Research on No-load Decision of Airport Taxi Based on Optimization Model

Xiang Yifei

School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei, 430072

Keywords: no-load problem; parking space; optimization model; Monte Carlo simulation

Abstract: With the facilitation of transportation, taxis have become one of the main modes of transportation, but no-load taxis will cause damage to revenue. Based on this, this article first determines the total passenger flow at the airport and the passenger flow in a day based on factors such as seasons and flights distribution law, so as to get the driver's decision plan. Then combine the processed data to estimate the parameters, and solve the model to get the no-load decision plan of the taxi. Compared with the actual taxi detention situation, the model accuracy rate of the scheme is close to 70%. After that, the influence of factors such as season, the probability of choosing a taxi, the number of pick-up points, and the average pick-up time on the driver's decision is studied, and the relevant factors and the practical significance of the model decision are given. Finally, three types are introduced. The common pick-up point setting mode is to establish a multi-objective optimization model. Through a large number of Monte Carlo simulations, it is finally found that the parallel-departure multi-lane departure mode is better than the other departure modes. In the best case, there are 12 berths and the actual berths in Hongqiao Airport are the same.

1. Introduction

Taxi is one of the main means of transportation for passengers arriving at and departing from the airport. As most domestic airports separate the delivery (departure) and pick-up (arrival) channels from the airport, drivers will face two options [1]. One is to go to the designated storage pool in the arrival area to wait in line for the passengers to return to the city, but this requires a certain time cost, and the waiting time is determined by the number of taxis in the storage pool and the number of passengers [2]; the second is to return to the urban area so that the taxi driver will bear the no-load cost and may lose the potential passenger income [3]. The number of flights arriving in a certain period of time and the number of vehicles in "storage pool" is certain information that can be observed by the driver [4]. Usually the driver's decision is related to his personal experience, but in reality, there are many determination and uncertainty factors that affect the taxi driver's decision, and their correlations are different. The effects are also different.

2. Establishment and solution of airport taxi decision model

2.1 Establishment of airport taxi decision model

2.1.1 Total passenger flow at the airport and its time distribution

The total passenger flow at the airport Tol is determined by the total number of flights on the day Nf and is often affected by the seasonal factor S and the holiday factor H. It is assumed that the impact of S and H on the total passenger flow is multiplicative, k is a constant coefficient, and the value is 1. The average number of passengers on a flight, the total passenger flow can be expressed as: To1 = kHSN.

After determining the total passenger flow, further consider the distribution of passenger flow during the day, taking into account that the flight frequency is artificially set to have a certain periodicity, and there is no situation where the flight is too thin or dense, so the conventional parking is used. The loose distribution to approximate the passenger flow distribution is not in line with the actual situation. If the flight data of the airport has been collected, only the time needs to be segmented, and the change law of the arrival passenger flow over time can be obtained by the kernel fitting method. Note that the density function of the obtained passenger flow distribution is c.

2.1.2 Passenger's choice of transportation mode

Considering the diversity of airport transportation, arriving passengers will first choose the mode of travel (such as taxi, subway, bus). Assuming that the probability of passengers choosing to take a taxi is P_0, we can know that the unit time is selected. The number of passengers who choose to take a taxi follows the binomial distribution with a multiple of E (C) .The number of passengers who take a taxi per unit time is a random variable v. The probability distribution of v is as follows:

$$P\{v=x\} = C_{E(C)}^{x} p^{x}(1-p)^{E(C)-x}, \quad \sim B(p)$$

2.2 Solution of the airport taxi decision model

First, select a crowded area near the departure point of the airport as a "node", and stipulate that the pick-up point for taxis leaving the airport is only at the node, that is, regardless of the situation where the taxi stops in the middle of the road section. After selecting the node, you need to determine the cost of the airport taxi to stay option, so as to judge the driver's decision based on the cost.

2.2.1 Cost of staying in line at the airport

For taxis located at the airport, assume that there are already two taxis in the storage pool at this time, the number of pick-up points is N, and the average time for one pick-up is t_av, combined with the unit time obtained in the previous section, choose the number of passengers V for the taxi. According to the "first come, first served" principle of the storage pool, if the passenger's entry speed is greater than the maximum number of pick-ups per unit of time at the pick-up point, the driver does not need to worry about the problem of no one riding, and it is easy to get the taxi to the storage at this time. The expected waiting time W for the pool to receive passengers.

