

# Research on metro passenger flow route selection model based on passengers' familiarity with road network

Yiting Liu<sup>1</sup>, and Shunping Jia\*

<sup>1</sup>School of traffic and transportation, Beijing Jiaotong University, Beijing

**Abstract:** Road network familiarity is a key attribute that affects passengers' travel route choice. This paper constructs a differentiated travel generalized cost function based on the passenger's road network familiarity and the influencing factors of route choice, and uses the Regret Theory to construct a route choice model. By setting passenger decision-making rule weights increase the flexibility of the model. The paper uses the method of combining RP survey and SP survey to conduct route selection behavior survey and calibrate model parameters. Finally, the prediction results before and after the passenger classification are compared with the survey data. The prediction error value is 5.98%, and the prediction accuracy after passenger classification is improved by 6.03%. The effectiveness of the prediction model is verified and the necessity of passenger classification is verified.

## 1 Preface

The research on traveler's route choice behavior has important theoretical significance and practical value for urban traffic planning and management. Among them, the generalized travel costs and passenger decision-making behavior will affect the traveler's route choice behavior.

In terms of the construction of the generalized cost function, domestic scholars have constructed a more realistic generalized cost function that considers various influencing factors through qualitative and quantitative analysis of passenger travel route selection behavior. The previous studies on generalized travel costs mainly focused on path attributes. For example, Su Juan<sup>[1]</sup> considered the transfer congestion delay when analyzing the passenger transfer time cost, and expressed the transfer congestion delay time by using the product of the departure interval and the additional congestion coefficient. Liu Jianfeng<sup>[2]</sup> introduced a penalty for the number of transfers in the generalized cost function. Xu Da<sup>[3]</sup> proposed the cost impedance of the network topology angle and the line dependent impedance. Chen Peiwen<sup>[4]</sup> proposed the concept of detour cost and described the deviation between the schematic diagram of the subway line and the actual subway network.

However, the generalized travel cost is also closely related to the attributes of travelers. Liu Jianfeng<sup>[5]</sup> et al. divided travelers into eight categories based on the cross-classification of passenger attributes. For all types of passengers, a generalized travel cost function was constructed based on travel time, transfer time, and number of transfers, but ignored The path attribute of congestion degree. Qin Zhipeng<sup>[6]</sup> classifies passengers

based on their familiarity with rail transit lines and proposes a path cost calculation method based on time factors and route map factors, but does not consider the impact of congestion and passenger transfer psychological perception amplification on path impedance judgment.

Traditional passenger route choice behavior research usually uses "expected utility theory, EUT" for analysis. With the deepening of research, some scholars began to consider the bounded rationality of passenger route choice, and proposed theories such as "Prospect Theory<sup>[7]</sup>", "Regret Theory", and "Reflection effect". The application of regret theory in route choice behavior research is still in its infancy.

Existing studies mainly apply regret theory to the study and analysis of passengers' travel behavior and emergency evacuation methods, but rarely apply to the route selection of rail transit passengers. Moreover, few studies consider the proportion of regret theory's influence on travel behavior and express this proportion through parameters.

In summary, existing studies mostly establish generalized travel cost models from the perspective of route attributes, and have less consideration of passenger attributes such as the familiarity of the passenger's road network. In view of this, based on the perspective of passengers, this paper classifies passengers based on their familiarity with the road network, and establishes differentiated generalized travel cost functions for different types of passengers. At the same time, the regret theory is introduced to establish a generalized random regret minimum path choice model, which can more flexibly characterize passenger travel behavior by adjusting the weight of passenger decision-making rules. Through actual survey data, model parameters are

<sup>1</sup>19120859@bjtu.edu.cn

\*Corresponding author: Tel: 010-51684109, E-mail: shpjia@bjtu.edu.cn

calibrated to verify the effectiveness of the model and reveal the route selection of different types of passengers behavior and has a certain theoretical and practical significance.

