Intelligent Automation And Soft Computing, 2020
Vol. 26, no. 3, 531–538
DOI: 10.32604/iasc.2020.013930

Classifications of Stations in Urban Rail Transit based on the Two-step Cluster

Wei Li[1, 2, 3], Min Zhou[1,*], Hairong Dong[1]

[1]State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China;
[2]Beijing Transportation Information Center, Beijing 100073, China;
[3]Beijing Key Laboratory of Comprehensive Transportation Operations and Service, Beijing 100073, China.

ABSTRACT
Different types of stations have different functional roles in the urban rail transit network. Firstly, based on the characteristics of the urban rail transit network structure, the time series features and passenger flow features of the station smart card data are extracted. Secondly, we use the principal component analysis method to select the suitable clustering variables. Finally, we propose a station classification model based on the two-step cluster method. The effectiveness of the proposed method is verified in the Beijing subway. The results show that the proposed model can successfully identify the types of urban rail transit stations, clarify the function and orientation of each station.

KEY WORDS: Two-step cluster; urban rail transit; station classification; time series; principal component analysis; spatial-temporal data analysis

1 INTRODUCTION

AS the key node in the urban rail transit network, rail transit stations are affected by the urban area function, the land use function, the traffic network structure and passenger flow state. Different types of stations have been formed, which have different functions in the urban rail transit network. Research on the classification of three urban rail transit stations will be very important for urban traffic managers and rail transit operators to analyze the characteristics of the citizen's travel and evaluate the construction of the rail transit infrastructure. The research will also facilitate the development of passenger flow forecasting and passenger flow control for different types of stations, which shows important reference significance.(Kim, & Kyoungok, 2018; Deng, J. , & Xu, M, 2015; Jun, M, 2015; Liu, S, 2018; Erik, V. T. C. , & Rodriguez, D. A, 2018; Chen, F, 2018; Li, L, 2018).

Scholars have listed two categories in the research of the classification of urban rail transit stations. One is to classify the station based on the characteristics of the connected node, which uses the station passenger flow, the type of station connection mode, the number of track lines, the main functions of the station service area, and the type of station land development as classification indicators (Fang, Z., et. al, 2011; Fu, Bo-feng, et. al, 2008). Another is to classify the station based on the operational data of the station, which uses the intention of the passengers and the station passenger flow as classification indicators. (Yanyan, C, 2017; Lijie, Y. U., 2014).

With the development of Big Data Technology and related research, the research on the urban rail transit stations classification has gradually developed from the previous research through fieldwork data to the large-scale urban traffic big data, (Tang, Q. M. , 2012; Fang, L., 2015; Yao Enjian, 2018; Wang Zi-jia, 2018; Xiang-Nan, L, 2015) Roth (2011) et al. carries out the research on identification of the multi-center urban structure of London through the analysis of smart card data. CaoRui (2016) et al. identifies the type of land use around the stations in Shenzhen using the traffic card data and the time distribution characteristics at each station. Xu Zhirong (2016) uses Shanghai smart card data to analyze the characteristics of occupation-occupancy relationship and the citizen's commuting and identifies the type of station and commuting characteristics. Yin Qin et al. (2016) uses Beijing smart card data to extract the passenger flow characteristics to identify the types of stations and obtains passenger flow and spatial distribution patterns of 8 different types of stations.

CONTACT Min Zhou zhoumin@bjtu.edu.cn

In the above listed researches, the rail transit station smart card data shows massive, high-dimensional, information redundancy and other characteristics, thus the key research direction for the classification of the rail transit stations is extracting the passenger flow characteristics effectively and improving the effective utilization of the original data. In addition, the extracted feature values have different influences on the rail transit station classification results. It is necessary to deeply analyze the extracted feature values and classify them with regards to importance to ensure the rail transit station classification accurately. Aiming at the above problems, this paper applied a two-step cluster method to establish a classification model of urban rail transit stations based on the time series passenger flow characteristics and network structure features. Experimental verification analysis was carried out by the Beijing rail transit network data.

2 RAIL TRANSIT STATION CLUSTERING ANALYSIS METHOD

2.1 Clustering Indicator Analysis and Extraction

In this paper, the feature values such as network structure, time series and the passenger flow model of stations were calculated, and specific quantitative analysis was carried out through the network structure and real smart card data. The clustering indicator categories are shown in Figure 1.

