - a$ .



**Figure 6.** An illustration of the direction of MDP model. (a) Possible actions; (b) Arrival direction.

### 3.3. State Transition

When the driver selects an action, there are two possible conditions. If the driver fails to pick up [13] a passenger, the driver will continue to select an action so as to go to neighbor grid to find passengers. If the driver successfully picks up a passenger, the driver will transport the passenger to the destination. Because of fuel consumption, the driver who fails to find passengers will have negative earnings.

**Figure 7** illustrates the aforementioned state transition process. The current state of the driver is  $s = (i, t, d)$ . From **Figure 7**, it can be known that the driver selects an action, there are two possible conditions.

The first possible condition is that the driver successfully picks up a passenger in grid  $j$  after  $t_{seek}(j)$ . In the circumstances, the customer may select to go to one of the grid cells as destination (denotes as  $k$ ) at a probability  $P_{dest}(l, k)$ . The driver will go to grid  $k$  to drop off passengers. Hence, we use  $t_{driver}(j, k)$  to represent the time that spends on driving from grid  $j$  to  $k$ . The driver will earn a fare of  $D_y(l, k)$  Yuan, where represents the expected profits from grids  $j$  to  $k$ . The driver will start seeking from  $k$  again. Thus, the current state of the driver in grid  $k$  is  $s = (j, t + t_{seek}(j) + t_{driver}(j, k) + t_{driver}(j, k), 0)$ .

The second possible condition is that the driver fails to pick up a passenger in grid  $j$  after  $t_{seek}(j)$ . Then the driver will take an action from nine allowable actions to go to other grid so as to find passengers. For instance, we assume that the driver acts  $a = 9$  ( $\nearrow$ ), then the ride-hailing drivers will end up in state  $s' = (j, t + t_{seek}(j), 1)$  (from the bottom left grid,  $\nearrow$ ).

To sum up, the driver in any state  $s_0 = (i, t, d)$ , the driver takes an action to go from current grid to another grid. With the probability  $P_{pickup}(j) \times P_{dest}(j, k)$ ,  $k \in L$ , the driver will transition to state

$$s_1 = (j, t + t_{seek}(j) + t_{driver}(i, j) + t_{driver}(j, k), 0), \text{ and get a reward} = D_y - \alpha(d_{driver}(j, k) + d_{seek}(j) + d_{driver}(i, j)).$$

With the probability  $1 - P_{pickup}(j)$ , the state of driver will transition to state

$$s_2 = (j, t + t_{seek}(j) + t_{driver}(i, j), 10 - a).$$

Get a Reward =  $-\alpha(d_{seek}(j) + d_{driver}(i, j))$ , which is negative.

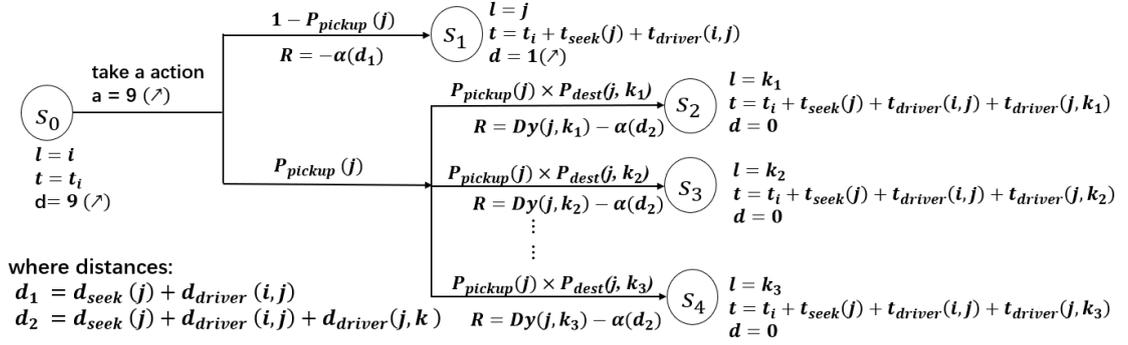


Figure 7. MDP state transition.

