International Journal of Engineering Research And Management (IJERM)
ISSN: 2349- 2058, Volume-10, Issue-12, December 2023

Characterization of Transit Travel Behavior of Potential Category Groups

Caixia Chen, Xiaokun Wang, Chen Zhang

***Abstract*—Based on the consideration of individual heterogeneity, it is necessary to analyze the characteristics of bus transit travel behaviour of different latent class groups. Firstly, by constructing a latent class model for distinguishing travel groups, the questionnaire was designed from four aspects: household attributes, personal attributes, travel attributes, and suggestions for the development of urban buses. Then, based on the 2,000 valid questionnaires recovered, 10 indicators, such as the total number of people in the household, the total income of the household, the gender of the individual, the age of the individual, and the time spent on walking, etc., were selected as exogenous variables, and a latent categorization was made based on the indexes of the goodness of fit for the travel groups. The results of the study show that: the bus travel group can be divided into three latent classes, with the student group aged 6 to 17, the male worker group aged 23 to 34, and the female worker group aged 35 to 49, accounting for 6.8%, 42.8%, and 50.4% of the total respectively. There are some differences in the travel behaviour of the three groups, and by analyzing the travel characteristics of each group, suggestions are made for the development of urban buses in terms of rationally setting up bus routes and stops, adjusting the operating hours and implementing "just-in-time buses", improving the environment for riding and transferring, and reducing the cost of travelling on buses.**

***Keywords*—Traffic engineering, Bus transit, Travel behavior, Individual heterogeneity, Latent class model**

I. INTRODUCTION

Promoting the priority development of public transportation is an important way to build a resource-saving and environmentally friendly society and realize sustainable development. In order to avoid the huge economic losses caused by urban traffic congestion, it is necessary to analyze and study the travel behavior characteristics of different public transport travel groups.

In recent years, it has become a research hotspot to consider the effect of traveler heterogeneity on travel mode choice. Heterogeneity was first proposed by Böckenholt and Dillon [1]. Chen [2] divided the surveyed population into three potential categories through the potential category model and analyzed the characteristics of each category group. Tao W C et al [3] selected gender, age, travel purpose, and average number of trips as exogenous variables to study the market segmentation of intercity high-speed railroad passenger transportation by constructing a potential category model. Li D L et al [4] selected exogenous variables from both socio-economic attributes and psychological factors, constructed a potential category model, and provided a basis for the study of road passenger heterogeneity through probabilistic parameterization and model estimation. Zhou Y B et al [5] used the potential category model to segment the residents near rail transit stations to provide ideas for the planning and layout of transportation facilities around rail transit stations.

For the study of public transportation mode choice, certain research results have also been achieved at home and abroad. Liu J R et al [6] used validation factors to analyze the impact of the new crown epidemic on the bus travel behavior of the elderly. Fresh Yu J C [7] studied the impact of the number of daily trips and travel time on commuters' travel mode choice. Tian L [8] investigated the effects of the number of daily trips and age on the travel mode choice behavior of the elderly. He L et al [9] obtained survey data through scenario hypotheses to study the travel mode preferences of different groups of people for public transit, and put forward corresponding policy recommendations for each group. Wang L W [10] et al. classified commuters into standard commuters, flexible commuters, and two kinds of low-frequency commuters by using a clustering method. commuters and two kinds of low-frequency passengers, and the analysis concluded that it is feasible to improve the level of public transportation services based on passengers' travel modes. Li P et al [11] derived the correlation between passenger flow and influencing factors through clustering analysis to provide reference opinions for public transportation development. In the study of group classification, cluster analysis is the most commonly used method, however, traditional cluster analysis often has some shortcomings [12], such as the effect of clustering cannot be guaranteed, and the cost of calculation time is large.

In summary, the established literature on public transport mode choice is not yet comprehensive, in order to more carefully analyze the public transport mode choice behavior of urban residents, it is necessary to carry out a heterogeneous study on the characteristics of travel behavior of different travel groups. This paper takes the potential category model as the basis, based on 2000 valid questionnaires recovered in Yingkou City, Liaoning Province, China, and takes the variables corresponding to the respondents' family attributes, personal attributes, and public transit travel attributes as the exogenous variables, and conducts a segmentation research on the categories of travelers choosing public transit travel, so as to enhance the accuracy of the model and the

Manuscript received November 30, 2023

Caixia Chen, School of Traffic and Transportation Engineering, Dalian Jiaotong University, Dalian, Liaoning, China
Xiaokun Wang, School of Economics and Management, Dalian Jiaotong University, Dalian, Liaoning, China
Chen Zhang, School of Civil Engineering and Transportation, Northeast Forestry University, Harbin, Heilongjiang, China

reasonableness of the categorization.

II. LATENT CLASS MODEL

A. Probabilistic parameterization and model construction

Latent class model (LCM) consists of exogenous variables and latent variables, and is mainly used to study the relationship between exogenous variables and latent variables. Compared with traditional cluster analysis, LCM has the advantage that there is no need to preset the number of clusters, and the optimal number of clusters can be determined by various goodness-of-fit indicators [13].

The probability of the latent variable X at the t level., the probability of indicating that the user belongs to category t in the classification of residential travelers π_t^X is the latent category probability. Assuming that there are three exogenous variables, A、B、C whose level values are i、j、k, the conditional probabilities of each exogenous variable are $\pi_{it}^{\bar{A}X}$, $\pi_{jt}^{\bar{B}X}$, and $\pi_{kt}^{\bar{C}X}$ respectively. The levels of each exogenous variable are independent of each other.

The objective function of the latent class model is:

$$\pi_{i,j,k}^{A,B,C} = \sum_{t=1}^{T} \pi_t^X \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \tag{1}$$

The constraints are:

$$\sum_t \pi_t^X = 1.00 \tag{2}$$

$$\sum_t \pi_{it}^{\bar{A}X} = \sum_t \pi_{jt}^{\bar{B}X} = \sum_t \pi_{kt}^{\bar{C}X} = 1.00 \tag{3}$$

$\pi_{i,j,k}^{A,B,C}$ is the joint probability. In general, the magnitude of the potential category probability π_t^X is used to indicate the weighting at different levels of the potential category variable X, which can be used to indicate the size of the subdivided category group, while the conditional probabilities are used more to explain the characteristics and practical significance of the investigators included in each potential category.

