International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:14, No:9, 2020

Analysis of Residents' Travel Characteristics and Policy Improving Strategies

Zhenzhen Xu, Chunfu Shao, Shengyou Wang, Chunjiao Dong

***Abstract*—**To improve the satisfaction of residents' travel, this paper analyzes the characteristics and influencing factors of urban residents' travel behavior. First, a Multinominal Logit Model (MNL) model is built to analyze the characteristics of residents' travel behavior, reveal the influence of individual attributes, family attributes and travel characteristics on the choice of travel mode, and identify the significant factors. Then put forward suggestions for policy improvement. Finally, Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) models are introduced to evaluate the policy effect. This paper selects Futian Street in Futian District, Shenzhen City for investigation and research. The results show that gender, age, education, income, number of cars owned, travel purpose, departure time, journey time, travel distance and times all have a significant influence on residents' choice of travel mode. Based on the above results, two policy improvement suggestions are put forward from reducing public transportation and non-motor vehicle travel time, and the policy effect is evaluated. Before the evaluation, the prediction effect of MNL, SVM and MLP models was evaluated. After parameter optimization, it was found that the prediction accuracy of the three models was 72.80%, 71.42%, and 76.42%, respectively. The MLP model with the highest prediction accuracy was selected to evaluate the effect of policy improvement. The results showed that after the implementation of the policy, the proportion of public transportation in plan 1 and plan 2 increased by 14.04% and 9.86%, respectively, while the proportion of private cars decreased by 3.47% and 2.54%, respectively. The proportion of car trips decreased obviously, while the proportion of public transport trips increased. It can be considered that the measures have a positive effect on promoting green trips and improving the satisfaction of urban residents, and can provide a reference for relevant departments to formulate transportation policies.

***Keywords*—**Travel characteristics analysis, transportation choice, travel sharing rate, neural network model, traffic resource allocation.

I. Introduction

WITH the rapid development of the economy and society, the number of motor vehicles increases year by year, and the proportion of residents relying on cars to travel increases accordingly, which also leads to more and more serious congestion on urban roads. Residents choose to travel by car, largely because of insufficient public transportation resources, unable to meet a lot of travel demand. By analyzing the characteristics of residents' trips and digging out the significant factors that affect residents' trips, we can accurately grasp the rules of residents' trips, predict the travel demand, optimize the travel structure, and provide some references for formulating traffic policies and rationally allocating resources.

Domestic and foreign scholars have done a lot of research on residents' travel behavior, and the non-lumped model is the most widely used, mainly including multiple Logit model, nested Logit model, paired combination Logit model, mixed Logit model and so on [1]; in addition, improved models of NL model and Probit model are also adopted. Ben-Elia et al. [2], based on considering the panel effect, established a hybrid Logit specification discrete selection model to study the influence of information and experience on drivers' path selection behavior. Liya et al. [3] constructed a two-layer NL model, in which the characteristics of travelers, travel characteristics and service level of travel mode were selected as the utility variables, and they respectively took departure time and travel modes as the lower layer of the model to analyze residents' travel choice behavior. Van Can [4] established several Probit models to study the travel choice behavior of tourists in Nha Trang, Vietnam. Long et al. [5] analyzed the effect mechanism of family attributes and personal attributes on the subjective attitude of low-income commuters and adopted the MNL model with latent variables to study the effects of various attributes on their choice of travel mode. Dapeng et al. [6] established a two-layer Logit model based on travel behavior, studied the influencing factors of inter-city passenger traffic structure, and predicted the future with example data. Wei et al. [7] took dynamic factors as independent variables and built an MNL model to analyze the impact of these factors on residents' car rental choice behavior. Xiaomei et al. [8] constructed and estimated the layered Logit model to study the behavior characteristics of the joint selection of inter-city transportation mode and departure date under the policy of free highway during holidays, and simulated different schemes, and proposed the suggestion of formulating time-sharing preferential policies. Acheampong [9] used the Logistic regression model to analyze the influence of the type of land used by commuters in African cities on their attendance patterns. There are also some scholars adopt mixture model or other methods to study the residents travel behavior, Bing [10] combined with the survey data, such as fusion to extend the MIMIC model and planned behavior theory as the core to the discrete choice model of multinomial Logit model as the core to

Zhenzhen Xu is with Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China (phone: 13161092525; e-mail: 18120927@bjtu.edu.cn).

Chunfu Shao (Corresponding author), Ph.D., Professor, is with Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, School of Traffic and Transportation, Beijing Jiaotong University, 3 Shangyuan Cun, Haidian District, Beijing, China, 100044, (phone: 86-010-51688236, Fax: 86-010-51688236, e-mail:: cfshao@bjtu.edu.cn).

