Analysis of Transfer Route Choice on Loop Line with Automatic Fare Collection Data

Hongchang Tian[1] and Hezhou Qu[2]

[1]School of Transportation and Logistics, Southwest Jiaotong Univ., Chengdu, Sichuan 611756, China; National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Chengdu, Sichuan 611756, China (corresponding author). E-mail: tianhongchang2013@qq.com
[2]School of Transportation and Logistics, Southwest Jiaotong Univ., Chengdu, Sichuan 611756, China; National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Chengdu, Sichuan 611756, China. E-mail: hzqu@swjtu.edu.cn

ABSTRACT

The loop line plays an important role in meeting the transfer demand and enhancing the accessibility of urban rail transit network. After operation of loop line, the actual effect on transfer flow and passengers' route choice are crucial for metro management. Some factors may greatly affect route choice on loop line, which are helpful to estimate an individual's decision-making and improve transfer service quality. The massive data in automated fare collection (AFC) systems opens up a new opportunity to understand the transfer effect on loop line. In this paper, an analysis framework is firstly developed to define typical transfer sub-network formed by secant lines and loop line. Then, a qualitative analysis is conducted to explore the potential impact of spatio-temporal information by passenger route choice in sub-networks. To analyze transfer route choice, a spatio-temporal route match model by mining AFC data is proposed. The case of a typical loop line in Chengdu Metro is applied to demonstrate the capability of the proposed model. Results shows that passengers' transfer route choice behaviors on loop line could be estimated, which is useful to analyze spatio-temporal distribution of transfer flow, and evaluate the transfer service quality of loop line.

Key words: urban rail transit, loop line, transfer route choice, spatio-temporal, automatic fare collection (AFC) data

1 INTRODUCTION

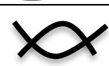
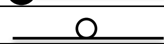
Urban rail transit (or called metro) has been largely applied in many cities of the world. Nowadays, metro networks are expanded rapidly in many metropolises of China to satisfy a huge demand for passengers. By the end of 2018, there were total length of 5,761.4 km metro lines in 35 cities of Mainland China (CAMET, 2019).

Loop line is a kind of urban rail transit lines, which approximately forms a circle. London had the first loop line in 1884, and 41 loop lines or pan-shaped loops comes into service in 32 metropolises worldwide now (Rohde, 2017). It is noted that up to 50% metropolises whose scale of metro network is over 100 km have built loop lines or proposed at least one loop line. The length of existing loop lines is about 20 km en average, but ranges from 3 km (People mover, Miami, United States) to 57.1 km (Line 10, Beijing, China). At least 5 loop lines are now proposed or under construction in China.

In fact, not all loop lines have the same shape and most of them approximate to rectangle or precisely, inscribed polygon of a circle. Therefore, it is more meaningful to distinct loop lines from its features of function. The relatively large number of loop lines built or under construction, proves that loop lines are a feasible option to enhance the usability of existing metro networks, especially in terms of transfer convenience and relieving congestion. The chessboard and radial metro network without a loop line often cause several big transfer hubs in the central area of metropolis. Big transfer hubs have been proven overcrowded and inefficient in western cities. The loop line orbiting the city core or the central business district (CBD), could diverge big transfer flow before entering the city center and other lines could meet in smaller transfer stations within the circle.

Table 1 shows different types of loop lines. Some lines (i.e. terminal loop and spiral) mainly satisfy the demand in some special geographical areas. Some lines (i.e. pan-shaped loop and operationally split circle) look like a circle but operate only in one direction. The isolated loop line is compared to be more reliable. An isolated circle crossed by secant lines and diameter lines is considered the best efficient type of metro networks, since it provides network connections between sub-centers or satellites and gives more transfer choices into or out of city center. Such connections assist passengers avoiding overcrowded at big transfer hubs, reducing travel time and even reducing transfer times.