## 2 Passenger classification and construction of generalized travel cost function

### 2.1 Passenger classification based on road network familiarity

According to the data analysis of rail transit passenger travel behavior survey, it can be seen that the difference of passengers' familiarity with the road network leads to the difference in the basis of their route selection. Passengers who are familiar with the road network have rich travel experience and can choose a reasonable travel route based on experience and the actual operation of the road network. For passengers who are not familiar with the road network, such as migrant workers, tourists, etc., they need to inquiry subway line diagrams and map APP software to select travel routes.

Based on this, the passengers on the rail transit network are classified according to their familiarity with the network:

- (1) Type a: Passengers who are familiar with the road network and rely on their own experience to travel;
- (2) Type b: Passengers who are not familiar with the road network and need to use the map APP to complete the trip;
- (3) Type c: Passengers who are not familiar with the road network and need to use the subway line diagram or inquiries to complete the trip.

### 2.2 Construction of generalized travel cost function

#### 2.2.1 Generalized travel costs for passengers of category a and b

Passengers of two types a and b can obtain richer travel route information through their own experience or map app. When choosing a route, it mainly considers the travel time, the number of transfers, the transfer time and the degree of congestion.

(1) Travel time cost

$$T_{od}^{n,k} = \sum t_r^k + \sum t_s^k \quad (1)$$

Among them:  $\sum t_r^k$  represents the running time of the train on the  $k$ -th route, and  $\sum t_s^k$  represents the stopping time of the train on the  $k$ -th route.

(2) Transfer time cost

$$E_{od}^{n,k} = \sum_i E_i^k = \sum_i \alpha^n \cdot (n_{od}^{ik})^{\beta^n} \cdot \left( t_i^k + \frac{H_i^k}{2} \right) \quad (2)$$

Among them,  $t_i^k$  represents the transfer travel time at

the  $i$ -th transfer station,  $H_i^k$  represents the train departure interval,  $\alpha^n$ ,  $\beta^n$  represent the perception coefficients of the  $n$ -th passenger transfer time and the number of transfers respectively, and  $n_{od}^{ik}$  represents the cumulative number of transfers at the transfer station  $i$ .

(3) Congestion penalty cost

The congestion cost coefficient is:

$$w_{od}^{i,k} = \begin{cases} 0 & E_i \cdot C_p \leq Z_l \\ A \left( \frac{E_i \cdot C_p - Z_l}{Z_l} \right) & Z_l < E_i \cdot C_p < C_p \\ A \left( \frac{E_i \cdot C_p - Z_l}{Z_l} \right) + B \left( \frac{E_i \cdot C_p - C_p}{C_p} \right) & C_p \leq E_i \cdot C_p \end{cases} \quad (3)$$

Among them,  $w_{od}^{i,k}$  represents the congestion coefficient of the path  $k$  between  $od$  in section  $i$ ,  $E_i$  is the average full load rate of the section,  $Z_l$  is the number of subway seats on the line  $l$ ,  $C_p$  is the approved number of passengers on the subway of line  $l$ , and the congestion penalty coefficients of A and B are based on relevant literature,  $A=0.5$ ,  $B=0.8$  [8].

The congestion penalty cost for path  $k$  is:

$$\psi_{od}^k = \sum_i^n (T_{od}^i)^k \cdot w_{od}^{i,k} \quad (4)$$

The generalized travel cost of passengers in categories a and b is:

$$\begin{aligned} C_{n,od}^k &= T_{od}^{n,k} + E_{od}^{n,k} + \psi_{od}^{n,k} \sigma_n \\ &= \lambda_n \cdot \left( \sum t_r^k + \sum t_s^k \right) + \sum_i \alpha^n \cdot (n_{od}^{ik})^{\beta^n} \cdot \left( t_i^k + \frac{H_i^k}{2} \right) \\ &\quad + \left[ \sum_i^n (T_{od}^i)^k \cdot w_{od}^{i,k} \right] \cdot \sigma_n \end{aligned} \quad (5)$$

Among them:  $\lambda_n$  and  $\sigma_n$  are the perception coefficients of the  $n$ -th passenger on the travel time and the degree of congestion.