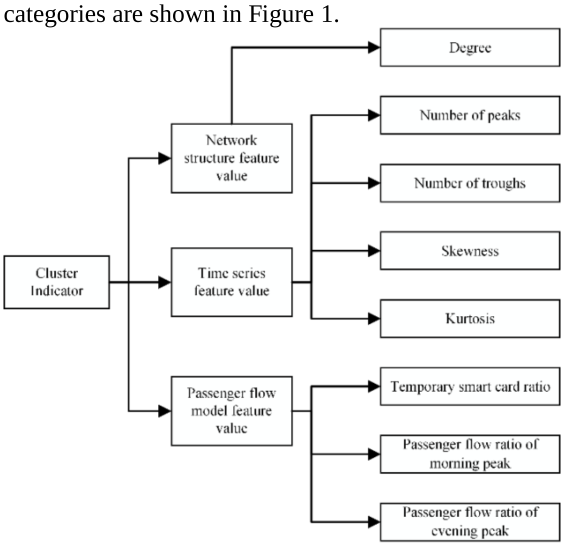

Figure 1. Cluster Indicator Items.

2.1.1 The Network Structure Feature Value Extraction

According to the network characteristics of the rail transit stations, the Complex Networks theory was applied in this paper and the degree of the rail transit stations was extracted, which is the basis of the cluster analysis. The degree of the rail transit stations represents the number of stations connected to the station i.

Cluster indicator 1: The degree in which is the basic concept of characterizing a station in the complex network and can reflect the position and role of the station in shown in the network. In a complex network, degree K_i indicates that station i has the adjacent number of stations. In general, the value of the degree is larger, the influence of the station in the network is greater. The calculation formula is shown in (1).

$$K_i = \sum_{j \in N} \mathrm{a}_{ij} \tag{1}$$

In which (1), the station network is represented by an adjacency matrix $A = (a_{ij})_{N\times \mathrm{N}}$. If station i is connected to station j , then $\mathrm{a}_{ij} = 1$; On the contrary, then $\mathrm{a}_{ij} = 0$.

2.1.2 The Time Series Feature Value Extraction

The morphological characteristics and structural features of the passenger flow time series of the rail transit stations are taken as classification feature values. The morphological characteristics, mainly describe the characteristics of the time series morphological changes, including eigenvalues such as rise, fall, extremum, and transition of time series. The structural features, mainly describe the characteristics of the intrinsic variation mechanism of the time series, including basic statistical characteristics such as mean, variance, skewness, and kurtosis, time domain characteristics such as trend, seasonal wave, and spectral density.

Based on the morphological and structural characteristics of the time series and the specific passenger flow characteristics, this paper extracts the number of peaks, troughs, skewness and kurtosis of the time series of station as the feature values of the time series. These feature values are used as the basis for the cluster analysis.

Cluster indicator 2: The number of peaks, which can describe the morphological characteristics of a time series. According to the time series chart of the inbound and outbound stations of a station, the number of peaks of the station is obtained.

Cluster indicator 3: The number of troughs, which can describe the morphological characteristics of a time series. According to the time series chart of the inbound and outbound stations of a station, the number of troughs of the station is obtained.

Cluster indicator 4: The skewness, which can describe the structural characteristics of the time series and describe the symmetry of the overall distribution of the passenger flow time series. When the skewness is equal to 0, the time series distribution is the normal distribution. When the skewness is greater than 0, the time series distribution is right-biased. When the skewness is less than 0, then the time series distribution is left-biased. The calculation formula is shown in (2).

$$\alpha=\frac{1}{\text{n-1}}\sum_{\text{i}=1}^{n}(x_i-\mu)^3/\sigma^3 \quad (2)$$

In which (2), x_i represents the time series, μ is the sample mean, σ represents the standard deviation.