B. Parameter Estimation and Model Fit

(1) Maximum Likelihood Estimation

Solving the value of each parameter in the LCM model mainly adopts the method of great likelihood (Maximum Likelihood), and the EM (Expectation-Maximization) algorithm is often used in the iterative process. Assuming that in an LCM model, the potential variables have a potential category, the maximum likelihood function is:

$$\hat{\pi}_{ijk}^{ABCX} = \hat{\pi}_t^X \hat{\pi}_{it}^{\bar{A}X} \hat{\pi}_{jt}^{\bar{B}X} \hat{\pi}_{kt}^{\bar{C}X} \tag{4}$$

$\hat{\pi}_t^X$ is the probability estimate and $\pi_{it}^{\bar{A}X}$, $\pi_{jt}^{\bar{B}X}$, and $\pi_{kt}^{\bar{C}X}$ are the maximum likelihood estimates (MLE latent class probability) of the latent class probabilities of the exogenous variables (conditional probabilities). By summing the probability estimates of each exogenous variable under each potential class, we can obtain the maximum likelihood estimate of the joint probability, and the expected probability that can be used to test the fitness of the model is:

$$\hat{\pi}_{ijk}^{ABC} = \sum_{t=1}^{T} \hat{\pi}_{ijk}^{ABCX} = \sum_{t=1}^{T} \hat{\pi}_t^X \hat{\pi}_{it}^{\bar{A}X} \hat{\pi}_{jt}^{\bar{B}X} \hat{\pi}_{kt}^{\bar{C}X} \tag{5}$$

If (4) is divided with (5), the weight of each level of each exogenous variable on a potential category is obtained, which can be used as a comparison of the relative importance of each exogenous variable on a potential category [14].

$$\hat{\pi}_{ijk}^{ABC\bar{X}} = \hat{\pi}_{ijk}^{ABCX} \,/\, \hat{\pi}_{ijk}^{ABC} \tag{6}$$

$\bar{X}$ in $\hat{\pi}_{ijk}^{ABC\bar{X}}$ denotes the sum of probabilities for each potential category, and $\hat{\pi}_{ijk}^{ABC\bar{X}}$ denotes the weight of each level of each exogenous variable in a potential category.

(2) Model fitness test

In order to reduce the heterogeneity existing in the original data, this paper determines the optimal potential number of categories for the goodness-of-fit metrics by segmenting the original dataset using the latent category model, including Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Adjusted Bayesian Information Criterion (ABIC), Entropy, LMRT (VUONG-LO-MENDELL-RUBIN LIKELIHOOD RATIO TEST) and BLRT (PARAMETRIC BOOTSTRAPP-ED LIKELIHOOD RATIO TEST).The smaller the value of AIC, BIC, and ABIC, the better the results of the model fit.If the three values are reduced synchronously, the reduction rate of BIC can be considered as the basis for judgment. The higher the entropy value, the better the effect of potential classification, and when the entropy value is closer to 1, the heterogeneity among potential categories is the largest.LMRT and BLRT are the reported P-values, and when the value is less than 0.05, it indicates that the classification is significant and the accuracy of the classification is greater than 90%, which is superior to the classification model that is smaller than the number of categories.

C. Latent Class

The ultimate goal of using the latent class model (LCM) is to categorize the collected data into appropriate latent classes, and according to Bayesian theory, the travelers are categorized into different latent classes by calculating the posterior probability of each data. The categorization probability is calculated as in the formula below:

$$\hat{\pi}_{ijk}^{\bar{X}ABC} = \hat{\pi}_{ijk}^{ABCX} \,/\, \sum_{t=1}^{T} \hat{\pi}_{ijk}^{ABCX} \tag{7}$$

$\hat{\pi}_{ijk}^{\bar{X}ABC}$ denotes the joint probability of each of the exogenous variables for that traveler; the posterior probability that a given user belongs to each of the potential categories can be computed, and the size of the posterior probability can subsequently be used to determine the potential category to which the user should belong.

III. QUESTIONNAIRE DESIGN AND SURVEYS

The questionnaire design mainly investigates four aspects of travelers' family attributes, personal attributes, travel attributes, and suggestions for the development of urban public transportation, and the variables corresponding to family attributes, personal attributes, and travel attributes are used as exogenous variables to categorize public transportation travelers, and the specific contents of the survey are shown in Table 1.

Table 1. Levels of travelers' categorical attributes

Attributes	Variable	Level
Family attributes	Total number of persons in household	1 for [1,2]; 2 for [3,4]; 3 for [5,6]; 4 for 7 and above
	Annual household income	Less than 20,000 yuan for 1; [20,000,40,000] yuan for 2; [40,001,60,000] yuan for 3; [60,001,100,000] yuan for 4; more than 100,000 yuan for 5
	Nature of residence	1 for household registration; 2 for non-household registration for more than half a year; 3 for non-household registration for less than half a year; 4 for less than a month for business trips, travel, and family visits
	Private car ownership	1 for none; 2 for [1,2] vehicles; 3 for 3 or more vehicles
Personal Attributes	Sex	1 for male; 2 for female
	Age	1 for [6,14]; 2 for [15,17]; 3 for [18,24]; 4 for [25,34]; 5 for [35,39]; 6 for [40,49]; 7 for [50,59]; and 8 for 60 and over;
	Occupation	1 for students; 2 for employees; 3 for teachers; 4 for civil servants; 5 for self-employed; 6 for medical personnel; 7 for farmers; 8 for others.
Travel Attributes	Walking time	1 for less than 3 minutes; 2 for [3,5] minutes; 3 for [6,10] minutes; 4 for [11,15] minutes; 5 for 15 minutes or more
	Ride Frequency	Rarely rides (no rides or [1,2] rides) is 1; [3,5] rides a week is 2; [6,10] rides a week is 3; more than 10 rides a week is 4
	Number of trips yesterday	1 for [1,2] trips; 2 for [3,4] trips; 3 for [5,6] trips; 4 for 7 or more trips