Shengyou Wang, Ph.D., Candidate, and Chunjiao Dong, Ph.D., Professor, are with Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, School of Traffic and Transportation, Beijing Jiaotong University, 3 Shangyuan Cun, Haidian District, Beijing, China, 100044.

build a hybrid selection model, analysis of people's social demographic characteristics and psychological impact variables influence on mode choice behavior; Yufeng et al. [11] comprehensively considered external and internal factors and adopted Structural Equation Model (SEM) to explore the relationship structure between various factors and the choice of urban residents' travel modes.

Most of these studies are limited to traditional models, which are only of high reference value when analyzing the current situation, and have the problems of poor prediction accuracy and practicability. Therefore, they cannot be used to evaluate the expected effect after the implementation of optimized policies. As machine learning is gradually applied to urban traffic problems such as traffic flow prediction and travel behavior analysis, scholars at home and abroad have found that they have a better ability to classify data through research. Xiugang et al. [12] applied the SVM model to vehicle collision prediction and proved that the model was faster and more effective than the traditional binomial prediction model. Wang et al. [13] compared the confidence band estimators of the SVM model and the Logistic regression model, and found that the classification accuracy of SVM was higher than that of the Logistic regression model. Shengyou et al. [14] constructed an NL model and a SVM model based on the residents' travel data within the scope, analyzed and predicted the residents' travel mode selection behavior before and after the implementation of the optimization and improvement measures, and verified the effectiveness of the model and the optimization effect of the improvement measures. Jialin et al. [15] applied LSTM to urban road traffic speed prediction. Yu et al. [16] used LSTM to predict short-term traffic flow on urban roads, which reduced the error of other methods. These studies prove that the prediction results of the neural network model are more accurate than the traditional regression model. Therefore, the MNL model is used to analyze the basic characteristics of residents' travel, and then the SVM model and MLP deep learning model are introduced to predict and evaluate the effect after the implementation of the policy, and the results are more convincing.

II. Resident Travel Characteristic Model

A. The MNL Model

Regression analysis is often used to analyze the correlation between explanatory variables and explained variables. When the explanatory variables are not of the complete numerical type and the explained variables are of multiple types, multiple Logistic regression model is adopted. In this paper, the basic utility function can be expressed as:

$$U_{aj} = V_{aj} + \beta_{aj} \quad (1)$$

where U_{aj} represents the utility function of a traveler a choosing the j^{th} travel mode, V_{aj} represents the observable influencing factors, namely the determined term, and β_{aj} represents the unobservable factors, namely the random term. When the random term takes a different distribution hypothesis, different models can be derived from it. According to the stochastic utility maximization theory, the probability of traveler a choosing the j^{th} travel mode can be expressed as:

$$P_{aj} = P\left(U_{aj} > \max U_{al}, j \neq l, j、 l = 1,2,3,4\right) \quad (2)$$

In this paper, the explanatory variables are travel modes, and there are four choices. To study the factors affecting residents' choice of travel modes, the MNL model is constructed, and the explanatory variables selected include individual attributes, family attributes, and travel characteristics, with a total of 12 variables. Three assumptions need to be made before building the model: (1) travelers always choose the most effective mode of travel. (2) the determined term and the random term are mutually independent; (3) the random terms are independent and uniformly distributed, and all obey the Gumbel distribution. Then the probability of traveler a choosing the j^{th} travel mode can be expressed as:

$$P_{aj} = \frac{\exp\left(V_{aj}\right)}{\sum_{j=1}^{4}\exp\left(V_{aj}\right)} \quad (3)$$

The determined term V_{aj} is considered as a combination of linear functions of various influencing factors, and its expression is as follows:

$$V_{aj} = \beta_0 + \beta_1 X_{aj1} + \beta_2 X_{aj2} + \cdots + \beta_{12} X_{aj12} \quad (4)$$

where β_0 is the constant term, $X_{aj1} \sim X_{aj12}$ are the 12 variable values that affect the traveler's choice of travel mode, and $\beta_1 \sim \beta_{12}$ are the undetermined coefficients of each variable value. Therefore, the probability of traveler a choosing the j^{th} travel mode is:

$$P_{aj} = \frac{\exp\left(V_{aj}\right)}{\sum_{j=1}^{4}\exp\left(V_{aj}\right)} = \frac{\exp\left(\beta_0 + \beta_1 X_{aj1} + \beta_2 X_{aj2} + \cdots \beta_{12} X_{aj12}\right)}{\sum_{j=1}^{4}\exp\left(\beta_0 + \beta_1 X_{aj1} + \beta_2 X_{aj2} + \cdots \beta_{12} X_{aj12}\right)} \quad (5)$$

The travel mode with the highest probability of being selected in the model result is considered as the final travel mode.