Cluster indicator 5: The kurtosis, which describes the structural characteristics of the time series and describes the steepness of the overall value distribution pattern of the passenger flow time series of a station. When the kurtosis is equal to 0, the time series is the same as the steepness of the normal distribution. When the kurtosis is greater than 0, the time series distribution is steep and is a peak. When the kurtosis is less than 0, the time series distribution is relatively flat, which is a flat peak. The calculation formula is shown in (3).

$$\beta=\frac{1}{\text{n-1}}\sum_{\text{i}=1}^{n}(x_i-\mu)^4/\sigma^4-3 \quad (3)$$

In which (3), x_i represents the time series, μ is the sample mean, and σ represents the standard deviation.

2.1.3 The Passenger Flow Model Feature Value Extraction

According to the statistics of the daily passenger flow, it is generally considered that the time of the morning peak and evening peak are different between working days, weekends and holidays.

Cluster indicator 6: The temporary smart card ratio, which describes the proportion of the passenger traffic at the rail transit station using temporary cards to enter the station during the statistical time.

Cluster indicator 7: The passenger flow ratio of the morning peak. The calculation formula is shown in (4).

$$p=\frac{Q_e}{Q_d} \quad (4)$$

In which (4), Q_e indicates the average number passing through the station at the morning peak (units per hour); Q_d indicates the number passing through the station throughout the day.

Cluster indicator 8: The passenger flow ratio of the evening peak. The calculation formula is shown in (5).

$$\text{q}=\frac{Q_l}{Q_d} \quad (5)$$

In which (5), Q_l indicates the average number passing through the station at night (in units of people per hour); Q_d indicates the number passing through the station throughout the day.

By extracting the above eight eigenvalues and performing the cluster analysis, we can find similar rail transit stations with similar shapes and intrinsic change mechanisms, and then classify and analyze the orbital stations.

2.2 The Passenger Flow Feature Extraction Based on the PCA

Before the cluster analysis, because each feature value reflects the partial information of the rail transit station in different aspects, it is easy to have the problem of the multivariate collinearity. It is necessary to reduce the dimension of the features value and extract the hidden common factors.

The Principal Component Analysis (PCA) is a multivariate statistical analysis method that uses the idea of dimensionality reduction to condense multiple indicator variables into a few principal components by linear transformation. Each transformed principal component indicator is independent of each other and is a linear combination of the original indicators. Principal component analysis is essentially a matrix transformation. The principal component indicators are obtained through the matrix transformation, and the principal component indicators are not required to have practical significance.

2.3 The Two-step Cluster Model

The two-step cluster model is an exploratory clustering method that uses statistics as distance indicators for clustering and can intelligently determine the optimal number of categories based on statistical criteria. Compared with other clustering algorithms, the results of the two-step cluster analysis are more accurate. The two-step cluster model includes two steps of pre-clustering and formal clustering.

2.3.1 Pre-clustering

This paper uses the following steps to build and modify the clustering feature tree.

Step 1: Construct a leaf node containing all the variable information of the observation value initiated by the root of the tree according to the first observation in the dataset.

Step 2: Each observation is added to an existing node or a new node according to the principle of similarity (the size of the distance measure).

Step 3: When all the observations are placed in the same way and formed into a clustering feature tree, the pre-clustering process ends.

The distance measurement model uses logarithmic similarity. The calculation formulas are shown in (6)-(8).

$$d(i,j)=\xi_i+\xi_j-\xi_{<i,j>} \quad (6)$$

$$\xi_v = -N_v(\sum_{k=1}^{K^A}\frac{1}{2}\log(\hat{\sigma}_k^2+\hat{\sigma}_{vk}^2)+\sum_{k=1}^{K^B}\hat{E}_{vk}) \tag{7}$$

$$\hat{E}_{vk} = -\sum_{i=1}^{L_k}\frac{N_{vkl}}{N_v}\log\frac{N_{vkl}}{N_v} \tag{8}$$

In which (6), $d(i,j)$ represents the distance between the two clusters of i and j ; the $<i,j>$ index represents a new cluster generated by combining the clusters i and j .

In which (7), K^A represents the number of continuous variables; K^B represents the number of categorical variables; N_v represents the total number of records in the v cluster; $\hat{\sigma}_k^2$ represents the estimated variance of the k-th continuous variable in all the data sets; $\hat{\sigma}_{vk}^2$ represents the estimated variance of the k-th continuous variable in the v cluster.

In which (8), N_v represents the total number of recorded data in the v cluster; L_k represents the number of k-th categorical variables; N_{vkl} represents the number of v clusters, and the categorical variable k is divided into l groups.