### 3.4. MDP Parameters

The pickup probability of passengers  $P_{pickup}(j)$  calculation. Firstly, we divided the map of Beijing into  $30 \times 30$  grids. Secondly, we projected passenger pick-up and drop-off point data and GPS trajectory points of empty drivers onto the map of Beijing. Finally, we use spatial connection to match points to each grid, thus we can know that the number of pick-up and drop-off points in each grid and the number of empty cars passing through the grid. Hence,  $P_{pickup}(j)$  is calculated by dividing the pickup points of a grid by the number of empty cars passing through the grid. Let  $n_{pickup}$  be the number of pickup points in the grid and  $n_{pass}$  the number of idle taxis crossing the grid, therefore the pickup probability can be expressed as:

$$P_{pickup}(j) = \frac{n_{pickup}(j)}{n_{pass}(j)} \tag{2}$$

The passenger destination probability  $P_{dest}(j, k)$  calculation. When the driver successfully picks up a passenger in grid  $j$  after  $t_{seek}(j)$ , the passenger will go to the  $k \in L$  grid with probability  $P_{dest}(j, k)$ . Hence, in order to calculate the destination probability, we first calculate the number of passenger orders from grid  $j$  to grid  $k$ , (denote as  $n_{j \rightarrow k}(j)$ ). Secondly, we use  $n_{j \rightarrow k}(j)$  to divide by  $n_{pickup}$ . Hence, finally the  $P_{dest}(l, k)$  can be written as:

$$P_{dest}(j, k) = \frac{n_{j \rightarrow k}(j)}{n_{pickup}(j)} \tag{3}$$

The driving time  $t_{driver}(j, k)$  denotes time spend on driving from grid  $j$  to grid  $k$ . So, we are able to take the average of all driving times from grid  $j$  to grid  $k$  as an approximation of the  $t_{driver}(j, k)$ . The driving distance  $d_{driver}(j, k)$  is also calculated as the average distance from grid  $j$  to grid  $k$ .

The seeking time  $t_{seek}(j)$ , we calculate all drivers' search time for passengers and it is about 300 meters/min, hence, we set the searching time  $t_{seek}(j) = 1$ . The searching distance  $d_{seek}$  is set as 300 meters.

### 3.5. Solving MDP

In our MDP model, our goal is to maximize revenue in the current time slot.

From Sub 3.2, we simulated the driver's one-hour MDP search process, therefore, when  $t > 60$ , our model will stop. For each  $a$ , if the driver takes an action  $a$  in state  $s$ , the  $V'(s, a)$  function represents the maximum expected return in the current time slot.  $V(s)$  is the maximum expected return of state  $s$ . Finally,  $V'(s, a)$  is calculated as:

$$\begin{aligned} V'(s, a) = & \left(1 - P_{pickup}(j)\right) \times \left[-\alpha(d_{seek}(j) + d_{driver}(i, j) + V(s_1))\right] \\ & \times P_{pickup}(j) + \sum_{k=1}^{|L|} P_{dest}(j, k) \times \left[D_y(j, k) \right. \\ & \left. - \alpha(d_{seek}(j) + d_{driver}(i, j) + d_{driver}(j, k)) + V(s_2)\right] \end{aligned} \quad (4)$$

Hence, the optimal policy  $\pi$  is defined as follows:

$$\pi(s) = \arg \max V'(s, a) \quad (5)$$

$$V(s) = V''(s, \pi(s)) \quad (6)$$

The pseudocode of DP algorithm is given in **Algorithm 1**. According to the data analysis, we use algorithms to assign values to the dynamic prices of each grid. When the dynamic price coefficient is determined. The algorithm first generates a state  $s = (l, t, d)$ ,  $s$  will try all actions in order to find the best action (line 5 - line 8). When the driver takes an action, the total expected reward of state is calculated using Equation (1) (line 9). Then get the best action based on expectations (line 10 - line 11), finally, output our optimal strategy (line 16).

**Algorithm 1** has time complexity, because  $|D|$  and  $|A|$  are small constant numbers, the complexity can be rewritten as  $O(|T| \times |L|)$ . In the same way, the space complexity is  $O(|T| \times |L|)$ .

---

#### Algorithm 1. Dynamic programming algorithm

---

**Require:**  $L, T, A, P_{pickup}, P_{dest}, d_{seek}, d_{driver}, t_{seek}, t_{driver}, D_y, p_k$

**Ensure:** The optimal policy  $\pi$

1:  $p_k$  is a  $|L| \times 1$ ,  $V$  is a  $|L| \times |T|$  matrix;  $p_k \leftarrow 0, V \leftarrow 0$

