Research on Airport Taxi Driver Decision Based on RBF Neural Network

Xianxin Jin, Tianshu Gao, Xu Liu, Decheng Kong
Harbin University of Science and Technology, Weihai, Shandong, 274300

Abstract. Affected by the combination of different factors, taxi drivers leaving the airport will have different decision-making schemes. This paper establishes mathematical models to optimize the decision of airport taxi drivers and the optimization of passenger-carrying problems between taxis and passengers. First take Shuangliu Airport as an example, and take the number of taxis and the number of passengers in the data into the RBF neural network decision model. The results obtained are compared with the real data. Finally, the cyclic test shows the changes in taxi driver decisions is more dependent on the travel season, number of flights, and holidays. Then a taxi passenger model based on M/M/S is established. There are two arrangements for the two parallel lanes of the airport "ride zone": the final solution is concluded that when the length of the boarding area is 50 meters, setting three "boarding points" has the highest total boarding efficiency. By introducing the average mileage division index of taxi drivers, the short- and long-haul taxi drivers are finally obtained after the optimized scheme. The returns have basically reached equilibrium.

1. Introduction
One of the main means of transportation for passengers leaving the airport is taxis. Taxi drivers who take passengers to the airport have two different options when leaving the airport: they go to the "storage pool" to queue for passengers to return to the city, but the waiting time of the "pool" depends on the number of taxis and passengers waiting in line [1]. The driver can choose to return to the city directly without load. However, the taxi driver may pay additional no-load fees and potential passenger benefits [2]. There are many reasons for the decision-making of taxi drivers, including certain determining factors and some uncertain factors [3]. Passengers who want to take a car after getting off the aircraft must enter the "riding area" to queue up, and the airport taxi management staff is responsible for taxis according to certain principles, taxis are released in batches and entered the "riding area" [4].

2. Establishment and solution of ahp model

2.1. Model Selection
When the taxi sends the customer to the airport, the taxi driver will make two decisions: driving into the storage pool to wait in line for the passengers returning to the city; returning to the city with an empty car to carry passengers.

The changing law of the number of passengers and the taxi driver's income are directly related to whether the driver is carrying passengers, so the above problems can be expressed by the taxi's passenger
carrying situation [5]. This paper determines the impact that affects the decision of taxi drivers. Factors: travel season, time period, major city activities, weather conditions, number of motor vehicles around the airport, number of flights, holidays, etc. Because there are too many factors, in order to avoid the interference of the special factors on the model, it is necessary to screen the determined influencing factors. A hierarchical analysis model is used to observe the influence of various factors on taxi drivers.

2.2. Establishment of AHP Model

![Hierarchy of Taxi Driver Decisions](image)

Figure 1. Hierarchy of Taxi Driver Decisions

Compare the influencing factors one by one to construct a consistent matrix between the criterion layer and the target layer

$$A = \begin{bmatrix}
1 & 1 & 4 & 2 & \frac{1}{2} & 1 & 1 \\
1 & 1 & 3 & 1 & \frac{1}{2} & 2 & 2 \\
1 & \frac{1}{4} & 3 & 1 & \frac{1}{2} & 1 & 1 \\
\frac{4}{3} & \frac{1}{4} & 4 & \frac{1}{2} & 1 & 1 \\
\frac{1}{2} & 1 & 4 & 1 & \frac{1}{2} & 2 & 2 \\
2 & 2 & 4 & 2 & 1 & 1 & \frac{1}{2} \\
1 & 2 & 4 & 2 & 1 & 1 & 1 \\
1 & 2 & 4 & 2 & 1 & 1 & 1 
\end{bmatrix}$$

(1)

2.3. Model Solution

For a consistency matrix, any column vector is a feature vector, and the consistency of the positive and reciprocal matrix is better [6]. First, the consistency matrix is normalized by the column vector, and then the row vectors are summed and normalized. The eigenvectors are obtained by the quantization process. The eigenvalues are obtained based on the eigenvectors, and the matrix is more consistent, so the weights of all influencing factors are obtained.

$$w = \begin{bmatrix} 0.1585 & 0.1099 & 0.0427 & 0.1055 & 0.1923 & 0.1800 & 0.2110 \end{bmatrix}^T$$

(2)
2.4. Model Conclusions
Because the weight of the important factor of the city's major activities is too low and much less than
the weight of other factors, this unimportant factor is ignored, and the influence mechanism that
ultimately affects the taxi driver's decision is: tourism season, time, weather, airport internal machine
Number of moving cars, number of flights, holidays.