2.3.2 Formal Clustering

The leaf nodes are combined using a merge clustering algorithm. According to the AIC (Akaike Information Criterion) criterion, each clustering scheme is compared. The smaller the AIC index, the better the clustering effect. The algorithm automatically selects the number of clusters with the smallest AIC index to optimize the clustering scheme. The AIC calculation formula is shown in (9).

$$AIC = -2\sum_{j=1}^{j}(\xi_j + 2J(2K^A + \sum_{k=1}^{K^B}(L_k - 1))) \tag{9}$$

In which (9), J represents the number of clusters; K^A represents the number of continuous variables used by the cluster; K^B represents the number of categorical variables used by the cluster; and L_k represents the number of the k-th categorical variable.

3 BEIJING RAIL TRANSIT SITE CLUSTERING EXAMPLE

3.1 Data Description

THE passenger flow data used in this paper is the 15-minute entry and exit statistics of the Automatic Rail Fare Collection System (AFC). The sampling time is from May 1, 2015 to May 14, 2015 (including the holiday; May 1st and 3rd). The passenger flow data is for 2 weeks, the data sampling interval is 15 minutes, and the sampling time is 06:00 to 23:00 every day. The main fields of the AFC show 15-minute inbound and outbound statistical records in Table 1. After data cleaning, there are a total of 266 rail transit stations in Beijing (the airport line is not counted).

Table 1. **Field Description.**

Field	Description
Start Time	Start time of the passenger
End Time	End time of the passenger
Station Location	Rail station name
Total Tickets Sold	Total number of tickets sold at the station
Total Entry Count	Total number of inbound passengers
Total Exit Count	Total number of outbound passengers

The daily passenger flow time series of each station of the rail transit can be composed of six types of data sets; workday inbound, workday outbound, weekend inbound, weekend outbound, holiday inbound, and holiday outbound. This paper analyzes the time series at the rail transit station level, which gets the characteristics of the morning and evening peak passenger flow distribution, working days, holidays and weekend passenger flow differences, and the ratio of one ticket pass.

3.2 Data Pre-processing

In order to fully reveal the data characteristics, it is necessary to standardize the data. In this paper, the z-score standardization method is used to standardize the observations.

According to the conversion formula, the time series $X=\{x_1,x_2,\cdots,x_m\}$ of length m is normalized to the time series $X'=\{x_1',x_2',\cdots,x_m'\}$.

$$x_i' = \frac{x_i-\mu_x}{\sigma_x} \tag{10}$$

In which (10), x_i represents an observation in the time series X ; x_i' represents an observation value in the time series X' ; μ_x represents the average of all observations in time series X ; and σ_x represents the standard deviation of all observations of the time series X . When x_i' is greater than 0, the observation is above average. When x_i' is less than 0, the observation is below average.

3.3 The Extraction and Processing Analysis Principal Component Index

In this paper, 40 initial feature values extracted according to the above method are selected for the cluster analysis of the rail transit stations, and the PCA method is used for dimensionality reduction to extract principal component indicators. Specific steps are as follows:

Step 1: The KMO and the Bartlett test were carried out to examine the structural validity. The results are shown in Table 2. The KMO test coefficient is 0.833, the significance is 0.000, and the structural validity is high. The PCA method can be used for data processing.

Table 2. The Test of the KMO and Bartlett.

KMO Test	Bartlett Test		
	Approximate chi-square	df	Sig
0.833	14196.098	780	0.000

Step 2: Forty initial feature values were subjected to the principal component analysis, and the feature values of each principal component and their contribution rates were calculated. The results are shown in Table 3. These 12 principal component indicators are extracted from the 40 initial feature values reflecting the characteristics of the rail transit station. These 12 indicators explain that 85.933% of the information in the original information, indicate that the original information has been well extracted. The number of the extracted principal components generally requires a cumulative contribution rate of >85%, so these 12 principal component indicators can be selected for further analysis.

Table 3. The Index Feature Value and Contribution Rate.