Table 1. Terminology and function of loop lines

Type	Diagram	Transfer Convenience	Quality of Relieving Congestion
Isolated circle		High	High
Shared circle		High	Middle
Pan-shaped loop		Middle	Middle
Operationally split circle		Middle	Middle
Terminal loop		Low	Low
Spiral		Low	Low

The objective of this study is to develop an effective method to analyze the route choice behavior on typical transfer sub-networks formed by an isolated loop line and secant lines. Considering the spatial (i.e. relative locations of OD pairs in the sub-network) and temporal (i.e. timetable, walking time and transfer time) constraints, a route match model is formulated considering spatio-temporal information obtained by passengers. The case of a typical loop line in Chengdu Metro is applied to test the proposed model and the transfer route choice behaviors on loop line is analyzed.

Table 2. Number of transfer stations in the loop line network

Number of lines	Network type	
	Circle-radial	Circle-chessboard
3(1 circle 2 lines)	5	5
4(1 circle 3 lines)	9	8
5(1 circle 4 lines)	14	11-12
6(1 circle 5 lines)	20	14-16
Formula	$N^{**} = \frac{(n^{*}-1)(n+2)}{2}$	$N = a^{\#}b^{\sim} + 2(a+b)$

*number of lines in the network **number of total transfer stations
number of lateral lines in the chessboard network
~ number of vertical lines in the chessboard network

2 ANALYSIS FRAMEWORK: TYPICAL TRANSFER SUB-NETWORKS

The network made by secant lines could be divided into 2 main types: chessboard network and radial network. The isolated loop line intersects with the existing network and forms several transfer stations, as shown in Table 2 (Cheng, 2014), which provides more alternatives for passenger to transfer among different lines.

4 typical sub-networks (Figure 1&2, using a circle to represent loop line, 2 points to represent OD pairs and two straight lines to represent secant lines) could be generalized from total metro networks to analyze passengers' route choice behavior. The angle between secant lines and the relative position of the origin and destination points (OD pairs) may influence the route choice behavior in these sub-networks.

(1) Parallel Transfer Sub-network: parallel lines sub-network and acute triangle sub-network (shown in Figure 1)

If several chessboard lines are partly put into operation, passengers may prefer choosing the circle route to cross the central area (Song, 2016). Acute triangle transfer network is composed of an arc of loop line and two edges of radial lines, between which the angle is inferior to right angle. Researchers have found that as the angle is inferior to 43° (Cheng, 2014); circle route is preferred over radial route from one edge to another.

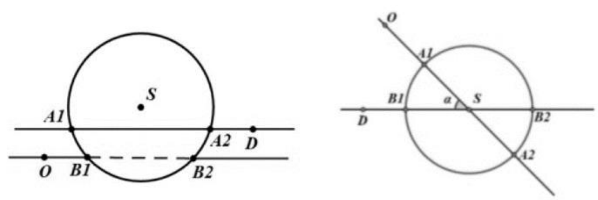


Figure 1. The topological form of parallel transfer sub-network

(2) Vertical Transfer Sub-network: right-angle sub-network and obtuse triangle sub-network (shown in Figure 2)

Theoretically, as radium is given, the arc will be longer than the sum of the two radii if the angle made by the arc and the center of the circle is larger than 115°. The smaller the angle is, the higher the probability of circle route chosen by passengers to cross the central area is. Some loop lines whose radium is longer than 6km (Cheng, 2014) would take more transfer flow between suburbs and main city areas.

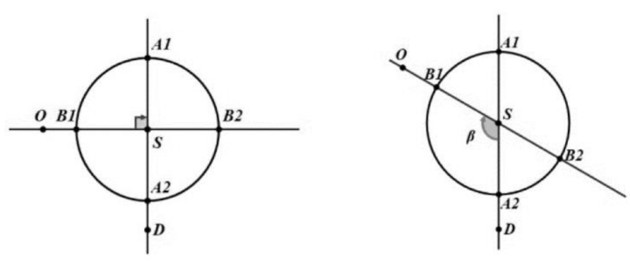


Figure 2. The topological form of vertical transfer sub-network

In reality, passengers would choose the "optimal" route in these sub-networks according to the spatio-temporal information factors consideration. A qualitative analysis will show the effect of factors on route choice.

