Rail Transit Travel Time Distribution and Prediction Based on Automatic Fare Collection Data

Ma Kun
Beijing Jiaotong University
School of Traffic and Transportation
Beijing, China
makunsy@sina.com

Wen Jiaxing
Beijing Jiaotong University
School of Traffic and Transportation
Beijing, China
13125732@bjtu.edu.cn

Wang Qi *
Beijing Jiaotong University
School of Traffic and Transportation
Beijing, China
wangqibjtu@163.com
* Corresponding Author

Abstract— With the development of urban rail transit line networking, precisely obtaining the travel time distribution and analyzing its characters have become very important. This paper is based on Automated Fare Card data to obtain the travel time distributions by analyzing the transaction records from Automated Fare Collection system. Beijing Metro is used as a case study. By choosing typical Origins to Destinations with different travel distances and transfer times, travel time distributions are generated and get fitted to the curves. A linear model is used to describe the travel time with different distance and transfer times. With consideration of the actual travel distance, transfer times, the time period and some other factors, a travel time prediction method is proposed. This proposed measure is used to predict the travel time from the time point when passengers swipe their smartcards at the entry gate to the time point when they reach the exit gate. The outcomes of this research are validated by the actual OD data from the Beijing Metro case.

Keywords-Rail Transit; Travel Time; AFC Data; Pedestrian Flow; Prediction Methodology

I. INTRODUCTION

With the development of the national economy, our cities are expanding, along with the population growth, traffic congestion becomes a big problem. Thus, rail transit starts playing a more important role in urban traffic development. Currently, weekday passenger flow is maintained around 800 million or above, and it is continuously growing. Rail transit, in this situation, will become a more important part of Beijing's transportation.

Passenger travel characteristic analysis is the basis of pedestrian flow prediction, operation plan and passenger organization. Yanshuo Sun and Ruihua Xu decomposed journeys into separate stages and got more accurate passenger travel time, then the path choice could be derived [1]. In 2012, they analyzed the travel time distribution and path choice behavior in "Rail Transit Travel Time Reliability and Estimation of Passenger Route Choice Behavior --Analysis Using Automatic Fare Collection (AFC) Data", etc. [2] In 2007, Martin, Catherine and Bruno Agard demonstrated the use of a smart card data to estimate the performance of a variety of transport methods in "Calculation of Transit Performance Measures Using Smartcard Data"[3].

Based on the AFC data research, from a macroscopic perspective, studying the laws of pedestrian flow varies with different period and space help us to have a better understanding of characteristics of the passenger flow in Beijing rail transit. On a microscopic point, the emphasis is focusing on studying rail transit travel time and its influential factors, and establishing travel time prediction models. These studies provide theoretical support to improve the level of service and operational efficiency, optimize network operation capabilities, and gain revenue for operating company and plan travel time for passengers.

II. THE CHARACTERISTICS OF URBAN RAIL TRANSIT TRAVEL TIME DISTRIBUTION

A. The Impact of Different Travel Distance on Travel Time

1) The Analysis of Short Travel Distance

a) The curve of travel time distribution of short travel distance has one peak. With the increase of travel time, the number of passengers increased at first, then decrease.

b) The travel time is within 30 minutes, the differences between the longest and the shortest travel time from every Origin to Destination(OD) are within 15 minutes. The differences are relatively concentrated.

c) The travel time which has the highest frequency nearly equaling to the average travel time.

d) According to the curve, the Normal distribution is selected for describing the travel time distribution.

2) The analysis of long travel distance

a) The curve of travel time distribution of long travel distance has one peak. With the increase of travel time, the number of passengers increased at first, then decreased. There is a higher probability of transfer that passengers choose ODs which has long distance. It leads to more travel time.

b) The travel time is within 60 minutes, the difference between the longest and the shortest travel time from every OD is about15 minutes. The differences are not concentrated.

c) The travel time which has the highest frequency nearly equals to the average travel time.

d) According to the curve, the Normal distribution or Polynomial distribution is selected for describing the travel time distribution.

3) Laws of travel time distribution related to travel distance

According to the discuss above, several laws could be summarized：

a) In general,with increased travel time, travel distance increases. The width of increase is not equal which is mainly affected by the differences of schedule between every line.

b) The curve of travel time distribution of travel distance has only one peak regardless the travel distance.

c) The average travel time nearly equals to the travel time which has the highest frequency.

B. The Impact of Different Transfer Times on Travel Time

In the situation of there are more than one path between one OD pair. If the difference of transfer times between different paths is one, the distance is within three stations. In most cases, when deducting the transfer times, travel distance increased one station. Considering the passenger's psychological requirements of reducing walking distance and waiting time, passengers who travel at a loop line will choose a path which has a small number of transfer times [4].