Indicators ID	Feature Value	Contribution Rate /%	Accumulation
1	13.704	34.259	34.259
2	5.041	12.602	46.861
3	3.373	8.431	55.292
4	2.480	6.201	61.493
5	1.864	4.660	66.153
6	1.492	3.730	69.882
7	1.364	3.411	73.293
8	1.245	3.113	76.406
9	1.132	2.830	79.237
10	.955	2.389	81.625
11	.901	2.253	83.878
12	.822	2.056	85.933

3.4 The Clustering Result Category Feature Analysis

3.4.1 The Clustering Result

In this paper, the two-step cluster is used to analyze the rail transit stations described by 12 principal component factors, and the appropriate number of station classifications is 6 categories, and the specific classification details of 266 rail transit stations are obtained. At the same time, type naming is carried out according to the characteristics of each type of the rail transit station. Details are shown in Table 4.

Table 4. Station Clustering Results.

ID	Type	Number	Proportion	Typical Station name
1	Transportation hub and tourism commercial type	20	7.5%	Beijing Railway Station, Beijing South Railway Station, Beijing West Railway Station, Tiananmen East Station, Tiananmen West Station…
2	Residential type	16	6.0%	Olympic Sports Center Station, Gaomidian South Station, Huangcun West Street Station…
3	Office type	51	19.2%	Chegongzhuang Station, Dongdaqiao Station, Donghuqu Station, Fuchengmen Station, Guozhan Station, Heping Xiqiao Station, Zhichunli Station, Zhongguancun Station…
4	Mixed living and office type	62	23.3%	Datun Road East Station, Gongxi West Bridge Station, Guogongzhuang Station…
5	Residential and office mixed type but partial living type	85	32.0%	Baliqiao Station, Beiyuan Station, Caofang Station, Communication University Station, Ciqu Station, Huilongguan Station, Huilongguan East Street Station…
6	Residential and office mixed type but partial office type	32	12.0%	Baishiqiao South Station, Beitucheng Station, Chaoyangmen Station, Chongwenmen Station, Dongdan Station, Wangjing Station, Xuanwumen Station…

3.4.2 The Clustering Result Category Feature Analysis

(1) The transportation hub and tourism commercial type.

The type of station is affected by large number of temporary passenger flows, and the peak of the working day is not obvious. During holidays, the number of passengers passing through the station has increased significantly compared to the working days and weekends. Typical sites are large train stations and tourist attractions.

(2) Residential type

Looking at the type of station, the morning peak of the working day is mainly the inbound passenger flow, the evening peak is mainly the outbound passenger flow, and the passenger flow of the inbound and outbound stations is single-wave type distribution, which has obvious tidal characteristics in the time distribution. Weekends and holidays are similar, with peak passengers below working days. The typical rail transit station is a station away from the city center on each route, reflecting the characteristics of the "sleeping city".

(3) Office type

Looking at the type of station, the morning peak of the working day is mainly the outbound passenger flow, the late peak is mainly the passenger flow of the inbound, the passenger flow of the inbound and outbound stations is a single peak type distribution, and the time distribution also has obvious tidal characteristics but the time distribution is the opposite of the office type. Weekends and holidays are similar, with peak passenger flow far below workdays. A typical station is an area with a lot of jobs in the city center.

(4) Mixed living and office type

The type of station has both residential areas and employment areas. The peak passenger flow is relatively average in the morning and evening, and the passenger flow in and out of the station is bimodal, and the tidal characteristics are not obvious. Weekends and holidays are similar, and the peaks are not obvious in the morning and evening. The typical station is a relatively balanced area for occupation and occupancy, and the comprehensive service facilities around the station are relatively complete.

(5) Residential and office mixed type but a partial living type

This type of station has a double-peak distribution on the working day, but the double peak distribution is not obvious compared with mixed living and office type. The peak passenger flow is not balanced in the morning and evening. Weekends and holidays are similar. There is a misplacement of the typical station for the occupation. There are residential areas and office areas around the station, but the proportion of the residential area is larger than the office area.

(6) Residential and office mixed type but partial office type

This type of station has a double-peak distribution on a working day, but the double peak distribution is not obvious compared with the mixed living and office type. The peak passenger flow is not balanced in the morning and evening. Weekends and holidays are similar. There is a misplacement of the typical station for the occupation. There are residential areas and office areas around the station, but the proportion of the residential area is larger than the office area.