B. SVM Model

SVM is a method to transform a nonlinear separable problem into a linear separable problem by mapping the sample space to a higher dimensional eigenspace, it will sample data as the foundation, to training to learn to find the regularity of data, using the correlation between data to forecast the result of the unknown. Linear classification problems are generally solved by adding the hyperplane method, while nonlinear classification problems need to be solved by adding kernel function to assist sample data mapping. For all sample data x_i, let,

$$f(x)=\omega^T\varphi(x)+b \tag{6}$$

Then the objective function can be expressed as:

$$\min\frac{1}{2}\|\omega\|^2$$

$$\text{s.t. } y\left(\omega^T\varphi(x)+b\right)\geq 1 \tag{7}$$

Its dual function is:

$$\max\sum_i a_i-\frac{1}{2}\sum_i\sum_j a_i a_j y_i y_j \varphi(x_i)^T\varphi(x_j)$$
$$\text{s.t.}\begin{cases}\sum_i a_i y_i=0\\ a_i\geq 0\end{cases} \tag{8}$$

To solve the dual problem, let the kernel function be $K(x_i,x_j)=\langle\varphi(x_i),\varphi(x_j)\rangle=\varphi(x_i)^T\varphi(x_j)$, then the above equation can be expressed as:

$$\max\sum_i a_i-\frac{1}{2}\sum_i\sum_j a_i a_j y_i y_j K(x_i,x_j)$$
$$\text{s.t.}\begin{cases}\sum_i a_i y_i=0\\ a_i\geq 0\end{cases} \tag{9}$$

The solution can be obtained as follows:

$$f(x)=\omega^T\varphi(x)+b=\sum_i a_i y_i K(x_i,x_j)+b \tag{10}$$

According to different sample data and requirements, there are four main kernel functions of SVM:

1. Linear kernel function:

$$K(x_i,\mathrm{x}_j)=\left(x_i^T x_j\right)^1 \tag{11}$$

2. Polynomial kernel function:

$$K(x_i,\mathrm{x}_j)=\left(x_i^T x_j\right)^d \tag{12}$$

where d is the degree of a polynomial, which is reduced to a linear kernel when $d = 1$.

3. Kernel function of radial basis:

$$K(x_i,\mathrm{x}_j)=\exp\left(-r\frac{\|x_i-x_j\|^2}{\sigma^2}\right)\cdots\cdots(\sigma>0) \tag{13}$$

4. Sigmiod kernel function:

$$K(x_i,\mathrm{x}_j)=\tanh\left(\lambda x_i^T x_j+\theta\right)\cdots\cdots(\lambda,\theta>0) \tag{14}$$

After the kernel function is introduced, the relaxation variable $\gamma\geq 0$ and the penalty factor C > 0 need to be introduced to process the sample data that is far from the normal position. The larger the value of C is, the greater the penalty for classification error is, then the objective function is modified as:

$$\min\frac{1}{2}\|\omega\|^2+C\sum_i\gamma_i$$
$$\text{s.t. } y_i(\omega^T\varphi(x_i)+b)\geq 1-\gamma_i \tag{15}$$

After writing the Lagrangian function, it is transformed into a dual problem to find the optimal parameter model.

C. MLP Deep Learning Model

MLP is also commonly referred to as an artificial neural network. It is a multi-layer feedforward neural network model, including input layer, output layer, and multiple hidden layers. It can obtain a target value by combining multiple feature values. The MLP structure diagram with three hidden layers is shown in Fig. 1.

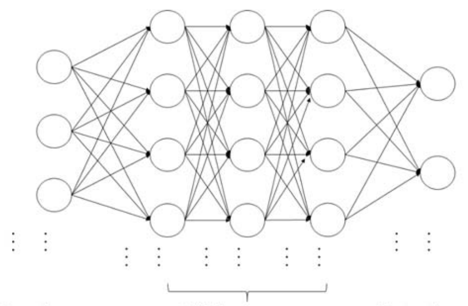

Fig. 1 MLP model structure diagram

In the diagram, the input layer is on the left, three hidden layers are in the middle, and on the right is the output layer. The MLP is fully connected from layer to layer. Input layer neurons are responsible for receiving information. If we input a 256-dimensional vector, we have 256 neurons. The hidden layer neurons are responsible for processing the input information. If the input layer is represented by the vector X, the output form of the hidden layer can be expressed as:

$$X_1=f(W_1X+b_1) \tag{16}$$

where W_1 is the connection coefficient, also known as the weight matrix. b_1 is the bigotry vector of the layer, and function f() is the nonlinear activation function. In this experiment, we adopted the ReLU function, and the hidden layer dimension is 512 dimensions.

The output layer neuron is responsible for the computer's cognition of the input information. The hidden layer to the output layer can be regarded as a multi-category logistic regression, so the output of the hidden layer is Softmax ($\mathrm{W}_4\mathrm{X}_3+\mathrm{b}_4$), where a is the output of the three hidden layers., Therefore, the above MLP model can be summarized as:

International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:14, No:9, 2020

$$f(X) = \text{Softmax}(W_4 * f(W_3 * f(W_2 * f(W_1 * X + b_1) + b_2) + b_3) + b_4) \quad (17)$$

The parameters used by MLP are the weight matrix W and the bias vector p of each layer. The determination of these parameters is the optimization problem to solve the optimal parameters. The loss function we adopted is the cross-entropy loss function, whose expression is:

$$L = \sum_{c=1}^{M} y_c \log p_c \quad (18)$$

where M is the number of categories, y_c is whether the variable is the same as the observation (if the same, it is 1), and the probability that the p_c observation sample belongs to category c.