2: **for**  $|L|$  to 1 **do**

3:      $p_L \leftarrow p_k$

4: **end for**

5: **for**  $t = T - 1$  to 1 **do**

6:     **for**  $l = 1$  to  $|L|$  **do**

7:         **for**  $d$  in  $\text{len}(d)$  **do**

8:             **for**  $a \in A$  **do**

9:                  $a_{\max} \leftarrow \text{find a } \max V'(s, a) \text{ can be computed by Equation (4)}$

10:                  $\pi \leftarrow a_{\max}$

11:                  $V(s) \leftarrow \max V'(s, a_{\max})$

12:             **end for**

13:         **end for**

14:     **end for**

15: **end for**

16: **Return**  $\pi$

---

## 4. Evaluation

In this section, we will evaluate our MDP model from two aspects: 1) Is there any change in the way of finding passengers after using our model? 2) Is the driver's income greatly improved than they did before after using our model?

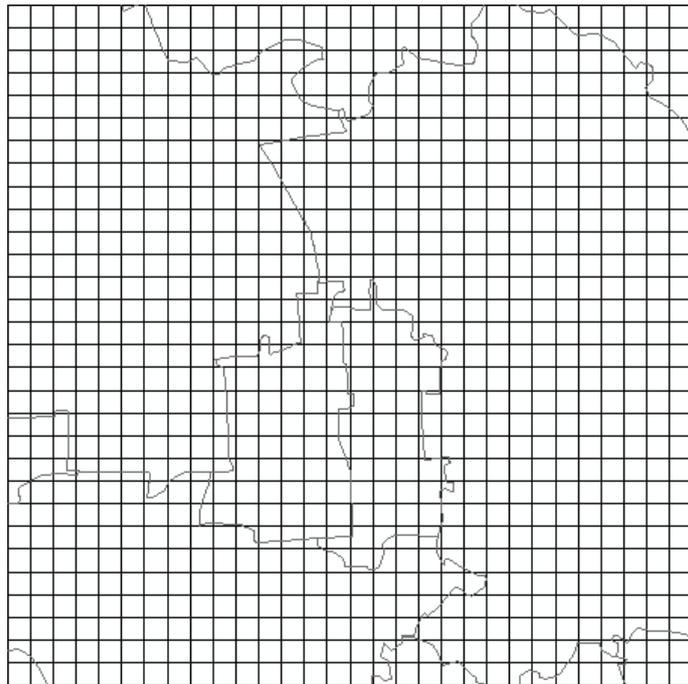
### 4.1. Seeking Strategy Evaluation

We conduct simulation based on MDP model and compare the simulation results with the passenger search methods of real drivers to verify the effectiveness of our recommendation.

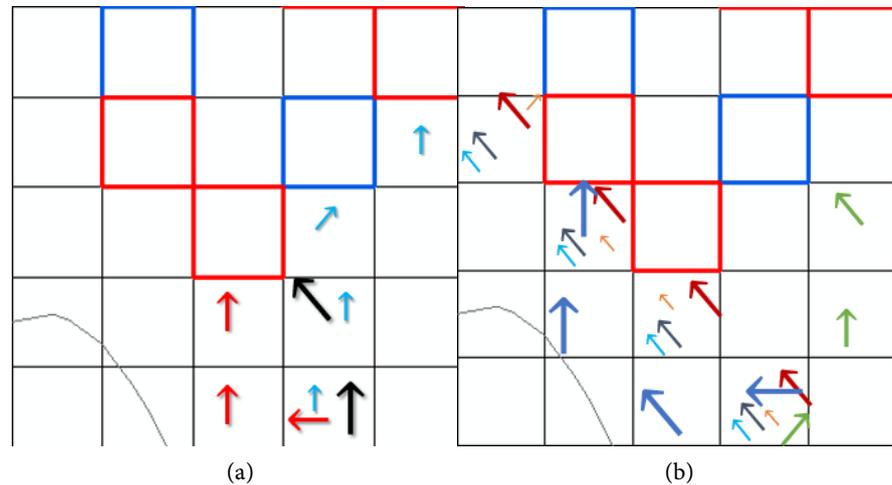
**Figure 8** stands for a map of Beijing. We will select two areas which are Fengtai Qu and Xicheng Qu in order to evaluate drivers' seeking strategies.

**Figure 9(a)** presents the search strategies of real drivers in the suburbs during the evening peak (The arrows of different sizes represent different drivers. Red and blue respectively indicate high price multiplier area, low price multiplier area.). It can be seen from **Figure 9(a)** that the majority of drivers miss local pickup opportunities because they are in a hurry to move to urban areas. It can also be found that a few drivers could find passengers by their experience, but there are too few experienced drivers to solve the problem of passenger hailing.