Note: Walking time is specifically defined as the walking time from work or school to the actual bus stop where you ride the bus

Table 2. Statistics of information on each variable under family attributes

Attribute	Variable Description	Sample Size	Sample Proportion
Total number of households (V2)	[1,2]	72	3.60%
	[3,4]	165	8.25%
	[5,6]	1117	55.85%
	7 or more	646	32.30%
Annual household income (V3)	Less than 20000 RMB	214	10.70%
	[20000,40000] RMB	403	20.15%
	[40001,60000] RMB	591	29.55%
	[60001,100000] RMB	530	26.50%
	More than 100000 RMB	262	13.10%
Private car ownership (V4)	No.	586	29.30%
	[1,2] vehicles	1297	64.85%
	3 or more vehicles	117	5.85%

Table 3. Statistics of information for each variable under personal attributes

Attribute	Variable Description	Sample Size	Sample Proportion

Age (V5)	[6,14]	54	2.70%
	[15,17]	55	2.75%
	[18,24]	21	1.05%
	[25,34]	534	26.70%
	[35,39]	591	29.55%
	[40,49]	634	31.70%
	[50,59]	101	5.05%
	60 and above	10	0.50%
Occupation (V6)	Students	106	5.30%
	Staff	1101	55.05%
	Teachers	202	10.10%
	Civil Servants	52	2.60%
	Self-employed	117	5.85%
	Medical personnel	23	1.15%
	Farming	15	0.75%
	Others	384	19.20%
Gender (V7)	Male	1136	56.80%
	Female	864	43.20%

Table 4. Information statistics for each variable under the travel attribute

Attribute	Variable Description	Sample Size	Sample Proportion
Nature of residence (V8)	Household registration	1649	82.45%
	Non-family with more than half a year's residence	304	15.20%
	Non-family members who live less than one month and less than half a year	39	1.95%
	Less than one month for business trip, traveling, visiting relatives	8	0.40%
Walking time (V9)	Less than 3 minutes	358	17.90%
	[3,5] minutes	557	27.85%
	[6,10] minutes	420	21.00%
	[11,15] minutes	182	9.10%
	More than 15 minutes	483	24.15%
Frequency of traveling (V10)	Rarely ride (no ride or [1,2] times)	1077	53.85%
	[3,5] times a week	242	12.10%
	[6,10] times a week	580	29.00%
	More than 10 times a week	101	5.05%
Number of trips yesterday (V11)	[1,2] times a week	1027	51.35%
	[3,4] times a week	680	34.00%
	[5,6] times	159	7.95%
	7 times or more	134	6.70%

In order to ensure the effectiveness and scientificity of the survey, the questionnaire survey was conducted offline in October 2021 in the downtown area of Yingkou City using stratified sampling method, and the scope of the survey included two administrative districts of Xicheng District and Zhanqian District. By analyzing the age and occupation division of the questionnaire survey on the public transportation travel sharing rate in related cities, the 2020 census of Yingkou City, the 2019 edition of the Report on the Analysis of the Data on the Number of Population by Age in Yingkou City, the Statistical Report on China's Talent Resources 2015 [15], and the post planning in the Yingkou

City Public Transportation Planning Research Report (2015 edition), the questionnaire survey was determined of age groups and occupations to achieve comprehensive coverage of all age groups, occupations and places of residence. Considering that the nature of residence may have a certain impact on the choice of travel mode, this paper adds the option of the nature of residence when designing the questionnaire, so as to more accurately analyze the attitudes of people with different nature of residence towards public transport behavior. In order to more carefully understand the public transport behavior of different groups, so as to provide decision-making support for the development of urban public transport, this paper on the "residents travel whether to choose the reasons for public transport", "public transport 'fast up' recommendations " and "the development of urban public transportation should focus on solving the problem" conducted a questionnaire survey.

III. Results and Discussions

A. Latent Class

Based on the survey data of Yingkou City, ten indicators such as total number of households, total household income, individual gender, age, walking time, etc. were selected as exogenous variables from three aspects: household attributes, individual attributes, and travel attributes, and the collected valid data were analyzed using Mplus software to estimate the parameters of the latent category model, and since the number of latent categories is unknown, the model was established by trying to establish the number of latent categories from 1 to 5, the corresponding model's goodness-of-fit indicator values of AIC, BIC, ABIC, and Entropy are shown in Table 5.

Comparing the values of each goodness-of-fit index, it can be seen from Table 5 that the indexes of goodness-of-fit are optimal when dividing the data set into three potential categories, with the largest heterogeneity between groups, significant P-value, and entropy value greater than 0.8, which indicates that the model is acceptable. Therefore, this paper divides the investigated traveling population into three potential categories, and the number of categories and category probability of each potential category are shown in Table 6. Then the great likelihood method is used for parameter estimation to get the conditional probability of each potential category under family attributes, personal attributes, and traveling attributes at different levels of the exogenous variables, and the results are shown in Tables 7-1, 7-2, and 7-3. According to the results in Table 6, it can be seen that there are 136 people in the first category group, accounting for 6.8% of the total number of surveyed people; 856 people in the second category group, accounting for 42.8% of the total number of surveyed people; and 1008 people in the third category group, accounting for 50.4% of the total number of surveyed people.

Analyzing the main characteristics of the three potential category groups, according to the data in Tables 7-1, 7-2 and 7-3, the characteristics of the three groups are as follows: Potential Category I is mainly a group of students aged 6 to 17, the family size of this group is 3 to 4 people and owns 1 to 2 private cars, the frequency of weekly public transportation is higher than that of the other two categories, and the time to walk to the bus stop is within five minutes; Category II group is a group of male workers aged 25 to 34, which rarely travel by public transit and have a walk time to the bus stop of more than 15 minutes; Category III group is a group of workers and other working women aged 35 to 49, which has an annual household income of $40,000 to $60,000, and has a walk time to the bus stop of three to five minutes. Other age groups continue to make up a small percentage of residents in each category.