III. Behavior Analysis of Travel Mode Selection

A. Data Preparation

In order to better study the travel characteristics of urban residents, a questionnaire survey was conducted in Futian district of Shenzhen in February 2019. The sample was randomly selected and covered all residential districts. The sample size was proportional to the population of each residential district. After investigation, a total of 4,673 questionnaires were collected, and a total of 3,446 valid questionnaires were obtained through a preliminary screening. After collating and summarizing the collected data, the proportional distribution of the six basic information such as gender and age of the surveyed traveling residents is obtained, as shown in Table I.

TABLE I
CHARACTERISTIC DISTRIBUTION OF RESIDENTS' TRAVEL INFORMATION

Basic information	Category	Sample size	Percentage (%)
Gender	Male	1974	57.3
	Female	1472	42.7
Age	[10~18)	47	1.4
	[18~25)	482	14.0
	[25~35)	1461	42.4
	[35~45)	967	28.1
	[45~)	489	14.2
Job	Student	205	5.9
	Full-time employee	2105	61.1
	Temporary employee	353	10.2
	Freelancer	520	15.1
	Retiree	263	7.6
Education	Junior high school or below	383	11.1
	Technical secondary school and high school	1208	35.1
	College and bachelor	1675	48.6
	Postgraduate or above	180	5.2
Income	<50000	692	20.1
	50000~100000	1328	38.5
	100000~200000	960	27.9
	>200000	466	13.5
Number of cars	0	2080	60.4
	1	1268	36.8
	2	98	2.8

As can be seen from Table I, the proportion of male respondents is 57.3%, slightly higher than that of female respondents (42.7%). Most of them are between 25 and 45 years old, accounting for 70.5%, which means that the majority of young and middle-aged travelers participated in the survey. The percentage of full-time travelers was the highest at 61.1%. The education level of respondents mainly focuses on middle and higher education, so the data obtained from the survey are representative to some extent. In this part of the traveling population, the proportion of the household annual income of "50,000-100,000" is the highest at 38.5%, followed by "100,000-200,000" at 27.9%. The data distribution is reasonable. 60.4% of respondents did not own a car, 36.8% owned one car and 2.8% owned two cars, within the normal range. According to the travel habits of local residents, there are four travel modes in the questionnaire, namely public transportation (including subway, bus, and taxi), private car, non-motor vehicle, walking, and other means. The results of the questionnaire show that the distribution of travel sharing rate, travel time and travel distance of the travelers in this part of the survey is shown in Fig. 2.

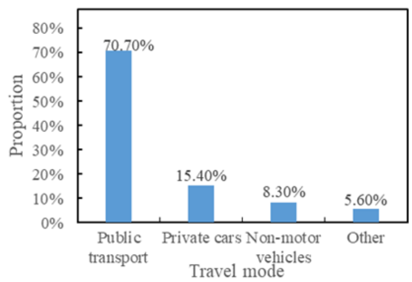

Fig. 2 (a) Distribution of travel mode sharing rate

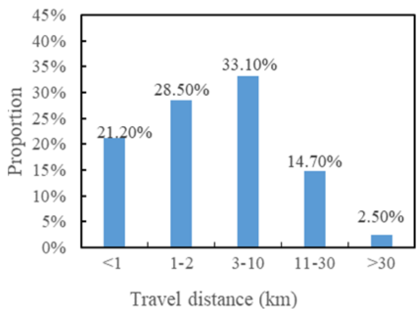

Fig. 2 (b) Distribution of travel distance

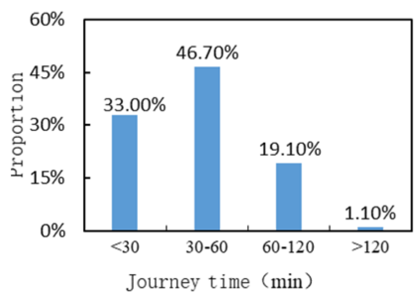

Fig. 2 (c) Distribution of journey time

It can be seen that most of the travelers surveyed chose public transportation, accounting for 70.7%. The second choice is private cars, accounting for 15.4%. Then non-motor vehicle travel accounts for 8.3%; walking and other travel modes was the lowest, at just 5.6%. It is seen that local residents prefer to choose public transportation, and the proportion of choosing non-motorized travel is very low; this is due to the lack of non-motorized lanes. If residents choose to travel by non-motorized vehicles, they will face many safety risks, and they will therefore switch to private cars and other relatively more convenient means to travel. In addition, more than 50% of residents travel more than 3 kilometers, and most of them still choose public transportation, which cannot solve the problem of "the last kilometer" in long-distance travel, which will also increase the journey time and bring inconvenience to travel.