#### 2.2.2 Generalized travel costs for passengers of category c

For category c passengers, the factors that affect their judgment of the travel route are the travel information that can be provided on the subway line diagram, including the number of stations passing by the route and the number of transfer stations. Passengers judge the travel time based on the number of stations, and judge the transfer time based on the number of transfer stations. The generalized travel cost of type c passengers from station o to station d is:

$$C_{n,od}^k = (\phi(M_{od}^k - 1)) \cdot \lambda_n + \alpha_n \cdot \mu \sum_i (n_{od}^{i,k})^{\beta_n} + \psi_{od}^{i,k} \sigma_n \quad (6)$$

Among them:  $\phi$  is the conversion factor for the

number of stations,  $\mu$  is the transfer conversion factor,  $M_{od}^k$  is the number of stations on the route. For the  $n$ -th passenger,  $\alpha_n$ ,  $\beta_n$ ,  $\lambda_n$ ,  $\sigma_n$  are the perception coefficient of transfer time, number of transfers, travel time, congestion degree.

### 3 Generalized random regret minimum route selection model

Before a passenger's journey: According to the part of the road network information that the passenger has, as well as the travel experience between the OD and their own choice preferences, they choose the route that they think has the least regret value. When the trip is over, the passenger will evaluate the route chosen for this trip and compare it with the unselected route, thereby generating feedback of regret or joy. This feedback will have a continuous correction effect on the subsequent route choice. In the regret minimization model, the utility of the selected path is not only related to the attributes of the path itself, but the utility of the unselected path will also affect the selected path. Therefore, the regret model can better explain route choice behavior.

Suppose there are  $n$  options in the path option set  $W$ , and each option has  $x$  attributes  $m$  to describe, and these attributes are comparable among all alternatives.

Assuming that there are  $v_m(i)$  and  $v_m(j)$  respectively the utility values of pat  $i$  and  $j$  with respect to attribute  $m$ , the passenger's regret perception of path  $i$  on attribute  $m$  is expressed by the utility difference, and the expression is as follows:

$$R_{ij}^m = (v_m(i) - v_m(j)) / v_m(i) \quad (7)$$

The introduction of parameter  $\theta_m$  indicates the degree of relevance of the attribute to the regret value of the path plan. The greater the absolute value of the parameter, the greater the influence of the attribute on route selection. When attribute  $m$  is travel time, transfer time, transfer times, and congestion degree, parameter  $\theta_m$  takes  $\lambda_n$ ,  $\alpha_n$ ,  $\beta_n$ ,  $\sigma_n$  respectively.

In reality, passenger travel behavior is heterogeneous, and the degree of bounded rationality of passenger path choice behavior is different, and the perception coefficient and regret weight of each path attribute  $m$  are also heterogeneous. Introduce the regret weight parameter  $\delta_m^n$ , which represents the regret weight of the  $n$ th passenger on the attribute  $m$ . Relative value of attributes between every two path schemes  $R_{ij}$ :

$$R_{ij} = \sum_{i \neq j} \sum_M \ln(\delta_m^n + \exp(R_{ij}^m \cdot \theta_m)) \quad i \neq j \quad (8)$$

When  $\delta_m^n=1$ , formula (8) is a random minimization selection model. When  $\delta_m^n=0$ , it shows a linear relationship and is the expected utility maximization selection model. The change of the regret weight coefficient can reflect the route choice behavior of different types of passengers under the multi-decision rules. When the selection scheme is extended to  $n$  paths, the introduction of  $\varepsilon_i$  represents the unobservable heterogeneity of the alternative paths in the regret model. The random regret value of path  $i$ :

$$R_i = \sum_{i \neq j} R_{ij} + \varepsilon_i \quad i, j \in W \quad (9)$$

Considering the volatility of factors influencing the route selection, establish a discrete route selection model. In this case, the logit model is used to obtain the selection probability for calculation. The probability of choosing path  $i$  is:

$$P_R(i) = \text{prob}(R_i < R_j, \forall i, j \in W \text{ 且 } i \neq j) \\ = \frac{\exp\left(-\left(\sum_{i \neq j} \sum_M \ln(\delta_m^n + \exp(R_{ij}^m \cdot \theta_m))\right) + \varepsilon_i\right)}{\sum_{j \in W} \exp\left(-\left(\sum_{i \neq j} \sum_M \ln(\delta_m^n + \exp(R_{ij}^m \cdot \theta_m))\right) + \varepsilon_i\right)} \quad (10)$$