Table 5. Goodness-of-fit of different categories of potential category models

classes	AIC	BIC	ABIC	LMRT	BLRT	Entropy
1	21943.608	22125.195	22007.681	— —	— —	— —
2	21376.707	21744.789	21506.585	0.0000	0.0000	0.707
3	20995.090	21549.666	21190.772	0.0000	0.0000	0.824
4	20904.045	21645.116	21165.532	0.3438	0.0000	0.831
5	20841.122	21768.687	21168.413	0.8201	0.0000	0.858

Table 6. Category probabilities for each potential category

Latent Classes	Number of categories	Probability of each category
1	136	0.068
2	856	0.428
3	1008	0.504

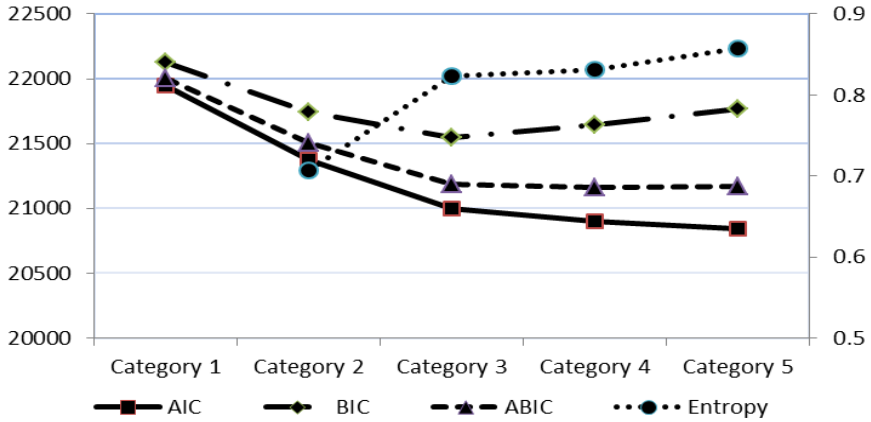

Figure 1. Results of potential category model evaluation indicators for different numbers of categories

Table 7-1 Conditional Probabilities under Family Attributes

Exogenous variables	Levels	Class1 (6.8%)	Class2 (42.8%)	Class3 (50.4%)
Total number of households (V2)	1	0.148	0.464	0.229
	2	0.606	0.419	0.676
	3	0.143	0.086	0.065
	4	0.104	0.031	0.031
Annual household income (V3)	1	0.178	0.135	0.092
	2	0.201	0.203	0.196
	3	0.207	0.380	0.241
	4	0.237	0.21	0.32
	5	0.177	0.072	0.152
Private car ownership (V4)	1	0.237	0.419	0.182
	2	0.709	0.493	0.788
	3	0.054	0.088	0.03

Table 7-2 Conditional Probabilities under Individual Attributes

Exogenous variables	Levels	Class1 (6.8%)	Class2 (42.8%)	Class3 (50.4%)
Age (V5)	1	0.44	0.008	0
	2	0.503	0	0
	3	0.028	0.028	0.006
	4	0	0.408	0.183
	5	0.015	0.243	0.392
	6	0	0.241	0.343
	7	0	0.064	0.072
	8	0.015	0.008	0.003
Occupation (V6)	1	0.988	0.003	0
	2	0.012	0.905	0.328
	3	0	0	0.189
	4	0	0	0.045
	5	0	0.009	0.118
	6	0	0	0.020
	7	0	0.016	0.002
	8	0	0.068	0.298
Gender	1	0.606	0.936	0.281
	2	0.394	0.064	0.719

(V7)				

Table 7-3 Conditional probabilities under trip attributes

Exogenous variables	Levels	Class1（6.8%）	Class2（42.8%）	Class3（50.4%）
Nature of residence (V8)	1	0.926	0.734	0.882
	2	0.059	0.227	0.107
	3	0	0.032	0.011
	4	0.015	0.007	0
Walking time (V9)	1	0.355	0.089	0.251
	2	0.430	0.137	0.342
	3	0.089	0.217	0.203
	4	0.084	0.123	0.087
	5	0.043	0.434	0.117
Frequency of traveling (V10)	1	0.030	0.822	0.759
	2	0.132	0.109	0.135
	3	0.676	0.042	0.044
	4	0.162	0.027	0.062
Number of trips yesterday (V11)	1	0.266	0.657	0.444
	2	0.555	0.160	0.442
	3	0.059	0.111	0.071
	4	0.12	0.073	0.044

B. Characterization of potential class groups

As for the main reasons for different potential categories of groups to choose public transportation, it can be seen from Figure 2 that the factors influencing the first category of groups to choose public transportation are ranked as follows: convenient > low cost > safe > suitable for short distances > comfortable > conducive to transferring > enjoying preferential treatment; factors influencing the second category of groups to choose public transportation are ranked as follows: low cost > convenient > suitable for short distances > safe > conducive to transferring > enjoying preferential treatment > comfortable; and factors influencing the third category of groups to choose public transportation are ranked as follows: low cost > convenient > suitable for short distances > safe > conducive to transferring > enjoying preferential treatment > comfortable. The order of factors influencing the third group to choose public transportation is: low cost > convenient > suitable for short distances > safe > conducive to transfer > comfortable > enjoy preferential treatment. It can be seen that "convenience", "low cost", "suitable for short-distance travel" and "safety" significantly influence the choice of public transportation. passengers' choice of public transportation. However, compared to the second and third groups of passengers, the first group of passengers prefers "comfort" because the first group of passengers contains the largest number of passengers over 60 years old. The second and third categories of riders preferred transit to be "transfer friendly" compared to the first category of riders because the second and third categories of riders are more likely to be in the [25,49] age range, and they preferred to reduce the time spent on peak hour trips.

Regarding the reasons for different potential groups to choose travel modes other than public transportation, Figure 3 shows that the factors affecting the first group to choose travel modes other than public transportation are in the following order: long waiting time > inconvenient transfer > overcrowding > unfree > unreasonable route planning > unreasonable station setups > insecurity; factors affecting the second group to choose travel modes other than public transportation are: long waiting time > overcrowding > transfer > inconvenient > unfree > station setups > insecurity. Inconvenient > Not free > Unreasonable station setting > Unreasonable route planning > Unsafe. Factors affecting the third group's choice of travel modes other than public transportation: long waiting time > overcrowding > inconvenient transfer > unreasonable route planning > unreasonable station setting > unfree > unsafe. Long waiting time, inconvenient transfer and overcrowding are the main factors influencing passengers not to choose public transportation. Compared to the factors of "unreasonable route planning" and "unreasonable station setting", the "lack of freedom" of taking public transportation will influence the first and second groups to choose other modes of transportation. Compared to the factor of "lack of freedom", the factors of "unreasonable route planning" and "unreasonable stops" significantly influence the third group to choose other travel modes than public transportation.