B. Model Fitting and Result Analysis

Based on the characteristics of the model, variables need to be redefined before data are introduced. A total of 12 explanatory variables including individual attributes, family attributes, and travel characteristics were extracted to analyze their effects on the explained variables. The definitions of each variable are shown in Table II.

TABLE II
VARIABLE DEFINITIONS

Category	Variable	Variable definitions
Individual attributes	Gender	1 = female, 2 = male
	Age	1 = [10,18)years old, 2 = [18,25)years old, 3 = [25,35)years old, 4 = [35,45)years old, 5 = [45, 100)years old
	Job	1 = student, 2 = full-time employee, 3 = temporary employee, 4 = freelancer, 5 = retiree, 6 = others
	Education	1 = junior high school or below, 2 = technical secondary school and high school, 3 = college and bachelor, 4 = postgraduate or above
	Income	1 = (0,50000], 2 = (50000,100000], 3 = (100000,200000]; 4 = (200000,500000], 5 = more than 500000
Family attributes	Number of cars	Continuous variable, unit: vehicle
	Family members	Continuous variable
Travel characteristics	Travel purpose	1 = work/school, 2 = to have a meal, 3 = medical/fitness, 4 = to visit friends and relatives, 5 = shopping and other things
	Departure time	1 = before 6:00AM, 2 = during 6:00AM-7:00AM, 3 = during 7:00AM-8:00AM, 4 = during 8:00PM-12:00AM, 5 = after 12:00AM
	Journey time	1 = within 0.5h, 2 = 0.5h-1h, 3 = 1h-2h, 4 = more than 2h
	Travel distance	1 = under 1 km, 2 = 1-2 km, 3 = 3-10 km, 4 = 11-30 km, 5 = more than 30 km
	Travel times	Continuous variable, unit: times

In order to better analyze the influence of variables, the four travel modes as the explained variables are defined as 1 = public transportation (including bus, subway, and taxi), 2 = private car travel, 3 = non-motor vehicle travel, 4 = walking and other modes. MNL model was calibrated with the help of SPSS.24 software. 12 influencing factors were added to the factor column, "travel mode" was added to the dependent variable column, variable attributes were adjusted, and reference categories were set. The system defaults to the last category of each variable as a reference. After parameter estimation, the model fitting effect is shown in Table III.

As can be seen from Table III, in the likelihood ratio test results of the calibrated MNL model and the model with the only intercept, the significance is less than 0.05, indicating that the final model has statistical significance and the model is established. The three pseudo-r square values of the model output are 0.352, 0.422 and 0.241 in turn, and the pseudo-r square values are generally between 0.3 and 0.5, which will neither be too low in accuracy nor over-fit. Therefore, the fitting degree of the model is relatively good, and it can be fitted reasonably according to different data.

According to the fitting results, the significance level of each explanatory variable is less than 0.05, so the null hypothesis of "removing a certain factor from the model does not affect the coefficient in the model" is rejected, indicating that the influence of these variables on the explained variables is significant.

The partial parameter estimation results of choosing "walking and other modes of travel" as the reference term are shown in Table IV. Among all the variables affecting residents' choice of travel mode, if the significance level corresponding to a variable is less than 0.05, it indicates that the variable has a very significant impact on residents' choice of this travel mode. If greater than 0.05, the variable should be removed from the model. B is listed as the coefficient of the variable in the model, and the positive and negative situation indicates that the variable plays a positive/negative role in the selection of the travel mode. Only the influential factors with significant effects are shown in the table.

The following conclusions can be drawn from the analysis of parameter estimation results:

(1) When the behavior reference items were determined by walking and other means, the residents' education level (= 1), number of cars owned (= 0/1), departure time (= 3) and travel times (= 2) had a significant influence on their choice of public transportation. The system estimate for households without cars was 1.537, indicating that residents without cars were more likely to use public transport; when residents travel later and less frequently, the probability of choosing public transportation is higher. The systematic estimate of the population's education level (= 1) is negative, indicating that less educated residents are more likely to travel by foot and other means.

(2) Gender, occupation, education level, population, income,

travel purpose, departure time and travel distance all have a significant influence on residents' choice of private car travel. The larger the family size, the more likely the residents are to choose private cars for travel because the carrying capacity of private cars can meet the needs of many people, which is more convenient and safe than other travel methods. The earlier the departure time, the farther the travel distance, residents are more inclined to choose private cars to travel because private cars are more suitable for long-distance travel, faster speed; however, residents with lower education levels and income generally do not choose private cars for travel. On the one hand, they may not be able to afford the cost of a car due to their limited income; on the other hand, their daily demand for private cars is not high, so they tend to choose other ways of travel.