**Figure 9(b)** presents the search strategies of MDP agent in the suburbs during the evening peak (The arrows of different sizes represent different agents.). From **Figure 9(b)**, it can be seen that our algorithm will preferentially recommend agents to search for passengers in the local area. Meanwhile, in the search for the passengers, our algorithm will give priority to areas with high dynamic prices. By comparing **Figure 9(b)**, we can find that after using our algorithm, agents



**Figure 8.** The grid division of the Beijing.



**Figure 9.** Comparison between simulated and real seeking strategies. (a) MDP agent seeking strategies; (b) Real driver seeking strategies.

will give priority to search for passengers locally, at the same time, our algorithm can alleviate the problem that it is difficult for passengers to take a taxi in the suburbs.

**Figure 10** presents the search strategies of MDP agent in the downtown during the evening peak (green background, pink and red background indicate low price multiplier, medium price multiplier and high price multiplier). It can be found from the above that our algorithm will recommend the driver to move slowly from the area of low dynamic price to the area of high dynamic price. In particular, our algorithm will not dispatch drivers from all regions to regions with the highest dynamic prices, but dispatch drivers in some regions to regions with the highest prices. The possible reason for this is that the use of dynamic pricing to dispatch drivers can prevent all drivers from flocking to a specific area with a high price multiplier, causing an oversupply situation in this area.

## 4.2. Revenue Evaluation

From Section 4.1, it can be known that when the driver uses our algorithm to search for passengers, our algorithm will recommend the driver to go to areas with high dynamic prices to obtain high-quality orders. Hence, we intend to use the quality of order acquisition to evaluate driver's revenue efficiency. The driver's revenue efficiency calculation, we divide the driver's income per order by the driver's working hours, the driver's working time is the sum of the time spent looking for customers and spent completing an order.

**Figure 11(a)** shows the profit comparison diagram of drivers before and after using our algorithm (Pink represents the benefits obtained by drivers in areas with medium price). From **Figure 11(b)**, it can be seen that driver's average profit per minute after using our algorithm is higher than before. Before using our algorithm, the average driver's minimum revenue per minute is 0.71 yuan. After using our algorithm, the maximum income of the average driver per minute is 0.8336 yuan. The maximum yield of our model can be increased to 17%.

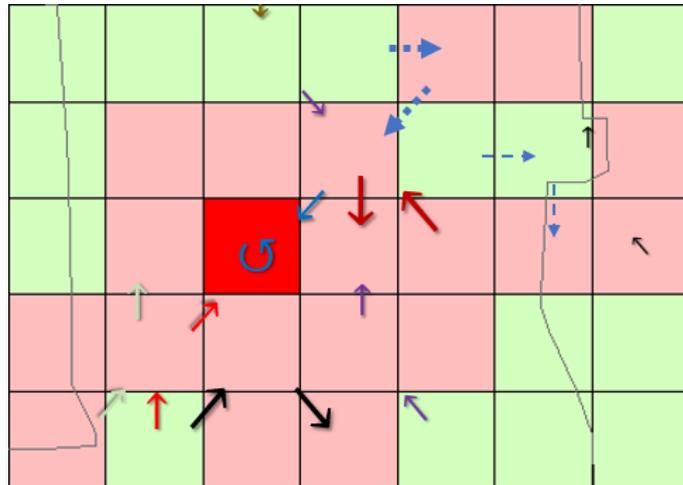
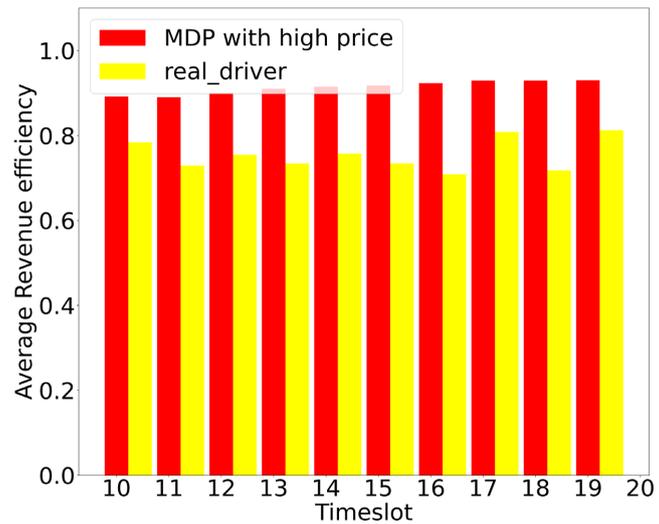
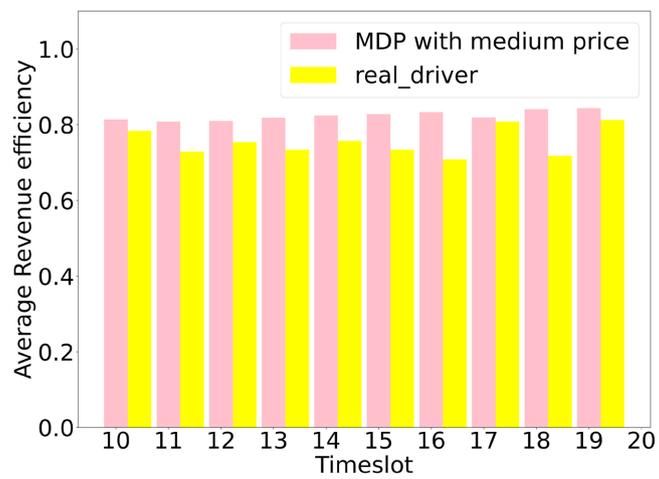