2.2.2 Cost of travelling to downtown

On the way from the airport to the i-node pick-up, the taxi needs to bear the no-load cost hi, at the same time, it will lose the potential gain w of pick-up at the airport and consume the waiting time cost t_i after arriving at the node. The node where the passenger cost is as low as possible. The cost expression of the taxi from the airport to the i-th node is as follows:

$$E_i = h_i + \beta w + t_i$$

In fact, the waiting time t is difficult to determine and is often judged by the driver empirically, but in general, the waiting time of the driver in the urban area is far less than the waiting time in the airport storage pool.

2.3 Taxi Decision

Compare the cost of staying in line at the airport with the cost of leaving the airport to a certain node. Experienced old drivers will choose the lower cost option, so the decision to get a taxi to stay is as follows:

$$\begin{cases} E_i < w & ,leave \\ E_i \geq w & ,stay \end{cases}$$

3. Establishment and solution of no-load decision model

3.1 Establishment of no-load decision model

In order to give the driver's specific decision-making plan, we found about 8 million GPS trajectory data of Shanghai taxis moving within 24 hours [5]. Through data cleaning and preprocessing screening, 600,000 meaningful data were obtained and calculated. We will calculate the actual data based on the model brought into to get each driver's decision plan. Compared with the real plan, it is found that the model prediction rate given in this article reaches 70%, which is much higher than the 5% effective prediction rate of the RBF neural network.

3.1.1 Data collection and preprocessing

On the CSDN website, we collected data on some taxis in Shanghai on February 10, 2007. The data included the vehicle number, vehicle GPS position (about 40 payment inspections), vehicle speed, vehicle passenger status, signal angle and other information, a total of 2217 taxis, a total of 4 million data. Some data and data format are as follows:

Table 1 Some taxi data

Taxi number	date	Detection time	longitude	latitude	Satellite angle	Speed	Passenger situation
105	2007/2/20	01:37:57	121.473300	31.231000	2	67	0
105	2007/2/20	01:38:13	121.474500	31.229300	8	67	0
105	2007/2/20	01:38:29	121.475300	31.228000	8	67	0
105	2007/2/20	01:38:47	121.475300	31.227100	6	35	0
105	2007/2/20	01:39:06	121.475000	31.227000	8	35	1
105	2007/2/20	01:40:07	121.471600	31.225800	2	12	1
105	2007/2/20	01:41:08	121.470800	31.225500	2	12	1

The storage pool is a point with a large range, and the actual location and length of the storage pool cannot be judged only by the location of the airport. In fact, taxi drivers have moved a large distance in the storage pool, and GPS data needs to be analyzed. The scope of the storage pool was determined through analysis. Through the collected driver passenger information and speed information, we selected multiple groups of passengers arriving near Hongqiao Airport to change their status from 1 to 0 (delivering passengers to the airport), and there were significant low-speed movement periods thereafter. The taxi data of (storage pool queuing), put multiple sets of GPS data on the satellite map and observe the actual geographic features in the satellite map. We finally determined that the range of the storage pool is [up = 31.197143; left = 121.351729; right = 121.3418130; down = 31.190568], in the rectangular rectangle, located southwest of the outside of the airport.

Among the 4 million GPS data of 2217 taxis that have been collected, redundant data that is irrelevant to the research target needs to be eliminated. The irrelevant data mainly includes the following categories:

1) Zombie cars: The GPS data shows that there are taxis that do not move within one day, and some even reside near the storage pool and have abnormal passenger information, so the data of all zombie cars need to be deleted.

2) Taxi without airport: In practice, not all taxis will go to the airport, and taxis that do not go to the airport obviously should not be included in the research scope of this question and need to be eliminated.

3) Taxi heading to the storage pool but always loading passengers: at the GPS position, heading to the storage pool but not empty, that is, the taxi does not go to the storage pool to queue for passengers. In fact, because the storage pool is often located underground, the GPS coordinates of the taxis on the ground at the same location are the same as the GPS coordinates in the underground storage pool, so the taxi data near the storage pool but always in the state of carrying passengers need to be removed. After processing, the requirements are finally met. The number of taxis is 318, with a total of 600,000 pieces of data.

3.2 Solution of the no-load decision model

3.2.1 Determination of model parameters

1) Total passenger flow and distribution

After searching the data, this article found the Shanghai Hongqiao Airport flight schedule data that is specific to every minute in 2016, but because the taxi GPS tracking information comes from 2007, in order to ensure timeliness, it is necessary to use the 2016 time information to check the flight time in 2007.