(3) Gender, departure time, journey time and travel distance have a significant influence on residents' choice of non-motor vehicle travel. Compared with walking, women are more likely to choose non-motor vehicles for travel, because the speed of non-motor vehicles is faster than walking and with the popularity of sharing bikes, non-motor vehicles are convenient and flexible in travel, suitable for short-distance daily travel. When the travel time and distance are shorter, it is more reasonable for residents to choose the way of walking. However, due to the lack of local non-motorized lanes, the overall proportion of residents choosing non-motor vehicles for travel is relatively low. In this case, to meet their daily travel needs, residents are more likely to choose public transportation or private cars for travel.

The model has the highest accuracy in predicting the travel choice tendency of public transportation. The overall accuracy of the model is 72.8%, and the prediction effect is mediocre.

TABLE III
EFFECT OF MNL MODEL FITTING

Model fitting information						
Model	Model fitting conditions		Likelihood ratio		Pseudo R-Square 2	
	-2 log likelihood	Chi-square	degree freedom	Significant	Cox and Snell	.352
Only intercept	6153.142				Nagelkerke	.422
Final	4657.665	1495.477	126	.000	McFadden	.241

TABLE IV
ESTIMATION OF MNL MODEL PARAMETERS

Mode of transport		B	S.E.	Wald	D.F.	Sig.	Exp(B)
Public transport	Intercept	13.943	1996.153	.000	1	.994	
	[Education = 1]	-1.610	.585	7.569	1	.006	0.200
	[Car = 0]	1.537	.448	11.760	1	.001	4.652
	[Car = 1]	.957	.452	4.486	1	.034	2.604
	[Departure time = 3]	.779	.291	7.165	1	.007	2.180
	[Travel times = 2]	1.110	.556	3.987	1	.046	3.033
Private cars	Intercept	1.825	2749.057	.000	1	.999	
	[Gender = 1]	.724	.196	13.673	1	.000	2.062
	[Job = 4]	1.054	.402	6.867	1	.009	2.868
	[Education = 1]	-3.132	.639	24.053	1	.000	0.044
	[Education = 2]	-2.094	.587	12.743	1	.000	0.123
	[Education = 3]	-1.205	.574	4.401	1	.036	0.300
	[Population = 3]	.825	.359	5.291	1	.021	2.282
	[Population = 4]	.885	.356	6.172	1	.013	2.424
	[Income = 1]	-1.231	.381	10.429	1	.001	0.292
	[Income = 2]	-.820	.346	5.608	1	.018	0.440
	[Travel purpose = 2]	1.515	.544	7.747	1	.005	4.548
	[Travel purpose = 3]	1.154	.498	5.371	1	.020	3.171
	[Departure time = 1]	2.008	.540	13.828	1	.000	7.451
	[Departure time = 2]	1.657	.426	15.152	1	.000	5.246
	[Departure time = 3]	1.517	.411	13.655	1	.000	4.561
	[Departure time = 4]	1.063	.419	6.446	1	.011	2.895
	[Travel distance = 3]	1.969	.720	7.480	1	.006	7.164
Non-motor vehicles	Intercept	-1.003	1.492	.452	1	.502	
	[Gender = 1]	.401	.204	3.874	1	.049	1.493
	[Departure time = 3]	1.656	.426	15.120	1	.000	5.239
	[Departure time = 4]	1.662	.426	15.195	1	.000	5.269
	[Journey time = 1]	-2.968	.743	15.965	1	.000	0.051
	[Journey time = 2]	-2.659	.740	12.924	1	.000	0.070
	[Journey time = 3]	-1.555	.763	4.148	1	.042	0.211
	[Travel distance = 2]	2.149	.961	4.997	1	.025	8.577
	[Travel distance = 3]	2.768	.972	8.104	1	.004	15.924

IV. EVALUATION OF POLICY IMPROVEMENT EFFECT

In this paper, the traditional non-lumped model (MNL) is used to analyze the travel behavior of urban residents, and the results show that the departure time, journey time and travel

International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:14, No:9, 2020

distance are all significant factors influencing residents' choice of travel behavior. To improve the efficiency, convenience and satisfaction of residents' travel, considering the policy orientation of green travel and environmental protection, it is proposed to put forward policy improvement suggestions from the aspect of transportation facilities allocation. In this paper, the SVM model and multi MLP deep learning model are respectively used for model fitting and parameter estimation, and the fitting accuracy of the MNL model is compared. The higher ones are selected for the effect evaluation after policy improvement.

When the SVM model is used for prediction, the radial basis kernel function with less deviation is selected in this paper to achieve higher prediction accuracy. The data sample is divided into two parts, one as training set and the other as verification set. In practice, it is divided into the training set and verification set according to 8:2, the value range of parameter C is set as [0.01,1000], and the value range of parameter a is set as [0.01,1000]. After parameter optimization, it is found that when C is set as 100 and a is set as 0.01, the classification accuracy is the highest, about 71.42%.