3. Establishment and solution of rbf neural network decision model

3.1. Model Establishment
For the change of taxi location and traffic around the airport, 500 vehicles are generated based on Monte
Carlo dynamic simulation under the main influencing factors of travel season, time, weather, number of
motor vehicles in the airport, number of flights, and holidays based on the analytic model. The decision
data of the car under specific conditions. The passenger results of the 500 taxis are used as training data.
Due to the large amount of data, the learning rate in the original BP neural network is a fixed value, so
the time required by the BP algorithm may be long and possibly falling into the localized optimal
solution [7]. The distance between the hidden nodes of the radial basis neural network and the input
pattern and the center vector is used as the independent variable of the function, and the radial basis
function is quickly used as an activation function, we choose the RBF-optimized BP neural network
model, which can achieve a fast convergence of weights in a shorter time and higher accuracy.

Loss function of least squares:

\[ \sigma = \frac{1}{P} \sum_{j} d_j - y_j c_j \]  

In the formula, x is an n-dimensional input vector, and c_j is the center value of the j-th radial basis
function, and it is consistent with the dimension of the input vector. \( \sigma \) is the normalized constant of the
ith center point of the basis function.
Training steps of radial basis neural network decision model:
1) Unsupervised self-learning training to solve the center and variance values of each basis function
of the hidden layer.
2) Supervised learning training, solving the changes in the weights between the hidden layer and the
output layer.

3.2. Model Solution
Using Matlab to run a neural network algorithm to obtain the results of 96 different combinations of
different influencing factors. We use the 500 data generated by Monte Carlo simulation as the initial
training data, and the initial weights are self-learned and related to the corresponding adjustment amount.
In addition, the cumulative calculation can get the new weight value, and so on until the error square of
the output layer meets the model's accuracy requirements.

3.3. Model Checking
Comparing the taxi driver decision results with 500 training data, it is concluded that the influence of
each influencing factor on the driver decision results is relatively consistent. The 500 training data error
convergence map and coefficient change map:
4. Analysis of influencing factors on driver’s decision dependence

The taxi flow distribution flowchart shows the specific steps to study the taxi traffic distribution of Shuangliu Airport. By collecting data, we get the taxi distribution data near Shuangliu Airport and its city, and determine the airport based on the geographical location of Shuangliu Airport. The number of taxis around and the distribution of taxi latitude and longitude are given. The number of passengers can be determined by the number of flights and the capacity of each flight.

4.1. Taxi Traffic Distribution at Shuangliu Airport

The statistics of the taxi traffic near Shuangliu Airport and the passenger carrying status can be obtained by consulting the statistical data. The reason for the taxi driver to decide whether to carry the passenger is not completely determined, and there may be subjective special circumstances. Passenger traffic statistics can be based on the number of flights, the number of passengers on each flight is determined. For the taxi traffic distribution in the data, the taxi density around the airport can be derived.

The geographical location of Shuangliu Airport is 30° 34'47" north latitude and 103° 57'02" east longitude. The total area of the airport is 1 million square kilometers. With the area of the airport as the radius, the number of taxis around the airport is calculated by matlab as 118.

4.2. Comparison of Model Decision Plan with Real Data

The specific time in the data can be used to find out whether it is currently the tourist season, the time of day, or whether it is a holiday. Therefore, we describe these three factors in detail in the results. The results and real results obtained by bringing the number of taxis into the above model are shown in Table 1:
Table 1. Comparison table of model decision plan and real data