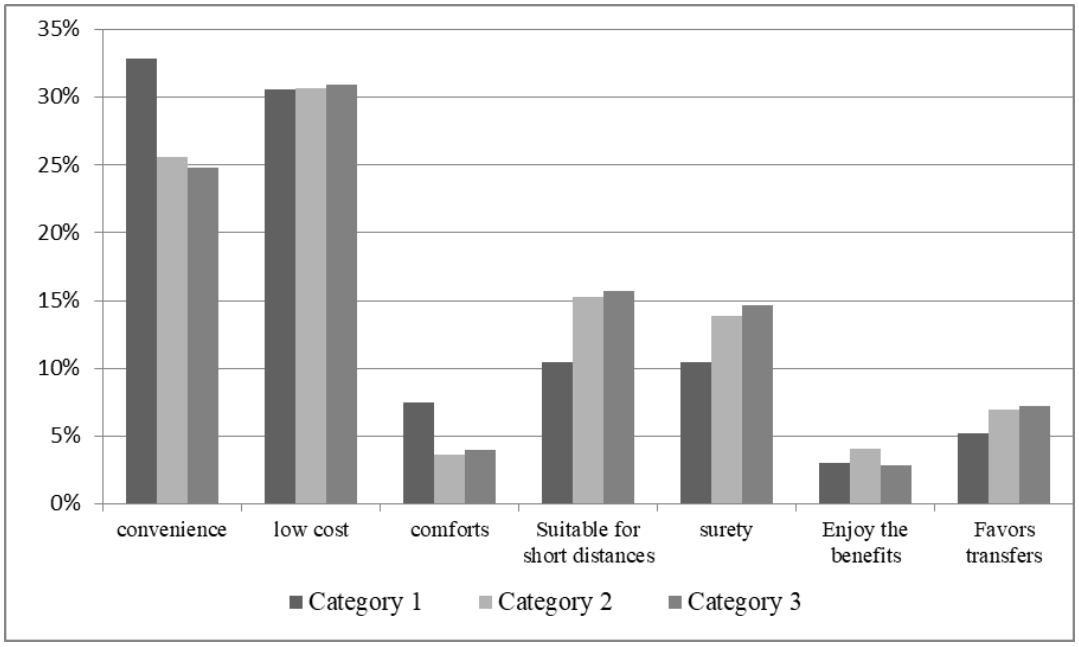

Figure 2. Main reasons for choosing public transportation for different groups

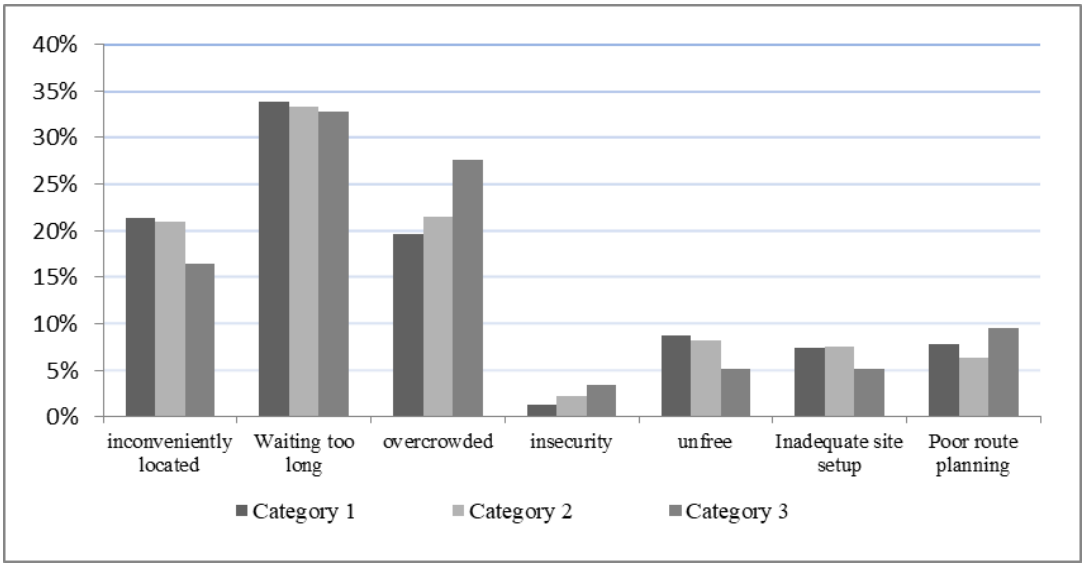

Figure 3. Reasons why different groups choose modes of travel other than public transit