When using MLP deep learning model for prediction, we use an Optimization Problem to solve the optimal parameter. Generally, the simplest method to solve optimization problems is Stochastic Gradient Descent (SGD), that is, all parameters are randomly initialized at first, and then iterative training is conducted to continuously calculate the gradient and update parameters until the error is small enough or the number of iterations is large enough. The output of the model shows that the prediction accuracy of the model can reach 76.42%.

The result pairs of the above three models are shown in Table V. Therefore, the MLP model with the highest accuracy is selected to predict the effect of policy improvement.

TABLE V
COMPARISON OF THE RESULTS OF THREE MODELS

Model	Accuracy
MNL	72.80%
SVM	71.42%
MLP	76.42%

In order to improve the travel convenience, reduce the wasted time in the journey, and strengthen the connection between various modes of transportation, the following two policy improvement suggestions are proposed from the perspective of transportation resource allocation:

(1) Add non-motorized lanes. The lack of non-motorized lanes on most roads in the city hinders residents' choice of non-motorized vehicle travel. Therefore, it is suggested to add non-motorized lanes and rationally plan the network structure of non-motorized lanes, so as to alleviate the "last kilometer" problem, enhance the convenience of residents' travel and improve their satisfaction.

(2) Improve the bus service frequency. Investigation found that there is public transportation passenger flow in the area of the peak time crowded phenomenon, part of the passengers could not get on in time and have to wait for the next bus, this will prolong the waiting time, lead to the waste of travel time. It is recommended for large passenger flow region or site, appropriate to improve the bus service frequency. Increasing the number of bus lines and subway lines can better share the traveling passenger flow pressure, and encourage people to use public transportation.

Assuming that scheme (1) can reduce the travel time of non-motor vehicles by 1 unit and the travel distance by 1 unit, and scheme (2) can reduce the travel time of public transportation by 1 unit, the MLP model is used to predict the proportion of the improved travel mode, as shown in Table VI.

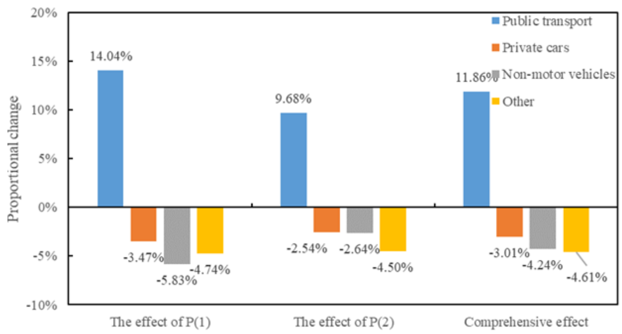

Fig. 3 Changes in the sharing rate of various modes of travel

TABLE VI
RESULTS CONTRAST

Travel mode	Sample (quantity/proportion)	The effect of P(1)	The effect of P(2)	Comprehensive effect
Public transport	2437(70.70%)	2920(84.74%)	2770(80.38%)	2845(82.56%)
Private cars	532(15.40%)	411(11.93%)	443(12.86%)	427(12.39%)
Non-motor vehicles	284(8.30%)	85(2.47%)	195(5.66%)	140(4.06%)
Other	193(5.60%)	30(0.86%)	38(1.10%)	34(0.99%)

As can be seen from Table VI, in the predicted results of travel modes, under the influence of scheme (1), the number of public transportation trips increased by 483, the number of private car trips decreased by 121, the number of non-motor vehicle trips decreased by 199, and the proportion of other trips decreased by 163. Under scheme (2), the number of public transportation trips increased by 333, the number of private car trips decreased by 89, the number of non-motor vehicle trips decreased by 89, and the proportion of other trips decreased by 155. Under the combined effect of the two schemes, the number of public transport trips increased by 408, the number of private car trips decreased by 105, the number of non-motor vehicle trips decreased by 144, and the proportion of other means of travel decreased by 159. Under different schemes, the change of share rate is shown in Fig. 3.

It can be seen that after the implementation of the two proposed policies, the proportion of public transportation trips increased by 14.04% and 9.69%, respectively, while the proportion of private car trips decreased by 5.83% and 2.64%, respectively, indicating that the two plans are conducive to promoting residents to choose public transportation and reducing the proportion of private car trips. In scheme (1), the addition of non-motorized lanes does not directly promote the proportion of non-motorized vehicle travel, the reason is that this measure allows the residents to more easily transfer between the bus, subway and bike in the process of the whole travel, which solves the problem of "last kilometer" travel, and guides people choose greener public transportation options; thus, positive effects indirectly reflected in the increase of the proportion of public transportation. Under the combined effect of the two schemes, the proportion of public transport trips will increase by 11.86%, while the proportion of private car trips will decrease by 4.61%. Therefore, adding non-motorized lanes and increasing public transport resources will play a significant role in improving the travel structure and promoting the development of "green" transportation.