Figure 10. MDP agent seeking strategies in downtown.

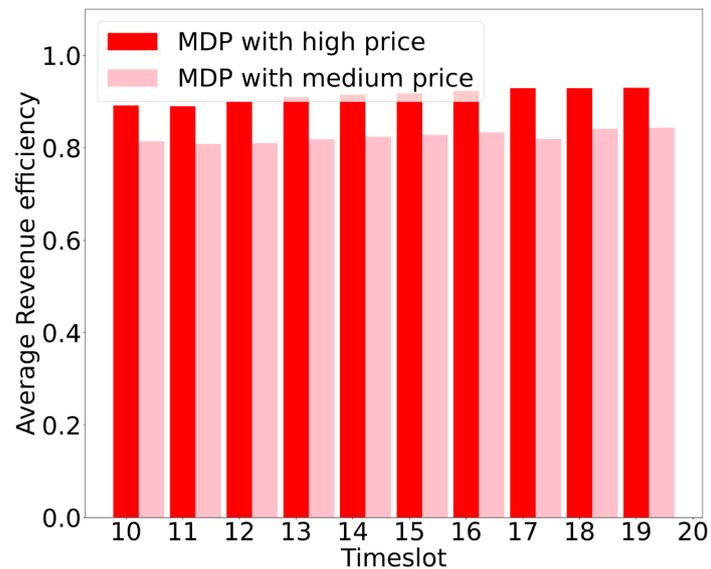


(a)



(b)

Figure 11. MDP agent and real driver's average revenue comparison. (a) MDP vs real driver's average revenue efficiency; (b) MDP vs real driver's average revenue efficiency.



**Figure 12.** MDP with high price vs MDP with medium price.

**Figure 11(b)** shows the profit comparison diagram of drivers before and after using our algorithm (red represents the benefits obtained by drivers in areas with high price). From **Figure 11(a)**, it can be seen that driver average profit per minute after using our algorithm is higher than before. Before using our algorithm, the average driver's minimum revenue per minute is 0.71 yuan. the maximum income of the average driver per minute is 0.91 yuan. The maximum yield of our model can be increased to 28%.

**Figure 12** shows the benefits obtained by the driver in finding passengers in different dynamic price regions after using our algorithm. Red represents the benefits obtained by drivers in areas with high price and pink represents the benefits obtained by drivers in areas with medium price. The maximum yield of our model can be increased to 9%.

## 5. Conclusions and Future Work

In this paper, we design a Markov Decision Process (MDP) model to answer “how to use dynamic prices to help drivers in seeking for passengers”. We first show the importance and need to do that by analyzing real service data. We then design a Markov Decision Process (MDP) model based on passenger order and car GPS trajectories datasets, and take into account dynamic prices in designing rewards. Results show that, on the one hand, when searching for passengers in the suburbs, our model can guide drivers to areas with high dynamic prices in front of them and improve drivers' utilization rate. On the other hand, when searching for passengers in the urban area, our model will also guide the driver to slowly cruise to the dynamic high price area. In the dynamic high price zone, our model will dispatch drivers reasonably, which can prevent all drivers from flocking to a specific area with a high price multiplier, causing an oversupply situation in this area. Finally, compared with things with the drivers before using

the model, the maximum yield after using it can be increased to 28%.