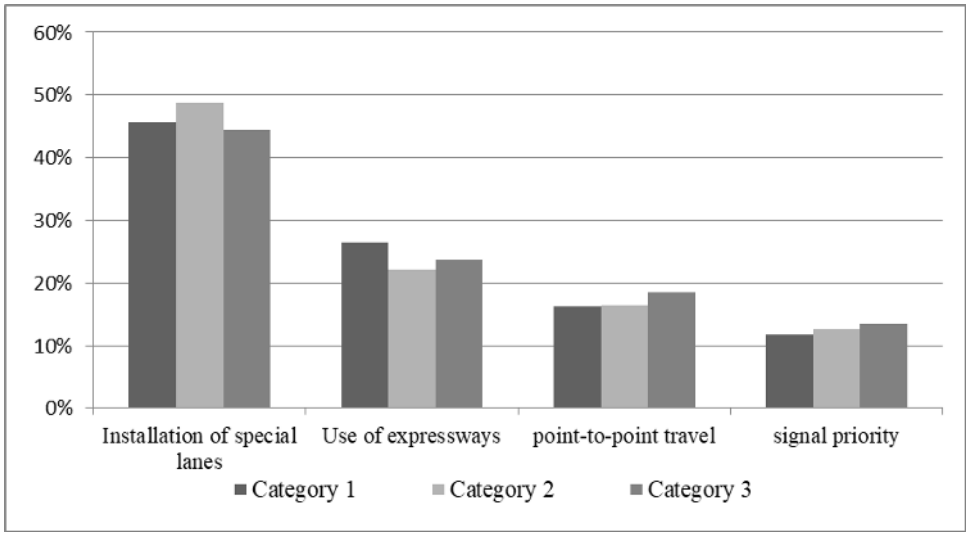

Figure 4. Suggestions from various groups to "fast-track" transit development

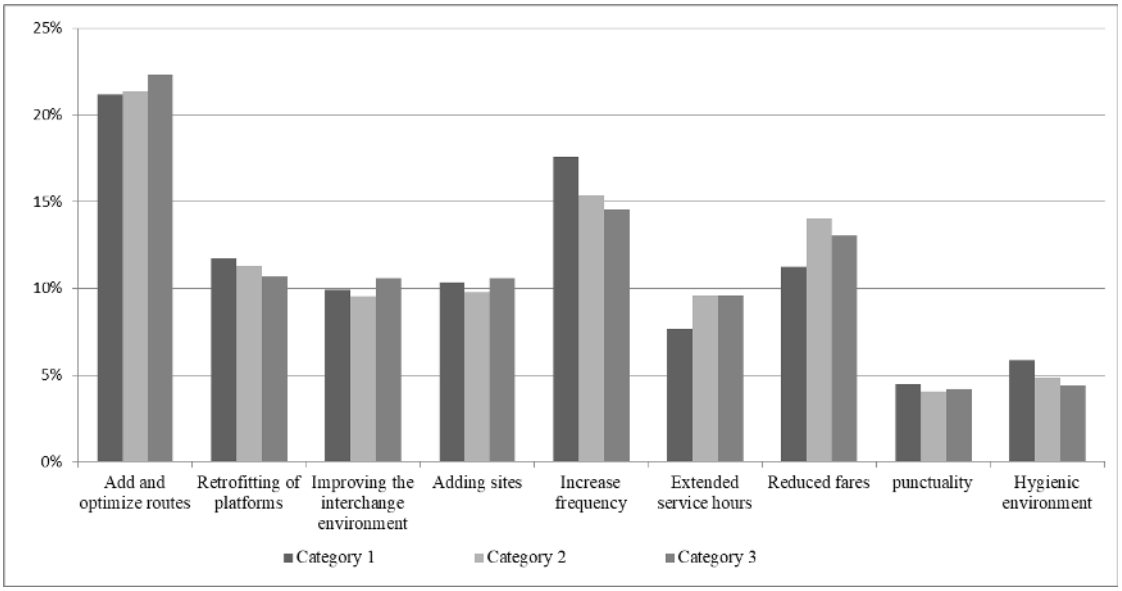

Figure 5. Issues that groups believe should be the focus of transit development

As can be seen from Figure 4, "setting up dedicated lanes" is the most important factor that should be considered, followed by more bus routes "utilizing expressway" systems, then "point-to-point" routes with fewer stops, and then more "signal priority" at intersections. "The second is that bus routes should make more use of the expressway system, followed by "point-to-point routes" with fewer stops, and then more "signal priority" at intersections.

As shown in Figure 5, the first group believes that the development of public transportation should focus on solving the following problems: increase and optimize routes > increase frequency > renovate platforms > reduce fares > increase stops > improve the transfer environment > extend service hours > environmental hygiene inside the vehicle > on-time departure; the second group believes that the development of public transportation should focus on solving the following problems in order of degree: increase and optimize routes > increase frequency > renovate platforms > reduce fares > increase stops > improve transfer environment > extend service hours > environmental hygiene inside the vehicle > on-time departure The second group thinks that the development of public transportation should focus on solving the problems, in order of their degree: increase and optimize routes > increase frequency > reduce fares > renovate platforms > increase stations > improve the transfer environment > extend service time > environmental health inside the bus > on-time bus departures; the third group thinks that the development of public transportation should focus on solving the problems, and the degree of their degree, in order of their degree: increase and optimize routes > increase frequency > reduce fares > renovate platforms > increase stations = improve the transfer environment > extend service time > environmental health inside the bus > on-time bus departures. service hours > environmental hygiene inside the bus > on-time train departures. "Increase and optimize routes" to reduce detours and "increase frequency" are the issues that all three groups think bus development should focus on. Second, the first group prioritizes "platform improvement" to improve waiting comfort over "fare reduction"; the second group considers "fare reduction" more important than "platform improvement" to improve waiting comfort. The second group believes that "reducing fares" is more important than "renovating platforms" to improve waiting comfort; excluding the importance of "reducing fares" and "renovating platforms", the third group believes that "improving the interchange environment" and "improving the waiting environment" are more important than "reducing fares" and "renovating platforms". In addition to "reducing fares" and "remodeling stations", the third group considers "improving the transfer environment" and "adding more stations" to be equally important, followed by "extending service hours".

Convenience" and "low cost" of public transportation are the main reasons for the three groups to choose public transportation; "longer waiting time" is the main reason for the three groups to choose travel modes other than public transportation; "increase and optimize the service time" is the main reason for the three groups to choose public transportation; "increase and optimize the service time" is the main reason for the three groups to choose public transportation. Longer waiting time" is the main reason for the three groups to choose other modes of travel than public transport; "increase and optimize routes", "increase frequency" and "reduce fares" are the issues that the three groups think should be given the highest priority in the development of public transport. In addition, for the first group, "inconvenient transfer" is the main reason for them to choose other travel modes, and this group thinks that the development of public transportation should focus on solving the problem of "reconstructing the platforms", which is because this group mainly consists of the students who have high frequency of weekly public transportation, and they pay more attention to the transportation. For the second group, "inconvenient transfer" and "overcrowding" are the main reasons that prompted them to choose other modes of travel, and this group believes that the development of public transportation should focus on solving the problem of "extending service hours". Extending the service time", because the walking time to the bus stops is longer for this group, which is inconvenient for them; for the third group,

"overcrowding" is the main reason that prompts them to choose other modes of transportation, and this group thinks that the development of public transportation should focus on solving the problem of "improving the transferring time". This group believes that the development of public transportation should focus on "improving the transfer environment" and "increasing the number of stops", because this group is predominantly female and pays more attention to the comfort of riding the bus.