| Dimension | longitude  | date       | time | Actual passenger | Model results |
|-----------|------------|------------|------|-----------------|---------------|
| 30.3404   | 103.5690   | 2018.08.02 | 10:15| Yes             | Yes           |
| 30.3499   | 103.5560   | 2018.08.18 | 14:12| Yes             | Yes           |
| 30.3390   | 103.5701   | 2018.09.05 | 08:34| Yes             | Yes           |
| 30.3402   | 103.5587   | 2018.09.29 | 21:30| no              | Yes           |
| 30.3421   | 103.5803   | 2018.10.12 | 12:30| Yes             | Yes           |
| 30.3409   | 103.5529   | 2018.10.30 | 16:21| no              | no            |
| 30.3522   | 103.5698   | 2018.11.10 | 10:08| Yes             | Yes           |
| 30.3447   | 103.5701   | 2018.12.01 | 12:21| no              | no            |
| 30.3426   | 103.5758   | 2018.12.22 | 20:40| no              | Yes           |
| 30.3501   | 103.5692   | 2019.01.06 | 08:10| Yes             | Yes           |
| 30.3432   | 103.5752   | 2019.02.28 | 13:20| no              | no            |
| 30.3532   | 103.5609   | 2019.03.12 | 17:20| no              | no            |
| 30.3433   | 103.5669   | 2019.04.20 | 15:14| no              | no            |
| 30.3400   | 103.5752   | 2019.05.06 | 10:34| Yes             | Yes           |
| 30.3534   | 103.5689   | 2019.06.29 | 19:26| no              | no            |

4.3. Analysis of The Rationality of The Model and Its Dependence on Relevant Factors

4.3.1. Model criteria. 1) The model takes into account the impact of time, holidays, travel season and other factors on the decision of taxi drivers.

2) Exclude influential factors in exceptional circumstances in the taxi driver's consideration criteria, such as: taxi drivers do not randomly pick up and drop off passengers at the airport, but accept appointments.

3) Considering the subjective thinking of taxi drivers, for example: taxi drivers may choose to work overtime at night to get additional income.

4.3.2. Model accuracy. According to the above results, it can be concluded that the accuracy of the model can reach 86.667%. Basically, it can be concluded that the taxi driver's decision for different decision factors under different influencing factors.

4.3.3. Dependence of related factors. The travel season, time period, weather, the number of motor vehicles in the airport, the number of flights, and the number of holidays that affect the decision of taxi drivers are weighted by 0.1585, 0.1099, 0.1055, 0.1093, 0.1800, and 0.2110. Two different results can be obtained by changing the time period. The same is true for other factors. When the travel season, number of flights, and holiday conditions change, most of the driver's decisions change; when the time period, weather, and the number of moving cars changes, the driver's decision may not change. Therefore, the tourist season, the number of flights, and the holidays are more dependent on the taxi driver's decision.

5. M/m/s-based taxi passenger carrying model

5.1. Model Construction Principles
The situation of taxis queuing passengers and passengers queuing in the airport at the airport is collectively referred to as the queuing phenomenon. This article obtains the average waiting time of all
passengers by changing the number of "boarding points" of the taxi to ensure the total boarding efficient.

Given the average passenger load factor of a taxi, the service intensity between all taxis and passengers in the area is calculated, that is, the average service time in the unit of taxi service. The airport's "ride zone" "for two parallel lanes, compare the setting of" boarding points "on both sides of the parallel lanes and the" boarding points "on one side. It is found that when the" boarding points "are set on one side, taxis carry passengers more efficiently.

1) One side lane is set up with "boarding point", and the other lane is used as the lane for taxis. The taxis will not queue up when entering the "storage pool".
2) Taxi is less likely to wait due to taxi stops in front of passengers when leaving.

5.2. Model Description

Before the model is established: the corresponding service intensity $\rho^*$ must be calculate to determine if it meets prerequisites for $\rho^* < 1$. The service intensity of the taxi is:

$$\rho^* = \frac{\lambda}{s\mu}$$  \hspace{1cm} (4)

Passenger arrival rates $\lambda$ can be derived from airport statistics and the average passenger load factor of each taxi, the number of "boarding points" $s$ is continuously modified to find the minimum average waiting time. When $s = 1$, $\rho^* = 0.667 < 1$; when $s = 2$, $\rho^* = 0.333 < 1$; when $s = 3$, $\rho^* = 0.222 < 1$, etc. Although the value of the service intensity obtained by the solution is less than 1, it is not much different from 1, so the queue phenomenon will still occur in the established M/M/S model, but this phenomenon is not obvious and the arrival of taxis and passengers is subject to three conditions:

1) Smoothness. In a certain time interval, the number of passengers arriving is only related to the length of the boarding area and has nothing to do with the position where the passengers get off the plane.
2) Persistence. Within a certain time interval, the number of passengers arriving is related to the queue leader, and has nothing to do with whether the passenger chooses to ride.
3) In general, the probability of reaching multiple passengers at the same time is very small, so it can be ignored.