For future work, we will introduce multi-agent reinforcement learning to study the influence of dynamic price on driver seeking; at the same time, probabilistic model is introduced to simulate the fluctuation of dynamic multiplier.

### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

### References

- [1] Guo, S., Chen, C., Wang, J., Liu, Y., Xu, K., Yu, Z., Zhang, D. and Chiu, D.M. (2020) Rod-Revenue: Seeking Strategies Analysis and Revenue Prediction in Ride-on-Demand Service Using Multi-Source Urban Data. *IEEE Transactions on Mobile Computing*, **19**, 2202-2220. <https://doi.org/10.1109/TMC.2019.2921959>
- [2] Garg, N. and Ranu, S. (2018) Route Recommendations for Idle Taxi Drivers: Find Me the Shortest Route to a Customer! *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, New York, July 2018, 1425-1434. <https://doi.org/10.1145/3219819.3220055>
- [3] Gao, Y., Jiang, D. and Xu, Y. (2018) Optimize Taxi Driving Strategies Based on Reinforcement Learning. *International Journal of Geographical Information Science*, **32**, 1677-1696. <https://doi.org/10.1080/13658816.2018.1458984>
- [4] Xu, Z., Men, C., Li, P., Jin, B., Li, G., Yang, Y., Liu, C., Wang, B. and Qie, X. (2020) When Recommender Systems Meet Fleet Management: Practical Study in Online Driver Repositioning System. *Proceedings of the Web Conference 2020*, New York, April 2020, 2220-2229. <https://doi.org/10.1145/3366423.3380287>
- [5] Powell, J.W., Huang, Y., Bastani, F. and Ji, M. (2011) Towards Reducing Taxicab Cruising Time Using Spatio-Temporal Profitability Maps. *Proceedings of the 12th International Conference on Advances in Spatial and Temporal Databases, SSTD'11*, 242-260. [https://doi.org/10.1007/978-3-642-22922-0\\_15](https://doi.org/10.1007/978-3-642-22922-0_15)
- [6] Rong, H., Zhou, X., Yang, C., Shafiq, Z. and Liu, A. (2016) 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*, New York, October 2016, 2329-2334. <https://doi.org/10.1145/2983323.2983689>
- [7] Shou, Z., Di, X., Ye, J., Zhu, H., Zhang, H. and Hampshire, R.C. (2019) Optimal Passenger-Seeking Policies on E-Hailing Platforms Using Markov Decision Process and Imitation Learning. arXiv:1905.09906 [cs.LG] <https://doi.org/10.1016/j.trc.2019.12.005>
- [8] Yu, X., Gao, S., Hu, X. and Park, H. (2019) A Markov Decision Process Approach to Vacant Taxi Routing with E-Hailing. *Transportation Research Part B: Methodological*, **121**, 114-134. <https://doi.org/10.1016/j.trb.2018.12.013>
- [9] Yuan, J., Zheng, Y., Zhang, L., Xie, X. and Sun, G. (2011) Where to Find My Next Passenger. *Proceedings of the 13th international conference on Ubiquitous computing*, New York, September 2011, 109-118. <https://doi.org/10.1145/2030112.2030128>
- [10] Yuan, N.J., Zheng, Y., Zhang, L. and Xie, X. (2013) T-Finder: A Recommender System for Finding Passengers and Vacant Taxis. *IEEE Transactions on Knowledge and Data Engineering*, **25**, 2390-2403. <https://doi.org/10.1109/TKDE.2012.153>
- [11] Guo, S., Chen, C., Liu, Y., Xu, K. and Chiu, D.M. (2018) Modelling Passengers'

- Reaction to Dynamic Prices in Ride-on-Demand Services: A Search for the Best Fare. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **1**, 1-23. <https://doi.org/10.1145/3161194>
- [12] Xie, X., Zhang, F. and Zhang, D. (2018) PrivateHunt: Multisource Data-Driven Dispatching in For-Hire Vehicle Systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **2**, 1-26. <https://doi.org/10.1145/3287074>
- [13] Yan, L., Shen, H., Li, Z., Sarker, A., Stankovic, J.A., Qiu, C., Zhao, J. and Xu, C. (2018) Employing Opportunistic Charging for Electric Taxicabs to Reduce Idle Time. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **2**, 1-25. <https://doi.org/10.1145/3287076>