1) Travel time frequency distributions of non-transfer, one time of transfer, two or more times of transfers are matched to the distributions of different travel distance.

2) The preliminary comparison found that with the increasing number of transfer times, the average travel time is longer than non-transfer case, which means the increase in the number of transfer times leads to an increase in travel time.

C. The Impact of Different Travel Period on Travel Time

Analyzing the travel time frequency distribution from each ODs, several laws could be found:

1) The travel time frequency distribution of morning and evening peak hours on weekdays (from Monday to Friday) are relatively stable, the curve has the same trend which have only one peak; travel time frequency distribution curve also has one peak on weekend.

2) The travel time frequency distribution of non-peak hours on weekdays is less stable than peak hours. On weekend, the number of passengers increased significantly, reducing the gap between weekdays.

Integrating the travel time frequency distribution of different period of one day and weekday or weekend from one week, some laws could be found:

1) Passenger's different choice of different travel period leading to the travel time frequency uniformly distributed.

2) The impact of different travel period to travel time was not significant.

Overall, no matter at the morning peak hour, evening peak hour, or non-peak hour, the travel time are roughly equal between a fixed OD. Only the left behind in a line or a station would lead to a significant increase in travel time.

III. THE BASIC LAWS OF PASSENGER TRAVEL

A. The Relationship of Travel Distance and Travel Time

TABLE 3-1.RELATIONSHIP OF TRAVEL DISTANCE AND TRAVEL TIME

OD	Spacing Distance	Average Travel Time	Transfer Times
Longze-Xierqi	1	7.82	0
	2		0
Jinsong-Jintaixizhao	3	10.56	0
Beigongmen-Zhongguancun	4	11.99	0
Jinsong-Tuanjiehu	5	14.51	0
Tiantongyuan-HuixinxijieBeikou	6	20.83	0
Guomao-Xidan	7	20.42	0
	8		0
Xizhimen-Beijingnan	9	25.26	0
	10		0
Sihui-Fuxingmen	11	29.21	0
Guomao-Gongzhufen	12	28.99	0
Tiantongyuan-Chaoyangmen	13	42.19	1
	14		0
Tiantongyuan-Jintaixizhao	15	44.66	1
Tiantongyuan-Fuchengmen	16	50.03	1
	17		0
Pingguoyuan-Yonganli	18	50.60	0
	19		0
	20		0
Pingguoyuan-Sihui	21	58.60	0

According to the above table, a statistical chart could be made. There are several ODs that include one transfer times, so the Fig.3-1 and Fig.3-2 made respectively as follows for fitting the distribution curve. The horizontal coordinate expressed as the travel distance, unit: station; Longitudinal coordinate expressed as the travel time [5], unit: min.
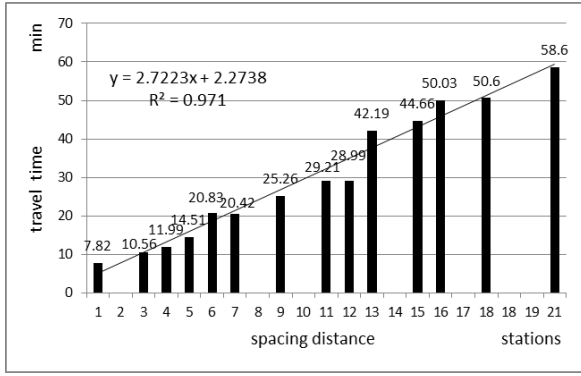
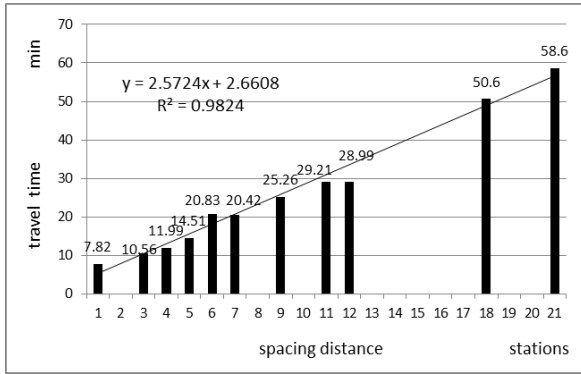Figure 3-1. Impact of Travel Distance on Travel Time (including transfer)

Figure 3-2. Impact of Travel Distance on Travel Time (exclusive of transfer)

Travel time distribution fitting curve including transfer shows that it is linear relationship of travel time and travel distance in Fig.3-1. The fitting formula and the value of R^2 are [6]:

$$Y = 2.7223x + 2.2738 \qquad (1)$$
$$R^2 = 0.971$$

Y expressed as the dependent variable, is travel time, unit: min;

X expressed as the independent variable, is travel distance, unit: station.