In February 2007, the throughput of Hongqiao Airport was 1.48 million. In 2016, the average monthly throughput was 4.08 million. By scaling and moving, it is estimated that the per-minute traffic density of the airport in 2007 is shown in Figure 1:

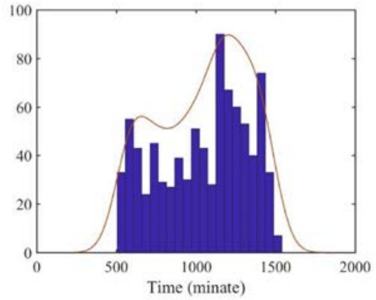

Figure 1 Changes in the flow of people at Hongqiao Airport over time

2) Seasonal factors

Taking the monthly throughput data of Hongqiao Airport from 2006 to 2016 for moving average calculation of the monthly seasonal factors, the seasonal factors are obtained through the calculation of the 10-year data.

3) Probability of choosing a taxi

Taking the probability that the passenger chooses to take the airport taxi as the transportation means p_0 is 0.08, according to the above conclusion, the probability distribution of the passengers is constant.

4) Node selection

Observe the driver loading points marked in the map. The passenger loading points near the airport are particularly concentrated in the three locations of Shanghai Zoo, Shanghai Studios, and Sinopec Gas Station, and the crowds are dense. Therefore, these three locations are selected as the roads near the airport network node, consider that the driver leaves the airport to pick up passengers at the above three places, and its latitude and longitude coordinates are as follows.

Table 2 Node latitude and longitude comparison table

Shanghai Zoo	Shanghai Studios	gas
[121.374768,31.196138]	[121.428907, 31.195134]	[121.467842,31.224222]

5) Cost of queuing for storage pool

According to the above model, we know the number of pick-up points in the single lane of Shanghai Hongqiao Airport (the number of cars that the pickup pool picks up at the same time) N = 4. The average boarding time t_{av} = 84 seconds, you can determine the taxi to the storage pool Time spent waiting in line to pick up passengers (x means x cars are waiting in front of this taxi)

$$w = \begin{cases} \frac{xt_{av}}{N} & v \geq \frac{N}{t_{av}} \\ \frac{xt_{av}}{vL} & v < \frac{N}{t_{av}} \end{cases}$$

6) Passenger cost to downtown

Calculate the distance between nodes to estimate the time it takes from the airport to each node, and plug it into the above model to determine the passenger cost of the taxi. β=-1

$$E_i = h_i + \beta w + t_i$$

3.2.2 Solution of taxi decision model

After determining various parameters, the model was used to simulate the choice of airport taxi drivers to stay and compare with actual data.

In the binary prediction, the prediction data is directly compared with the actual data to calculate the model prediction accuracy rate, which easily falls into the "false" accuracy rate. For example, when 270 out of 300 taxis choose to pick up passengers, the model predicts that all 300 cars will pick up, and the "false" accuracy rate will be 90%, but in fact the model does not explain the taxis leaving with no load. In order to resolve this contradiction, we use the strategy of grouping and adding.

Divide the actual passenger and unloaded taxis into two groups, and calculate the prediction accuracy rate for each group, respectively, as p, q, then the total prediction accuracy rate of the model is p + q-l. Based on this, the model operation results and the actual taxi stay status is shown in Table 3:

Table 3 Model run result detection

Actually	Model left	Actually	Model leaves	p	**q**	p+q-1
171	193	147	125	0.72	0.62	0.34

The accuracy rate of the results of the model comparison with the actual data reached 35%, and compared with 5% of the RBF neural network, it performed well.

3.2.3 Model sensitivity analysis to relevant factors

1) The influence of seasonal factors and passenger's choice of taxi probability on the model

For the seasonal factor, the passenger flow in different months is calculated by the seasonal factor, and the model is re-run. After calculation, the season has a limited impact on the model result, and the fluctuation range is generally within 10%. In winter, the number of taxis to choose to wait is the least, and the rental to choose to leave. The number of vehicles is the largest, reaching its peak in November, and the opposite is true in the spring and autumn seasons, which is inseparable from the seasonal change trend of the number of flights.