C. Recommendations for urban transit development based on classification results

(1) Reasonable setting of bus routes and stops

Convenience", "low cost", "suitable for short-distance travel" and "safety" are the main factors influencing the three groups to choose public transportation; the second group, accounting for 42.8% of the total sample, and the third group, accounting for 50.4% of the sample, prefer that public transportation "facilitates transfer". The second group, which accounts for 42.8% of the total sample, and the third group, which accounts for 50.4% of the sample, hope that public transportation is "conducive to transfer"; "unreasonable route planning" and "unreasonable station setting" are factors that significantly affect the third group's choice of public transportation. The factors such as "poorly planned routes" and "poorly set up stops" make it obvious that the third group chooses other modes of transportation than public transportation. Therefore, the rational arrangement of bus lines, optimize the station settings, optimize and adjust the bus network and station settings to meet the residents of the bus travel experience and service quality demand can prompt travelers to choose bus travel.

For the reasonable setting of bus lines and stops, it is recommended to consider and implement from the following aspects. Firstly, to understand the residents' travel demand, population distribution, employment and commercial center location and major traffic congestion areas and other information; secondly, based on the survey data and traffic studies, to choose the appropriate route layout, taking into account the road conditions, population density and major travel destinations, so that the bus routes can cover as much as possible of the population and important locations, and at the same time, it should be ensured that the transfers between the different routes are convenient and quick; then Regularly evaluate the operation of bus routes and stops, and make adjustments and optimizations based on residents' feedback and changes in traffic demand.

(2) Adjustment of operating hours and implementation of "just-in-time public transportation".

Excessive waiting time" is the main reason for the three groups to choose a mode of transportation other than public transportation, "increase frequency" is the common view of the three groups that public transportation development should focus on solving the problem, "extend the service time" is also the second and third group of residents to choose a mode of transportation other than public transportation. "Increase frequency" is the common view of the three groups that the development of public transportation should focus on solving the problem, "extend service time" is also the second and third groups that the development of public transportation should focus on solving the problem, which can be improved by rationally adjusting the operating hours of public transportation and the implementation of "just-in-time public transportation". For the adjustment of bus operating hours, extend the service time of the demand, can be considered and implemented from the following aspects. First of all, investigate the residents' demand, through questionnaires and other forms of understanding of the residents' demand for bus operating hours, according to the demand for rational planning of bus routes and operating hours; and then according to the results of the survey and operating costs, to determine the adjustment of the operating time period. In order to balance the residents' demand and the operation of bus companies, considering the cost and benefit of adjusting the bus service time, when implementing the above measures, it is recommended to adopt a differentiated operation mode for different time periods within the operation time period, so as to balance the residents' demand and the operation cost, and at the same time, to establish an effective feedback mechanism to collect the residents' opinions and suggestions on the bus service, and to make timely improvements and Optimization. At the same time, an effective feedback mechanism should be established to collect residents' opinions and suggestions on bus services, so that timely improvements and optimization can be made. It should be ensured that the adjusted operating hours can meet the needs of the residents without bringing too much economic pressure to the bus companies. The implementation of "just-in-time buses" is recommended to be considered and implemented from the following aspects. First, optimize road traffic, improve road traffic conditions is the premise of the implementation of just-in-time bus; Second, improve the level of bus service, to ensure that the normal operation of buses and buses on time; Third, the introduction of intelligent management system, real-time monitoring of the location of buses and driving conditions, timely detection of problems, and scheduling and processing; Fourth, to strengthen the operation and management, public transport enterprises should establish a scientific operation management system, to develop standards and requirements for just-in-time buses. Fourth, strengthening operation management, bus companies should establish a scientific operation management system and formulate standards and requirements for on-time bus departures.

(3) Improving the environment for traveling and transferring to other modes of transportation

The "lack of freedom" will influence the first and second groups to choose other travel modes, and the "renovation of platforms" to improve the comfort of waiting for buses and the "improvement of on-board sanitation" are the issues that the first group thinks should be urgently addressed in the development of public transportation. The first group thinks that the development of public transportation should be solved urgently, and the third group thinks that "improving the environment of public transportation" is a key issue to be solved in the development of public transportation. The following aspects can be considered and implemented to improve the bus environment. First, installing free WiFi and USB charging ports in buses, especially long-distance buses, to ensure that passengers can charge and maintain

communication during the journey; second, providing comfortable seats and air-conditioning systems, upgrading the seats in the buses to provide more comfortable seats and air-conditioning systems, and keeping the interior of the buses hygienic and tidy to improve the riding experience of the passengers; third, providing special facilities, such as child seats and barrier-free lanes, to ensure the safety of the buses and the environment of the disabled. and barrier-free access to ensure their travel safety and convenience; and fourthly, installing cameras and other safety equipment to safeguard passengers' travel safety. For the issue of improving the interchange environment, the following aspects can be considered and implemented. First, providing real-time ride information at bus stops and on vehicles to make it easier for passengers to know information such as bus arrival times and destinations, and to reduce waiting time; second, improving interchange facilities, and installing convenient facilities at interchange stations, such as barrier-free pathways, shelters, and guiding signs, to improve the smoothness and comfort of the interchange. By improving the ride and transfer environment, keeping the vehicle and station neat and clean, improving ride comfort, increasing passenger satisfaction, and enhancing the attractiveness of public transportation travel to the first and third groups of people.