3 QUALITATIVE ANALYSIS: SPATIO-TEMPORAL INFORMATION FACTORS

The spatio-temporal information factors influencing passengers' route choice in sub-networks could be sorted into 3 aspects: metro map, navigation and quality of service.

(1) Metro map

As mentioned in the second section, the circle route, which looks like shorter than secant lines, might be chosen by passengers. Most passengers could not obtain detailed information like real route distance and the whole timetable, and research result shows that passengers often trust the metro map more than their own travel experience, even for the most experienced passengers in metro system (Guo, 2011). A passenger may often count the number of stations from a transit map as a good approximation for the actual distance and in-vehicle time. This "reality" variable in model is useful to estimate passengers' real route choice behavior.

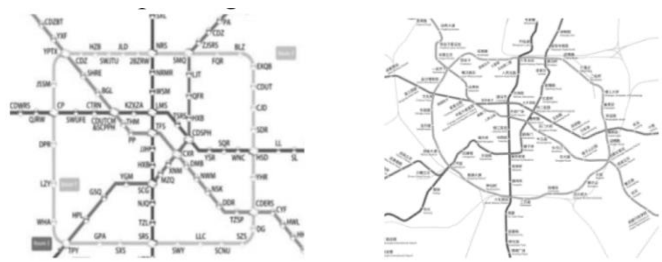


Figure 3. Schematic and geographical map of Chengdu Metro in central area

(2) Navigation

Nowadays, with the development of smartphones, passengers depend on the "optimal" route suggested by navigation application on travel, especially in case that the passenger visit the city for the first time or the metro network is too complex to be familiar with. Each application has their own preference to path planning, such as less travel time or less transfer times. Generally, the route with less travel time has more attraction to passengers.

(3) Quality of service

Quality of service includes transfer convenience, average waiting time, number of trains waited and level of crowding. Transfer convenience rely on the layout of transfer station and transfer flow organization. Waiting and crowding effect are not only determined by passenger flow and timetable, but also vary from person to person. Only frequent passengers might find an optimal route in their views.

4 SPATIO-TEMPORAL ROUTE MATCH MODEL

As discussed earlier, the route match model with the spatial and temporal constraints is presented to find the most possible route for each passenger in transfer sub-networks according to the AFC data.

4.1 Notations and assumptions

The notations of variables used for the proposed model is summarized in Table 3 and the assumptions considered in this study are described as follows.

(1) Trains run following the posted timetable.

(2) During a specific time period, average walking speed is assumed as a constant.

(3) Passengers would not be dwelled at stations except waiting for train.

(4) The route containing more than two transfers to reach their destination is negligible.

Table 3. Variable Definitions

Variable	Definition
t_O / t_D	Time swiping the card at the entrance of origin O or at the exit of destination D
$t^{i,l,u}_{S,dept}$ / $t^{i,l,d}_{S,dept}$	Departure time of the i^{th} train of line l at station S (origin O and transfer station S_t) on up/down direction
$t^{i,l,u}_{S,arrv}$ / $t^{i,l,d}_{S,arrv}$	Arrival time of the i^{th} train of line l at station S on up/down direction
$T^S_{walk,a/e/t}$	Walking time (access, egress and transfer walking time) at station S (origin O, destination D and transfer station S_t)
$\Delta T_{m,S_t}$	Liberal time for the m^{th} route at origin O and transfer station S_t
$\Delta T'_m$	Limited liberal time for the the m^{th} route
$T^l_{S1,S2}$	In-vehicle time on line l from station $S1$ to $S2$
r_m	Matching score for the m^{th} route

4.2 Passenger travel time and limited liberal time

Passenger travel time consists of walking (including access, transfer and egress walking time), waiting and in-vehicle time. In-vehicle time is constant since the route is given and trains follow the timetable strictly. Walking time is affected by the layout of each station and passenger walking speed, but it's a minor component compared to in-vehicle time (Qu, 2018). Minimum walking time at each station could be obtained by investigation. Therefore, the route match is mainly determined by the temporal "degree of freedom", which is the interval between the time point arriving to the platform and departure time point of the boarding train at origin and transfer station.