Travel time distribution fitting curve exclusive of transfer shows that it is linear relationship of travel time and travel distance. The fitting formula and the value of R^2 are:

$$Y = 2.5724x + 2.6608 \qquad (2)$$
$$R^2 = 0.9824$$

Y expressed as the dependent variable, is travel time, unit: min;

X expressed as the independent variable, is travel distance, unit: station.

R^2 are very close to 1 in the two fitting formula, Taking into account the increase in transfer times will lead to the increase in travel time, so the formula (2) is more reasonable. When other conditions are same, travel time increases linearly with the increase in travel distance, the formula is:

$$t = 2.5724x + 2.6608 \qquad (3)$$

T expressed as travel time only to consider travel distance, unit: min;

X expressed as travel distance, unit: station.

B. The Relationship of Transfer Times and Travel Time

The impact of travel distance on travel time including one time of transfer is shown in Fig.3-3. In this graph, horizontal coordinate expressed as the travel distance, unit: station; Longitudinal coordinate expressed as the travel time, unit: min. Yellow histogram represents the travel time without transfer, Black histogram represents the travel time with one time of transfer [7].
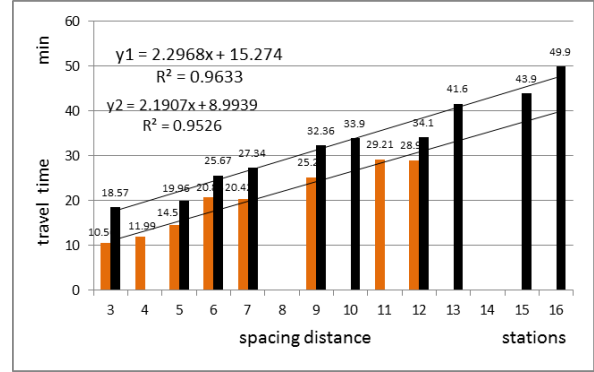

Figure 3-3. Relationship of Travel Distance and Travel Time

Fitting formula for the histogram in two cases:
$$Y1 = 2.2968x + 15.274 \qquad (4)$$
$$R^2 = 0.9633$$
$$Y2 = 2.1907x + 8.9939 \qquad (5)$$
$$R^2 = 0.9526$$

Y1, Y2 expressed as the dependent variable, is travel time, unit: min;

X expressed as the independent variable, is travel distance, unit: station.

Y1 regarded as travel time with transfer, Y2 regarded as travel time without transfer.

As shown in Fig.3-3, the two trend lines are in parallel. It shows that Travel time increases linearly with the increase of travel distance in the same condition.

The difference between the two trend lines at the Y axis is 6.3 min, it can be considered that 6.3min is caused by the transfer.

C. The Relationship of the Period of Travel and Travel Time

Based on the II.C, generally, the different period of one day and the different day of one week make a tiny effect on travel time. In the morning and evening peak hours, the crowd would cause the travel time 1 min longer than the peak-off hours [8]. There is a smaller difference of travel time between weekends and weekdays, which could be neglected.

IV. PASSENGER TRAVEL TIME PREDICTION

Travel time consumption is an important indicator about the characteristics of urban residents' travel [9]. The Prediction of Rail transit passenger travel time is a complex project. Beijing Rail Transit has the feature of intensive network, uneven distribution of population, and varieties of management systems and schedules of each line operation. These lead to the prediction of travel time is pretty tough.

According to historical data, grasping the regularities of travel time, comprehensive considering the analysis of

travel distance, transfer times, the period of travel and other factors are the basis of travel time prediction.

A. The Hypothesis of Travel Time Prediction

By summarizing the influence factors of travel time, a hypothetical formula for travel time prediction is [10]:

$$T=2.5724x + 2.6608+6.3y+z \qquad (6)$$

T expressed as travel time, unit: min;

X expressed as travel distance, unit: station;

Y expressed as the times of transfer, non-transfer was recorded as zero;

Z expressed as the travel time, the morning and evening peak hours was recorded as1, the peak-off hours was recorded as zero.

Note: Some stations are influenced by the crowd and other factors. There is a phenomenon of "left behind" during the peak hours, such as Tiantongyuanbei, Tiantongyuan and Huilongguan. At this point, the above formula does not apply.