(4) Reduce the cost of public transit travel

Compared with other modes of travel, the main reason for the three groups to choose public transportation is the "low cost" of public transportation, and more than 11% of the residents in the three groups believe that "reducing fares" is a key issue to be resolved. Consider reducing the cost of public transportation to attract residents to choose public transportation. It is suggested to solve the residents' demand for traveling cost from the following aspects. First, the time period discount, in a specific period of time, the implementation of bus transfer discount, such as the morning peak and the evening peak period; Second, cross-regional discount, across different administrative regions of the bus transfer, give passengers a certain discount; Third, the joint ticket discount, for the need to transfer passengers, to provide discounts for the joint ticket; Fourth, the long-term ride discount, for the long term public transportation passengers, can be given certain discounts, such as the purchase of monthly passes, seasonal passes, and other discounts. such as purchasing monthly, quarterly or annual passes.

IV. CONCLUSIONS

Considering the heterogeneity of the population choosing public transportation modes, this paper is based on the survey data of Yingkou City in October 2021, and introduces the potential category model to classify the travel group and analyze the public transportation behavior of different groups. According to the study, the main conclusions obtained in this paper are as follows.

(1) Based on the consideration of individual heterogeneity, by constructing the potential category model, 10 indicators such as the total number of households, total household income, individual gender, age, walking time, etc. are selected as exogenous variables, and applying the Mplus software, through the probabilistic parametrization of the exogenous variables, the public transit travelers are classified into three potential categories based on the indexes of the goodness of fit, which account for 6.8%, 42.8%, of the total number of people surveyed, respectively, 50.4%, which eliminates the heterogeneity of the data to a certain extent, and enables a more detailed analysis of the behavior of urban residents' public transport mode choice.

(2) "Low cost" and "long waiting time" are the main reasons for the three groups to choose public transportation. Regarding the suggestions for the development of urban public transportation, all three groups think that "setting up special lanes" is the most important factor to be considered in route planning; "increasing and optimizing routes", "reducing detours", and "increasing frequency" are the most important factors that should be addressed. "Increase and optimize routes", "Reduce detours" and "Increase frequency" are the most important issues that should be addressed. By analyzing the differences among the three groups, it can be seen that the first group pays more attention to the travel experience that public transportation itself can provide, the second group focuses on the impact that public transportation travel brings to the whole travel activity, and the third group pays more attention to the overall level of service of public transportation.

(3) Based on the analysis of the travel characteristics of each group, the commonalities and differences of the three groups, from the reasonable setting of bus routes and stops, adjusting operating hours and the implementation of "just-in-time buses", improving the environment for riding and transferring, and reducing the cost of public transport travel, four aspects of the development of urban public transport put forward a reasonable and operable proposal.

There are still some shortcomings in the study, this paper only analyzes the characteristics of different groups of public transportation behavior, and will be followed by integrating personal, social and environmental information from multiple sources to understand the reasons for the formation of different potential categories in a more in-depth manner.

REFERENCES

[1] Böckenholt U, Dillon W R. Some New Methods for an Old Problem: Modeling Preference Changes and Competitive Market Structures in Pretest Market Data. *Journal of Marketing Research,* 1997, 34(1):130-142..

[2] Chen J B. A study on heterogeneity of commuters' transportation mode choice based on potential category model [D]. *Nanjing University, China,* 2017-05-24.

[3] TAO W C, SU H Y, PENG S T, et al. Passenger Market Segmentation of Inter-city High-speed Railway Based on Latent Class Model. *Journal of Railway Science and Engineering*, 2021,18 (6): 1410-1417.

[4] LI D L, JI X F. A Model for Segmentation of Highway Passenger Groups Based on Latent Class . *Journal of Highway and Transportation Research and Development*, 2019,36(10):152-158.

[5] . ZHOU Y B, HUO Y M, ZHANG Y Y. Market Segmentation of Urban Rail Transit Commuter Based on Latent Class Analysis . *CHINA TRANSPORTATION OUTLOOK,* 2021,43(9):9-16.

[6] LIU J R, HAO X N, SHI W H. Impact of COVID-19 on the Elderly's Bus Travel Behavior. *Journal of Transportation Systems Engineering and Information Technology*, 2020,20(6):71-76.

[7] XIAN Y J C. Study on Commuter Activity Travel Behavior Analysis. *Shanghai: Shanghai Jiao Tong University*,2019.

[8] TIAN L. A Study on the Influence the Policy of Concessionary Bus have on the Travel Behavior of the Elderly. *Nanjing: Southeast University ,China*, 2017.

[9] HE L H, LI J, SUN J P. How to Promote Sustainable Travel Behavior in the Post COVID-19 Period: A Perspective from Customized Bus Services. *International Journal of Transportation Science and Technology,* 2023, 12 (1) 19–33.
[10] WANG L, CHEN Y, YU W, et al. Identification and Classification of Bus and Subway Passenger Travel Patterns in Beijing Using Transit Smart Card Data. *Journal of Advanced Transportation*, 2023, 2023(10):1-15.
[11] LI P, WU W , PEI X. A Separate Modelling Approach for Short-term Bus Passenger Flow Prediction Based on Behavioural Patterns: A Hybrid Decision Tree Method. *Physica A*, 2023, 616(2023):1-25.
[12] ZHANG H, WANG G H, ZHENG W F. Research on Customer Segmentation of Mobile Payment Based on MaxDiff and Latent Class Analysis. *Journal of Applied Statistics and Management,* 2017,36(03):506-517.
[13] Spearman C. General Intelligence Objectively Determined and Measured. *Am J Psychol.* 2019, 194(15):201-293.
[14] JIAO P P, LI R J, WANG J Y, et al. Causes Analysis on Severity of Elderly Pedestrian Crashes Considering Latent Classes . *Journal of Transportation Systems Engineering and Information Technology.* 2022, 22(5):328-336.
[15] Qiu H Z. Principles and techniques of latent category modeling. *Beijing:Education Science Press, China,* 2008.4.
[16] SUN Y P, SUN R. Build a Talent Classification Statistical System Based on Occupation Classification. *Chinese Personnel Science.*2022,10(58):21-34.

Caixia Chen：(1998-), Female, studying for master's degree, main research direction: urban transportation planning.
Xiaokun Wang: (1975-), Male, PhD, Professor, Main Research Direction: Transportation Planning, Intelligent Transportation System, Logistics System Planning.
Chen Zhang: (1996-), Male, PhD Candidate, Main Research Direction: Transportation Planning.