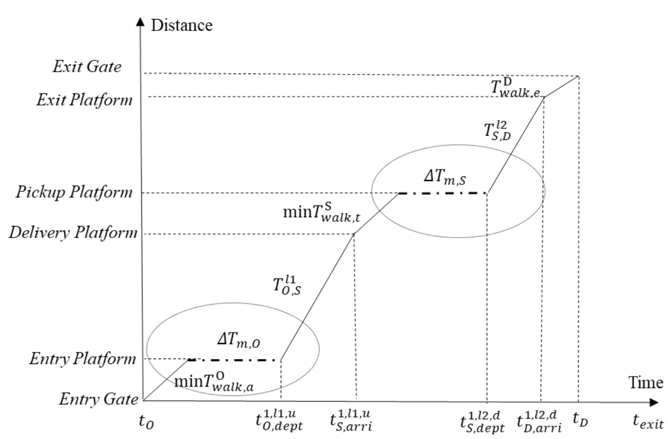


Figure 4. Space-time diagram of passenger travel trajectory with one transfer

This interval is named the "liberal time", formulated as Eq. 1 at origin and Eq. 2 at transfer station for a given route. "Liberal" means that this time slot varies over passengers' route choice and train timetable.

$$\Delta T_{m,O} = t^{1,l1,u/d}_{O,dept} - (t_O + \min T^{O}_{walk,a}) \tag{1}$$

$$\Delta T_{m,S} = t^{1,l2,u/d}_{S,dept} - (t^{1,l1,u/d}_{S,arri} + \min T^{S}_{walk,t}) \tag{2}$$

The space-time diagram for a passenger with 1 transfer and 2 liberal time is illustrated in Figure 4. For different routes of a given OD pair, the direct impact of increase of transfer times and layout of stations could be measured.

If a liberal time is short, it could be estimated that the passenger has strong temporal limit to take the given train and given route. The shortest liberal time called "limited liberal time" (Eq. 3), is the most restrictive part for a spatio-temporal route.

$$\Delta T_m^{'} = \min\{\Delta T_{m,S}\}, S \in \{O, S_t\} \tag{3}$$

The proportion of limited liberal time in the sum of liberal time represents the degree of spatio-temporal match limit to the estimated route. Thus, the "matching score" is defined as Eq. 4. The route with the highest matching score, in other words, with the strongest spatio-temporal restrict and least freedom to vary, has the highest possible to be chosen by the passenger.

$$r_m = \frac{\Delta T_m^{'}}{\sum_S \Delta T_{m,S}}, S \in \{O, S_t\} \tag{4}$$

If the absolute value of the difference between two scores is inferior to 0.001, count the number of stations of each route and choose the route with less stations as the most possible route for passenger.

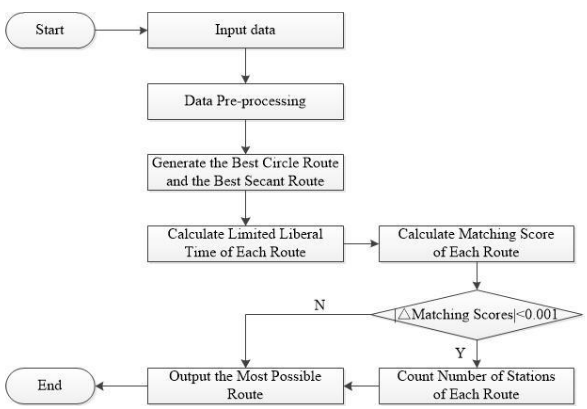


Figure 5. Flowchart of data processing

4.3 Data processing

The following procedure generates the most possible route for a given entry and exit transaction card record in the transfer sub-networks. Figure 5 shows the data processing procedure.

Step 1: Extract entry and exit transaction records distributed in transfer sub-networks from the AFC data. Input the train timetable, minimum walking time at each station, etc.

Step 2: Eliminate the abnormal trip data and the employee card records.

Step 3: For each OD pair, generate the best (least travel time) circle route and the best secant route suggested by map applications. Match them to the corresponding AFC data.

Step 4: Calculate the limited liberal time of best circle route and best secant route for each AFC data by Eq. 3.