B. The Validation of Hypothesis

Select actual data to fit the parameter from several pairs of typical OD for validating the formula. And the psychological factor of passengers should be considered.

Generally, 5min delay can be accepted. So, 5min is used as delay deviation range in this paper. It means formula was acceptable when the difference between test results and the actual data is within 5min that can be used to predict travel time. Validation table 4-1 as follows:

TABLE 4-1. VALIDATION STATISTICAL TABLE.

OD	Spacing Distance	Transfer Times	When to go	Forecast Travel Time	Actual Travel Time	D-Value
Houshayu-Wangjing	6	0	morning peak	19.10	29.8	-10.70
Qianmen-Dongwuyuan	8	1	morning peak	30.54	24.4	6.14
Sihui-Liujiayao	10	1	flat peak	34.68	30.3	4.38
Gongyixiqiao-Xidan	7	0	morning peak	20.67	20.7	-0.03
Dawanglu-Xidan	8	0	evening peak	24.24	21.6	2.64
Wudaokou-Beijingnan	12	1	flat peak	39.80	43.6	-3.80
Xihongmen-Dongwuyuan	15	0	flat peak	41.20	43.9	-2.70
Renmindaxue-Dongzhimen	9	1	flat peak	32.10	35.4	-3.30
Puhuangyu-Dongzhimen	8	1	flat peak	29.54	32.3	-2.76
Majiapu-Xidan	5	0	morning peak	15.50	16.8	-1.30
Liyuan-Dawanglu	11	1	morning peak	38.30	39.5	-1.20
Liujiayao-Chaoyangmen	8	1	morning peak	30.54	27.9	2.64
Beijingnan-Dongzhimen	11	1	flat peak	37.30	40.5	-3.20

The ODs in Table 4-1 are randomly selected from the raw data. All the difference between predicted values and actual values of ODs are within 5min except the difference of Houshayu − Wangjing, which the difference between predicted value and the actual value are 10.7min. So it can be considered that the prediction method is available. At the same time, according to the overall situation, the prediction value is less than the actual value.

V. CONCLUSIONS

By analyzing the Urban Rail Transit passenger travel characteristics, the basic laws of the passenger travel time have been concluded and the relationship between the factors that affect the travel time has been established. Finally, the prediction model of the passenger travel time has been proposed, and then validated it by actual data.

The paper contributes to distribution and prediction of rail transit travel time, and it has been carried out to further research on the rail transit travel time. Also it provides a reference for the estimation of passenger behavior in urban rail transit.

REFERENCES

[1] Yanshuo Sun, Ruihua Xu. Urban Rail Transit Travel Path Calculation Based On Automated Fare Collection Data[J]. Traffic and Transportation,2011,02:85-90.

[2] Yanshuo Sun, Ruihua Xu. Rail Transit Travel Time Reliability and Estimation of Passenger Route Choice Behavior Estimation Using Automatic Fare Collection Data[J]. Journal of the Transportation Research Board, 2012, Vol. 2275, pp. 58-67.

[3] Uniman, D. Service Reliability Measurement using Smart Card Data: Application to the London Underground[D]. Massachusetts Institute of Technology,2009.

[4] David L. Uniman, John Attanucci, Rabi G. Mishalani, Nigel H. M. Wilson. Service Reliability Measurement Using Automated Fare Card Data: Application to the London Underground[J]. Transportation Research Record: Journal of the Transportation Research Board. Dec 2010, Vol. 2143, pp. 92-99.

[5] Baohua Mao, Bingfeng Si, Zhili Liu. Urban Railway Traffic Network Management and TollAllocation:Theory andApplication[M].Beijing: Science Publishing House,2007.

[6] Jang,Wonjae. Travel Time and Transfer Analysis Using Transit Smart Card Data[C]. Transportation Research Board 2010 Annual Meeting CD-ROM, Washington, DC, 2010.

[7] Zhao, Jinhua, Rahbee, Adam, Wilson, N. H. M. Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems[J]. Computer-Aided Civil and Infrastructure Engineering, vol. 22, no. 5, (2007), pp. 376-387.

[8] Seddon P.A., Day M.P. Bus Passenger Waiting Times in Great Manchester[M]. Traffic Engineering and Control, (1974), 15:442-445.

[9] Qin Luo. Theoryand simulationanalysis ofpassenger flow distribution based on network operation for urban masstransit[D].Shanghai:Tongji University,2009.

[10] Bingfeng Si, Baohua Mao, Zhili Liu. Passenger flow assignmentmodel for urban railway traffic network under the condition of seamless exchange[J]. Journal of China Railway Society,2007,29(6):12-1.