## 4 Case study

### 4.1 Data Sources

Assuming that the interviewee is in the morning rush hour in Beijing on a certain day, the RP survey and the SP survey are combined to investigate the basic personal information of passengers, rail transit travel information, and travel willingness information in a simulated scenario. In the route selection attention, three OD of Beijing rail transit simulation scenarios are designed, namely "Jiaomen West-Yong'anli", "Muxidi-Ciqikou", and "Dazhong Temple-Lishuiqiao South". There are two effective paths between each OD, and the interviewee chooses the path. A total of 606 questionnaires were collected in the survey, of which 525 were valid questionnaires, and the questionnaire effective rate was 86.63%.

### 4.2 Parameter calibration

The train marshalling situation of the research line in this paper is 6B marshalling.  $Z_l=240$ ,  $C_p=1428$ . In class c passenger travel expenses, the average running time of Beijing rail transit trains between the two stations is used

**Table 1** Parameter Values in Generalized Random Regret Minimum route selection Model

Passenger type	$\lambda_n$	$\alpha_n$	$\beta_n$	$\sigma_n$	$\delta_1^n$	$\delta_2^n$	$\delta_3^n$	$\delta_4^n$
a	0.76	0.63	0.24	0.68	0.68	0.52	0.44	0.30
b	0.65	0.71	0.35	0.52	0.71	0.45	0.46	0.43
c	0.67	0.86	0.42	0.18	0.49	0.52	0.56	0.44

**Table 2** Information of the travel route of "Dazhong Temple-Lishui Bridge South"

OD	Dazhong Temple-South of Lishui Bridge				
Path	Path 1		Path 2		
Line	Line 13-5	Line 5	Line 13-10	Line 10-5	Line 5
Number of stations (pieces)	9	1	1	6	4
Travel time/ (min)	35	4	3	18	9
Peak departure interval/ (min)	2.5	2	2.5	2	2
Transfer time/ (min)	2	-	4	1	-
Average full load rate	0.86	0.29	0.48	0.51	0.63

as the number of stations conversion coefficient  $\phi$  (2.20min), and the average transfer time of Beijing rail transit interchange stations is used as the transfer conversion coefficient  $\mu$  (3.60min).

Use multiple linear regression analysis to estimate  $\lambda_n$ 、 $\alpha_n$ 、 $\beta_n$ 、 $\sigma_n$ 、 $\delta_1^n$ 、 $\delta_2^n$ 、 $\delta_3^n$ 、 $\delta_4^n$  are obtained by fitting the statistical results of the questionnaire survey data. In summary, the calibration results of each parameter are shown in Table 1.

### 4.3 Model validity analysis

The OD pair "Dazhong Temple-Lishui Bridge South" is selected as the research object. There are two effective paths between the OD pair, as shown in the table 2.

The questionnaire survey method is used to test the generalized random regret minimization route selection model built in this paper. Predict the route choices of different types of passengers, and compare and analyze them with the results of the survey.

**Table 3** Comparison of predicted value and true value of route selection model

Passenger type	Path 1		Path 2		error
	Predictive value	Actual value	Predictive value	Actual value	
a	61.19%	66.42%	38.81%	33.58%	5.23%
b	78.63%	70.78%	21.37%	29.22%	7.85%
c	70.77%	75.63%	29.23%	24.37%	4.86%

From the data in the table 3, it can be concluded that the calculation result of the effective route selection probability between "Dazhong Temple-Lishui Bridge South" is closer to the questionnaire survey result, and the average prediction error is 5.98%, which is relatively small. Therefore, the effectiveness of the generalized random regret minimization model can be verified.