For the probability of choosing a taxi, p_0, due to the large subjectivity of the previous selection, changing the value will affect the distribution of passengers who choose to take a taxi at the airport and re-solve the model. The higher the average probability of choosing a taxi for passengers. The more the taxis derived from the model that choose to go to the storage pool line up, the fewer taxis they choose to leave. The influence of this factor on the model results satisfies the law of supply and demand in reality.

2) The impact of the number of pick-up points and the average pick-up time on the model

In the previous model, the number of pick-up points and the average pick-up time were directly determined as N = 4 and t_{av} = 40 seconds according to the collected data. Consider these two quantities as variable and change the values of these two variables. These two factors both affect the taxi waiting time and passenger waiting time, and their impact on the model results is similar, so they are also analyzed together. It can be found that when there are less than 5 parking pick-up points. For each additional pick-up point, the number of taxis staying will increase, and the number of taxis leaving will decrease correspondingly, and the change will be large. When there are more than 5 pick-up points, the change in the number of pick-up points will affect the taxis determined by the model. There is no significant effect on the distribution of stays, the reason is that the work of the pick-up point was saturated before this time, and when the pick-up point further increased, the actual passenger arrival rate was already less than the maximum passenger arrival rate acceptable by the pick-up point, in the unsaturated working state, showing the status of cars and others, it is expected that if the season is changed or the probability of a passenger choosing a taxi, the saturation point of the pick-up point will be shifted left or right accordingly. It can be seen that if the average off time has been increasing, the number of taxi choose to leave will continue to increase up to 100% tend to choose the number of taxi car pool to stay go pick up. The reservoir will continue to decrease until it tends to 0%, which is clearly in line with the actual situation.

4. Establishment and solution of optimization model

4.1 Establishment of optimization model

Different "parking areas" have different parking methods and different numbers of parking spaces. Therefore, different parking methods need to be analyzed to find a boarding point setting scheme that takes both efficiency and safety into consideration. There are three common parking point settings. Mode: Sequential departure mode, separate departure mode, and parallel multi-lane departure mode. Sequential departure and separate departure modes are targeted at single lanes. In the case of double-sided lanes, the left and right sides of the train do not affect each other, which is a simple promotion of single lanes. The parallel multi-lane departure mode adds a guide channel in the middle of the two lanes, allowing passengers on the left and right sides to flow and find the emptyest car, which is more in line with the actual situation. For different modes of departure, we use Monte Carlo simulation and enumeration search to find the optimal number of berths at the pick-up point, and finally recommend a departure method based on the actual throughput of the airport.

4.1.1 Departure mode description

1) Depart in sequence

The departures in turn include a parking area that can accommodate m cars and a pick-up area that can accommodate n cars. Passengers wait for cars at the head of the pick-up area, and the management staff puts people in accordance with the empty cars in the pick-up area, ensuring the safety and order of passengers and the pick-up area. In the sequential departure mode, the rear taxi can only pass after the front taxi leaves, that is, if the first taxi is not loaded with passengers, even if all the back is fully loaded, all cars cannot drive. Passengers choose a taxi according to the distance. Due to the high traffic flow of taxis and the large passenger flow, it may cause equality between the two parties.

2) Departing separately

The difference between a separate departure and a sequential departure is that a harbour-type parking zone and a deceleration zone are set up on the roadside. There is no need to wait for the vehicle in front, but it also increases the safety risks of vehicles and people.

3) Departures in parallel with multiple lanes

The parallel multi-lane departure mode is a combination of the above two modes, including a bay-style parking spot and a straight-line pick-up point. Vehicles are divided into two lanes at the pick-up area, one enters the bay-type pick-up area, and one straight ahead waiting for the platform, all cars follow the sequential departure mode: only after the front car leaves. The parallel multi-lane departure mode can select multiple branch roads. After consulting the literature, it is found in hcm that when the number of parallel lanes is greater than or equal to two. This method is better than the first two departure methods.

4.1.2 Optimized model of berths at boarding points

The parking pool at a large transportation hub usually has a large capacity, so the restriction m in the parking area can be ignored. According to the discussion above, the optimization model for the number of parking spaces at the boarding point is as follows:

$$\min \quad N$$

$$s.t. \quad \begin{cases} N_o N \le l_s - l_b \\ t_t = f(D_t, D_p, N) \\ t_p = g(D_t, D_p, N) \\ y t_t + (1-y) t_p \le thold \\ N_o \ge ns \end{cases}$$

N represents the number of parking spaces set, N_0 represents the length of a berth, ls, lb represents the length of the pick-up area and the length of the deceleration area. t_t and t_p are the taxi demand Dt, the passenger demand Dp and the parking space a function of the number N. Because the analytical form of the function is difficult to determine, we use Monte Carlo simulation to simulate t_t, t_p. Thold represents the minimum average time to meet the working conditions. Yt_t + (1-y) t_p becomes the weighted average waiting time, where y = 0.5.