5.3. Model Establishment

In the taxi passenger model of M/M/S, the adaptive formula is listed according to the meaning of each model parameter. If a queuing phenomenon occurs, it is solved according to its average queue.

Pick-up point idle probability:

$$P_0 = \left[ \sum_{k=0}^{s-1} \frac{1}{k!} \left( \frac{\lambda}{\mu} \right)^k + \frac{1}{s!} \frac{1}{1-\rho^*} \left( \frac{\lambda}{\mu} \right)^s \right]$$  \hspace{1cm} (5)

Where $p_0$ is the idle probability at the pick-up point.

Average length of stay of passengers:

$$W_s = \frac{L}{\lambda} = W_q + \frac{1}{\mu}$$  \hspace{1cm} (6)

Where $W_s$ is the average stay time of passengers, and $W_q$ is the average wait time of passengers.
5.4. Model Solution
First will be the formula:

\[ \rho^* = \frac{\lambda}{\varepsilon \mu} \]  

(7)

Bring derived \( \rho^* \) into the expression \( p_0 \) to get the probability of the entire "storage pool" being free. Then further get the number of passengers waiting to enter the taxi, and bring the result into the expression to get the average waiting time for passengers. From the waiting time and the average team leader, the total waiting time for all passengers to get into the taxi can be obtained. Take matlab to solve and get:

![Figure 3. Time chart of passenger arrival and departure when "boarding point" is 1 (left) and time chart of passenger staying when "boarding point" is 1 (right)](image)

Figure 3 shows the time from when a passenger enters the "storage pool" to they leave by taxi and when they stay when there is only one "boarding point". It can be concluded that the time difference between the arrival and departure of passengers is too large.

![Figure 4. Passenger arrival and departure time chart when the "boarding point" is 2 (left) and passenger stay time chart when the "boarding point" is 2 (right)](image)
It can be seen from Figure 4 that when the "boarding point" increases, the difference between the arrival and departure times of passengers decreases to a certain extent, and the average stay time also decreases.

![Figure 4](image)

Figure 4. Time chart of passenger arrival and departure when "boarding point" is 3 (left) and time chart of passenger stay when "boarding point" is 3 (right)

According to Figure 5, it can be concluded that the difference between the arrival time and departure time of the passenger at this time has reached a minimum value. At this time, the passenger's stay time is the shortest.

According to the results of the model, it can be concluded that the number of "boarding points" is not as many as possible, but there is a limit. In a certain long "storage pool", the optimal boarding point tends to a specific value. According to the actual situation of the airport, it was concluded that three "boarding points" were set up at an average distance in the "storage pool" with a length of 50 meters, so that the total passenger efficiency was the highest.

6. Income equalization model

The premise for the establishment of this model is to find the mileage of each passenger in the airport and the total mileage of passengers on that day. Based on the short-haul and long-haul driver's mileage and revenue data of the taxi drivers at Shuangliu Airport as a sample. Increase the revenue of short-haul taxi drivers to balance the revenue among taxi drivers.

6.1. Model Establishment

use $N_x (x = 1, 2, 3)$ represents the total mileage of the xth taxi driver in a day, $x$ represents the number of samples, $n_{xm} (m = 0, 1, 2, 3)$ represents the mileage driven by the m-th passenger taxi driver in a day. Obviously:

$$N_x = \sum_{0}^{m} n_{xm}$$

Average the total mileage driven by all taxi drivers in a sample of group $x$ in one day, that is,
is used as a quantitative indicator to distinguish long-distance and short-distance taxis. A sample of $N_{xa} = 198$ is taken as the passenger load situation of a long-distance taxi in a day. After distinguishing short-distance taxis from long-distance taxis, the data of the short-haul taxis is reasonable. The average $n_{ma}$ is used to represent the average mileage of each long-haul taxi or short-haul taxi. After sorting the long-haul taxis and short-haul taxis, we have: The data of $n_{ma}$ represents the length of the mileage from long-haul or short-haul taxis to the $m$-th passenger in the last day. Based on these $m$ data, the $m$-th taxi driver's income can be obtained. Perform $m$ accumulations to get the driver's benefit $X_m$ at the accumulated time, that is:

$$X_m = x_m + X_{m-1}$$

(10)

With $X_m$ as the ordinate and time $t$ as the abscissa, the Matlab program can be used to obtain a comparison chart of short-distance drivers and long-distance drivers. It can be concluded that there is a large income gap between long-distance and short-distance taxis.