Figures 2-7 shows the passenger flow chart of each type of station.

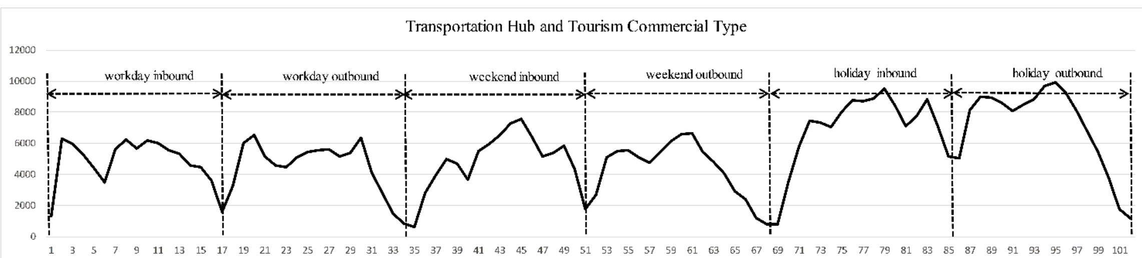

Figure 2. The Station Passenger Flow of the Transportation Hub and Tourism Commercial Type.

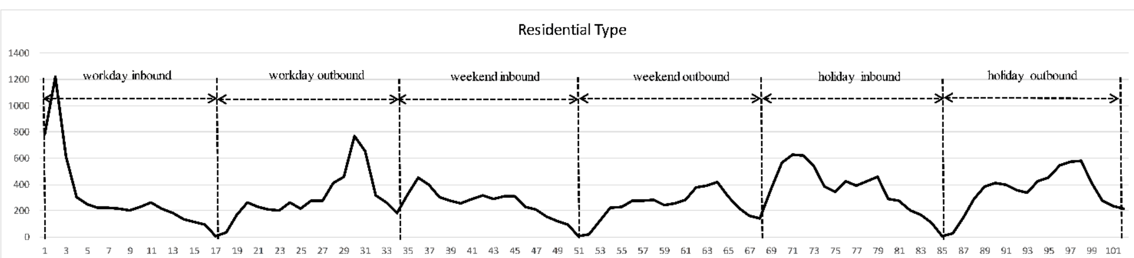

Figure 3. The Station Passenger Flow of the Residential Type.

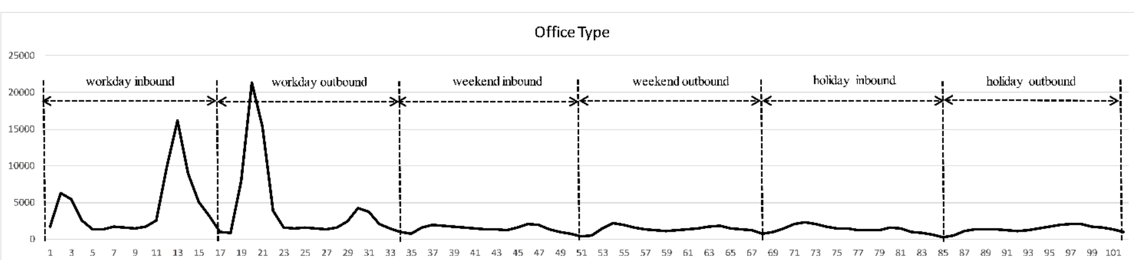

Figure 4. The Station Passenger Flow of the Office Type.

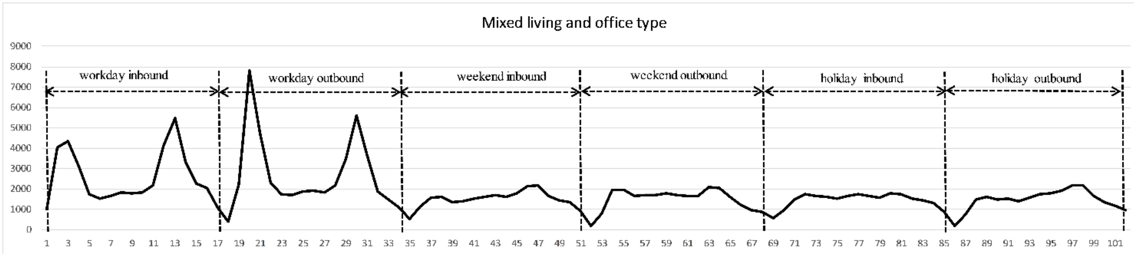

Figure 5. **The Station Passenger Flow of the Mixed Living and the Office Type.**