4.2 Monte Carlo Solution

4.2.1 Variable requirements

We use the parameters of the Shanghai Hongqiao Airport used in the previous solution to solve, and have the following rules for the variables used in the solution:

(1) Passengers and taxi drivers arrive at the storage pool in a Poisson distribution.

(2) The queuing mode is fcfs.

(3) Each vehicle carries only one passenger.

(4) Do not consider the time when the vehicle and passengers move, and do not consider the time when the vehicle enters and leaves.

4.2.2 Solving Results

(1) Sequential departure mode

According to the characteristics of the sequential departure mode, the two lanes of the two lanes do not interfere with each other, so we only need to solve the single lane model and then design the parking spot parking space. Given a simulation time of 1000s, the Monte Carlo simulation has randomness. After the statistics of the optimal solution frequency, find the number of berths with the highest stable frequency. First observe the relationship between the number of berths generated by the simulation and the weighted average waiting time to find the thold. After accumulation,tho1d = 10.35, and the optimal solution is 3. After 500 cycles in this way, the optimal solution frequency is obtained. When N = 4, it has the highest frequency, but N = 3 and N = 5 are also high. The number of berths used depends on the airport manager. The degree of estimation of passenger flow and traffic flow. When N = 4, t_t is 43.7540s and t_p is 6.6102s. In the case of two lanes, remember that the left and right traffic flows are the same. There are 6 parking spaces, 3 on each side, lined up along the road near the passenger point.

(2) Individual departure mode

Following the solution of the previous section, the highest optimal solution frequency is 3, that is, the number of berth points for the optimal efficiency and optimal safety of a unilateral lane is 3, and the number of double lanes is 6. At this time, the average passenger waiting time is 25.5758s, the average waiting time for a taxi is 47.7381s.

(3) Multi-lane mode

The driver can find the front and side empty spaces to move, and the passengers can enter the lane. Therefore, in the parallel departure multi-lane mode, the lanes are connected and cannot be analyzed separately. The two-lane data is input into the simulation model and the optimal solution is 3. The total number of parking spaces is 3 * 4 = 12. The average waiting time for passengers is 13.3120s and the average waiting time for taxis is 21.8197 seconds. The actual design of Hongqiao Airport is also12 parallel multi-lane parking spaces.

It can be found that this mode is better than the above two modes, which can greatly alleviate traffic congestion and reduce passenger waiting time. However, the parallel departure multi-lane mode requires a large amount of investment for reconstruction. For airports with low throughput, the above two modes are used. This way of starting can greatly reduce waiting time and be more economical and environmentally friendly.

5. Conclusion

The taxi detention decision model established in this article has a high degree of agreement with the actual situation. The impact of changing the relevant factors in the model on the model solution results can be found in real life and has practical value. However, this article is for the driver leaving the airport to go to the city. The process of picking up passengers in the area has been greatly simplified, and a limited number of nodes have been delineated, which has caused the truth of the prediction results to decline to a certain extent. After processing the data, there are only 300 taxis that meet the requirements. In future research, we need to increase the amount of data to make the results more accurate.

References

[1] Mu Chen, Zhao Xiangmo. Multi-agent simulation model of empty taxi search [J]. Journal of

Natural Science of Xiangtan University, 2012, 34 (01): 113-116.

[2] Mu Chen, Xuan Huiyu, Zhang Fa. Simulation research on empty car behavior in taxi operation [J]. Journal of Systems Engineering, 2008, 23 (05): 554-562.

[3] Wu Jiaorong, Li Ming, Liang Lijuan. Analysis of the management mode and efficiency of taxi pick-up point in comprehensive passenger terminal [J]. Traffic Information and Safety, 2012, 30 (04): 18-23.

[4] Jin Miao, Wei Ningji. Research on Shanghai Taxi System Planning [J]. Urban Public Transport, 2007 (09): 32-34.

[5] Jiang Qiyuan. Mathematical Model (Fifth Edition) [M]. Beijing: Higher Education Press, 2018.