V. Conclusion

This paper builds the MNL model research of Futian, Shenzhen district residents' travel behavior characteristics, using the random utility maximization theory, modeling parameter estimation by the SPSS software, takes walking and other travel modes as references, and shows that the time of departure, travel times, journey time and trip distance will significantly impact residents travel mode choice. That is to say, the earlier the departure time, the higher the probability that residents will choose a private car to travel, and the later the departure time, the higher the probability that residents will choose a bicycle/electric vehicle; the fewer the number of trips, the higher the probability that residents will choose public transportation. The shorter the distance, the more the residents tend to walk; the longer the travel distance, the more the residents tend to travel in private cars or non-motorized vehicles. From the perspective of optimizing the resources distribution of traffic based on the above results, we suggest two kinds of policies to improve: To assess the effect of the policy after implementation, the introduction of SVM model and MLP deep learning model used for predicting the residents travel mode under the optimization measures. The results show that the proposed policy improvement suggestions can help increase the proportion of public transportation trips, increase the flexibility of various modes of transportation, and enhance residents' satisfaction when traveling, and can provide a certain reference for relevant departments to formulate transportation policies. Since residents' travel behaviors are diverse and uncertain and may be influenced by many factors, deeper factors should be further considered. We will optimize of urban residents' travel environment from various aspects in future studies.

Acknowledgments

This research was supported by the Fundamental Research Funds for the Central Universities (NO. 2019YJS107) and National Natural Science Foundation of China (No. 51678044).

References

[1] Ming He, Xiucheng Guo, Jiangyu Ran, Cairui Wu, Wei Zhu, Chaoping Liu. Forecasting Rail Transit Split with Disaggregated MNL Model (J). Journal of Transportation Systems Engineering and Information Technology, 2010, 10(2).
[2] Eran Ben-Elia, Yoram Shiftan. Which road do I take? A learning-based model of route-choice behavior with real-time information (J). Transportation Research Part A, 2010,44(4).
[3] Yang Liya, Shao Chunfu, Haghani A. Nested Logit Model of combined selection for travel mode and departure time (J). Journal of Traffic and Transportation Engineering, 2012,12(02):76-83.
[4] Vo Van Can. Estimation of travel mode choice for domestic tourists to Nha Trang using the multinomial probit model (J). Transportation Research Part A, 2013,49.
[5] Cheng Long. Mode choice of low income commuters and the evaluation of their improvement measures in big cities (D). Southeast University, 2016.
[6] Ge Dapeng. Study on the traffic mode split of Metropolitan intercity passenger transport based on travel behavior analysis (D). Chang'an University, 2017.
[7] Jiang Wei. Construction and application of resident's car rental trip choice model based on dynamic factors (D). Chongqing Jiaotong University, 2017.
[8] Lin Xiaomei, Shao Chunfu, Dong Chunjiao, Wang Shengyou. The behavior characteristics of Inter-city travel under the implication of expressway Toll-free policy during holiday (J). Journal of Transportation Systems Engineering and Information Technology, 2019, 19(02):247-254.
[9] Ransford A. Acheampong. Spatial structure, intra-urban commuting patterns and travel mode choice: Analyses of relationships in the Kumasi Metropolis, Ghana (J). Cities, 2020,96.
[10] Xu Bing. Research on the ICLV model of intercity travel mode selection (D). Beijing Jiaotong University, 2018.
[11] Liu Yufeng, Qian Yizhi, Hu Dawei, Wang Laijun, Li Lu, Ma Zhuanglin. Correlation analysis of travel mode choice for urban residents in different urban size based on structural equation model (J). Journal of Chang'an University (Social Science Edition), 2018, 38(05):87-95.
[12] Xiugang Li, Dominique Lord, Yunlong Zhang, Yuanchang Xie. Predicting motor vehicle crashes using Support Vector Machine models (J). Accident Analysis and Prevention, 2008, 40(4).
[13] Xing Wang, Xin Wang, Zhaonan Sun. Comparison on Confidence Bands of Decision Boundary between SVM and Logistic Regression (P). INC, IMS and IDC, 2009. NCM '09. Fifth International Joint Conference on, 2009.
[14] Wang Shengyou. Analysis of trip characteristics and optimal design of traffic resources on Beijing Metropolitan Area (D). Beijing Jiaotong University, 2018.
[15] Yan Jialin, Xiang Longgang, Wu Huayi, Sun Shangyu. Urban road traffic speed prediction based on LSTM (J). Geomatics World, 2019, 26(05):79-85.
[16] Cao Yu, Wang Cheng, Wang Xin, Gao Yueer. Urban road short-term traffic flow prediction based on Spatio-temporal node selection and deep learning (J/OL).Journal of Computer Applications: 1-10(2020-01-13).