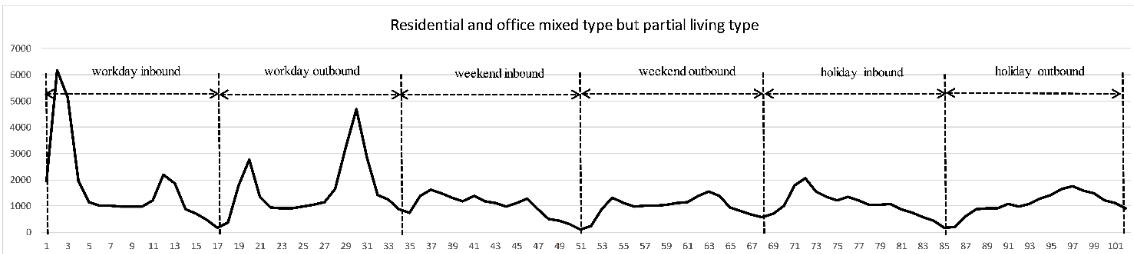

Figure 6. **The Station Passenger Flow of the Residential and Office Mixed Type but Partial Living Type.**

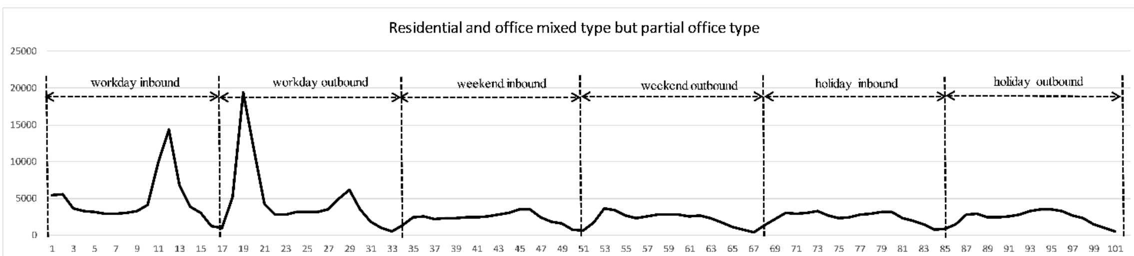

Figure 7. **The Station Passenger Flow of the Residential and Office Mixed Type but Partial Office Type.**

4 CONCLUSION AND DISCUSSION

TAKING Beijing rail transit station as an example, this paper calculated the feature values of the station network structure, time series, and passenger flow model, and got 40 feature values based on the actual urban rail transit smart card data in May 2015. Through the urban rail transit network structure and the real smart card data for specific quantitative analysis, the principal component analysis method was applied to carry out data dimensionality reduction on the feature values, and 266 stations were merged together by the two-step cluster method. The stations were divided into 6 categories, namely the transportation hub and tourism commercial type, residential type, office type, mixed living and office type, residential and office mixed type but partial living type, and residential and office mixed type but partial office type. The results shown that the model proposed in this paper can effectively identify the types of stations. The method is significant for the research of the different types of station passenger flow prediction and can provide useful reference for the urban rail transit management.

5 DISCLOSURE STATEMENT

THIS work is supported jointly by the State Key Laboratory of Rail Traffic Control and Safety (Contract No. RCS2019ZK001 and RCS2020ZZ002) and Postdoctoral Science Foundation of China (2019M660440).

6 REFERENCES

Camille, R., Moon, K. S., Michael, B., & Barthelemy, M. (2011). Structure of urban movements: polycentric activity and entangled hierarchical flows. *PLoS ONE*, 6(1), e15923-.

Chen, F., Qibing, W. U., Zhang, H., Sanbing, L. I., & Zhao, L. (2009). Relationship analysis on station capacity and passenger flow: a case of Beijing subway line 1. *Journal of Transportation Systems Engineering and Information Technology*, 9(2), 93-98.