### 4.4 Comparison of selection results before and after passenger classification

When passengers are not classified, the perception coefficients of various types of passengers on the factors affecting route selection are not considered. Combined with the survey data, the perception coefficients of travel time, transfer time, transfer times, and degree of congestion are 0.7, 0.5, 0.6, and 0.5, respectively, and the corresponding regret weights are 0.5, 0.7, 0.6, and 0.8. Based on generalized random regret Minimize the route selection model, calculate the route selection probability of unclassified passengers, and compare them with those after classification. The results are shown in Table 4:

**Table 4** Comparison of predicted value of route choice probability before and after passenger classification

OD	Path	Before classification	After classification	Actual value
Dazhong Temple-South of Lishui Bridge	Path 1	73.64%	67.61%	65.87%
	Path 2	26.36%	32.38%	34.13%

It can be seen from the table that the predicted value of the selection probability after the passenger classification is closer to the actual selection probability than before the classification, and the prediction accuracy rate has increased by 6.03%. It shows that passengers with different levels of familiarity with the road network have significant differences in route selection. The classification of passengers according to their familiarity with the road network can more accurately predict the route choice behavior of rail transit passengers, and provide model support for subsequent passenger flow distribution research.

## 5 Conclusion

This paper constructs a generalized random regret minimum route selection model based on passengers' familiarity with the road network. Combined with the survey data, it can be known that the passengers' familiarity with the road network has a significant correlation with their travel route selection. Passengers with different levels of familiarity with the road network have different perceptions of the four factors of travel time, transfer time, transfer times, and degree of congestion, and have obvious heterogeneity of preferences. As the degree of familiarity with the road network increases, the perception coefficient of travel time and congestion increases, and the perception coefficient of transfer time and number of transfers decreases. Therefore, passengers are divided into three categories, and a differentiated generalized travel cost function is constructed. Taking into account the bounded rationality of passenger route selection, the regret theory is introduced to establish a travel route prediction model. The prediction error is 5.98%, the accuracy rate is high, and the model is effective. At the same time, the model is used to compare the prediction results before and after passenger classification. The prediction accuracy after classification is 6.03% higher than before classification. The classification of passengers according to the familiarity of the road network can more accurately describe the travel route selection behavior of rail transit passengers and improve the accuracy of route selection prediction, making the results more in line with reality.

Using the results of passenger classification and the model built in this paper to further analyze the distribution of rail transit passenger flow is the next research direction.

## References

1. Su Juan. Research on Passenger Flow Distribution of Urban Rail Transit[D]. Beijing Jiaotong University, 2009.
2. Liu Jianfeng. Modeling and Empirical Research on Urban Rail Transit Network Flow Distribution Based on Transfer [D]. Beijing Jiaotong University, 2012.
3. Xu Da. Research on Metro Passenger Flow Distribution Model Based on Logit [D]. Southwest Jiaotong University, 2016.
4. Chen Peiwen. Research on Urban Rail Transit Passenger Flow Distribution Method Based on Improved Logit Model[D]. Beijing Jiaotong University, 2018.
5. Liu Jianfeng, Si Bingfeng, Liu Xinhua, Ma Yilin. Research on Urban Rail Traffic Flow Distribution Model Based on Passenger Type[J]. Logistics Technology, 2011, 30(23): 101-103+107.
6. Qin Zhipeng. Research on the travel route choice of urban rail transit passengers under the network condition [D]. Beijing Jiaotong University, 2011.
7. Shi Guoqi, Jia Fuqiang. Research on dynamic route selection model based on prospect theory[J]. Shandong Science, 2016, 29(05): 111-116+123.
8. Gao Yufang. Research on dynamic allocation model of urban rail transit based on regret theory[D]. Beijing Jiaotong University, 2018.