6.2. Model Solution

In order to reduce this gap, the short-haul taxi group is taken as the research object, and the average passenger mileage per short-haul taxi is used as an indicator $\eta^*$. Compare the mileage of each taxi with $n_{xm}$, and compare $n_{xm} < \eta^* = 18.1$. The number of data to be extracted, this is the number of times that all samples need to get into the lane preferentially for $m$ times of solicitation. It is recorded as $Y_c$ ($c$ is a positive integer). Due to the short-distance priority entry scheme, when short-distance taxis enter after the lane, the distance that the next passenger encountered will follow the (0-1) distribution. It may be useful to express the short-distance taxi income this time as the average of short-distance revenue $p_s$ and the average of long-distance revenue behavior $p_l$. It expresses the actual income $Q$ of all samples of the original data after each passenger chooses a short-haul bus, after implementing the subsidy of the priority entry lane scheme, expressed as

$$Q_m = \frac{2}{X} \cdot Y_c \cdot P_a + P_s \text{ or } Q_m = \frac{2}{X} \cdot Y_c \cdot P_a + P_l$$

(11)

Finally, using the newly generated $Q_m$ as the ordinate and time as the abscissa, the balance of the short-distance taxi driver and the long-distance taxi driver's benefits are balanced, as shown in Figure 6.
As can be seen from Figure 6, the scheme basically balances the long- and short-distance benefits, and the scheme is indeed effective.

By distinguishing the indicators $N_{xa}$, two groups of long and short-haul taxis can be obtained, and the short-haul taxi group can determine whether the car should enter the priority taxiway. When the taxi mileage $n_{xm} < \eta^*$, the vehicle can enter the priority lane, and the revenue of the taxi has basically reached equilibrium after the optimization scheme.

7. Conclusion

This article analyzes the resource allocation between airport taxis and passengers, and uses the queuing theory to simplify the problem of matching passengers and taxis to queues, and provides a more optimal queuing scheme for multiple lanes in parallel. Using Monte Carlo to simulate the data. The specific data results are obtained, which make up for the lack of data and make the model more persuasive. The model makes a quantitative analysis of the taxi driver's income balance and gives a standard for balancing taxi income. This standard novel and reasonable, with stronger innovation and practical significance. We can extend it to the allocation of taxis and passengers at any location in the entire urban area and improve it, which has strong practical significance. The model can be combined with resource scheduling problems that are generalized to factories and workshops to give an optimal solution under limited conditions.

References

[1] Xu Xin, Li Xiao, Liu Xuedong. Application of Internet + in taxi resource optimization model[J]. Journal of Zaozhuang University, 2017, 34 (02): 95-109.
[2] Zhou Junwu, Sun Chuanyao, Wang Fuli. Research and implementation of radial basis function (RBF) network[J]. Mining and Metallurgy, 2001 (04): 71-75.
[3] Tang Luliang, Zheng Wenbin, Wang Zhiqiang, Xu Hong, Hong Jun, Dong Kun. Spatio-temporal distribution detection method of gps trajectories of urban taxis boarding and disembarking passengers[J]. Journal of Geo-Information Science, 2015, 17 (10): 1179-1186.
[4] Yang Xiaoyan. Application of queuing theory in traffic control system[D]. Qingdao University of Science and Technology, 2015.
[5] Peng Yingchun, Dong Sibin, Chang Wenhu. Using queueing model to measure the efficiency of hospital outpatient process[J]. Chinese Journal of Hospital Management, 2005 (12): 806-809.
[6] Zhang Achai. Research on simulation optimization of queuing system performance based on
monte carlo method [D]. Hangzhou Dianzi University, 2011.

[7] Huang Xia, Sun Jin. Research on the calculating man-hour shift scheduling based on cost-benefit balance model——taking m shipping company as an example[J]. Science and Technology Innovation Herald, 2012 (09): 239-240.