Deng, J., & Xu, M. (2015). Characteristics of subway station ridership with surrounding land use: A case study in Beijing. *2015 International Conference on Transportation Information and Safety (ICTIS). IEEE.*

Erik, V. T. C., & Rodriguez, D. A. (2018). The ridership performance of the built environment for BRT systems: evidence from Latin America. *Journal of Transport Geography*, S0966692317303526-.

Fang, Z., Tu, W., Li, Q., & Li, Q. (2011). A multi-objective approach to scheduling joint participation with variable space and time preferences and opportunities. *Journal of Transport Geography*, 19(4), 0-634.

Fang, L., Lei, C., Xing, Q. (2015). Classification of urban rail transit station based on traffic transfer. *Journal of Luoyang Institute of Science and Technology (Natural Science Edition).*

Fu, Bo-feng. Wu Jiao-rong. Chen Xiao-hong. (2008) Method of Classification of Suburban Rail Transit Station Sites. *JOURNAL OF THE CHINA RAILWAY SOCIETY*, 2008, 30(6):19-23.

Jun, M., Choi, K., Jeong, J., Kwon, K., & Kim, H. (2015). Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. *Journal of Transport Geography.*

Kim, & Kyoungok. (2018). Exploring the difference between ridership patterns of subway and taxi: case study in Seoul. *Journal of Transport Geography*, 66, 213-223.

Liu, S., Yao, E., & Li, B. (2018). Exploring urban rail transit station-level ridership growth with network expansion. *Transportation Research Part D Transport & Environment.*

Li, L., Wang, Y., Zhong, G., Zhang, J., & Ran, B. (2017). Short-to-medium term passenger flow forecasting for metro stations using a hybrid model. KSCE *Journal of Civil Engineering.*

Lijie, Y. U., Yan, L. I., Kuanmin, C., School, H., & University, C. (2014). Using spectral clustering for urban rail station classification. *Journal of Transport Information and Safety.*

Qin, Y., Bin, M., & Liying, Z. (2016). Classification of subway stations in Beijing based on passenger flow characteristics. *PROGRESS IN GEOGRAPHY*, 35(1).

Rui, C., Wei, T. U., Baichong, C., Nianxue, L., Meng, Z., & Qingquan, L. I. (2016). Identification and analysis of home and work regions in the vicinity of metro stations using smart card data. *Journal of Geomatics.*

Tang, Q. M., Shen, L. P., & Tang, X. Y. (2012). Operation organization at urban rail transit station under large passenger flow. *Applied Mechanics and Materials*, 253-255, 6.

Wang Zi-jia. Liu Hai-xu. Taku Fujiyama. (2018) Metro Station Classification by Service Function Based on AFC Data and RF Method. *Journal of Transportation Systems Engineering and Information Technology*, 2018, 18(4).

Xiang-Nan, L. I. (2015). Classifying urban rail transit stations using cluster analysis. Railway Standard Design.

Xu Zhirong. (2016) Study on Job-Housing Relationship and Characteristic of Commuting in Shanghai: Based on the Perspective of Rail Transit Passenger Flow Data. *Shanghai Urban Planning Review*, 2016(2):114-121.

Yanyan, C., Nuo, Z., & Yao, L. U. (2017). A new classification approach to urban rail stations based on external connections. *Urban Rapid Rail Transit.*

Yao Enjian. Zhou Wenhua. Zhang Yongsheng. (2018) Real-Time Forecast of Entrance and Exit Passenger Flow for Newly Opened Station of Urban Rail Transit at Initial Stage. *China Railway Science*, 2018, 39(2).

7 NOTES ON CONTRIBUTORS

Wei Li is a Ph.D. student at Beijing Jiaotong University, Beijing, China. He received a Master's and bachelor's degree from Beijing Jiaotong University, Beijing, China, in 2009 and 2007 respectively. His main research field includes intelligent transportation systems.

Min Zhou received Ph.D. degree at Beijing Jiaotong University, Beijing, China in 2019. His current position is a Post-doctoral fellow at the Beijing Jiaotong University. His recent research focused on intelligent transportation systems.

Hairong Dong received Ph.D. degree at Peking University, Beijing, China, in 2002. Her current position is a Professor at Beijing Jiaotong University. Her recently research focused on control theory, and intelligent transportation systems.