Step 5: Calculate the matching score for each route by Eq. 4.

Step 6: Compare 2 matching scores.

Step 7: Output the most possible route (highest score) for corresponding AFC data.

5 CASE STUDY

Chengdu metro system in China consists of 6 lines 137 stations in 2018. Line 7 in Chengdu is a loop line, which is the third isolated circle in China after Line 2 and Line 10 in Beijing Metro. The total length is 38.61 km intersecting at 8 transfer stations.

The input data mainly comes from the AFC data and train timetables. The AFC data (April 2018) consists of Card ID, Category of Card (Single Trip or Tianfutong Metro Card), Entry/Exit Time and Station. It contains nearly 59.3 million transaction records incurred by nearly 8.67 million passengers. Abnormal data (nearly 2%, e.g. one trip cost more than 3h, trips with the same origin and destination) and routes with more than two transfers are excluded. Since each station has at least 2 entrances and 2 exits according to the investigation, the minimum walking time at one station is obtained based on comparison of walking time from each entrance and exit. Minimum transfer walking time is collected between platforms of two different lines.

Figure 6 shows the studied transfer sub-networks and the corresponding OD pairs in Chengdu Metro. The acute triangle sub-network covers a small range from a bus terminal to commercial districts nearby. The two vertical sub-networks connect Chengdu High-Tech Industrial development Zone in the south with the northwest area of Chengdu. Hi-Tech Zone is the biggest employment concentration district where hundreds of thousands of frequent passengers are to and from work. 9 stations in Hi-Tech Zone are in the top 25% of largest entering volume on weekday. The northwest area is a large residential zone and important science-education center, in which Xipu station (in obtuse-triangle) has largest entering volume during AM peak of weekday.

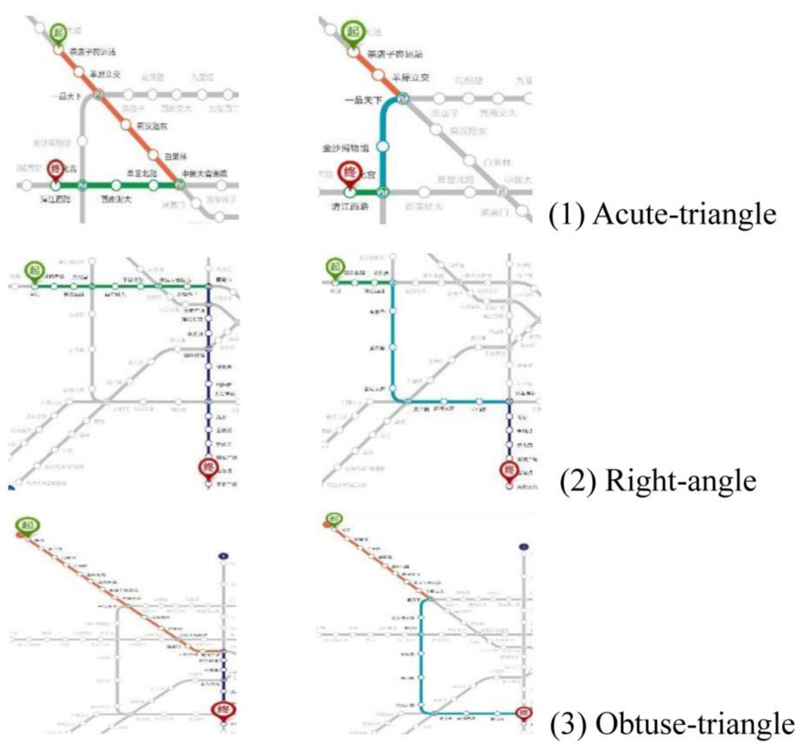


Figure 6. Transfer sub-networks and OD pairs in Chengdu Metro

5.1 Passenger route choice and transfer

Figure 7 illustrates the proportion of passenger route choice in transfer sub-networks. Less than 20% of passengers are believed to choose the circle route. In the obtuse sub-network, passengers more prefer to choose the circle route. This graph shows that angle does not have a major influence on circle route as estimated in the second section. Transfer time may be the most important factor in route choice.

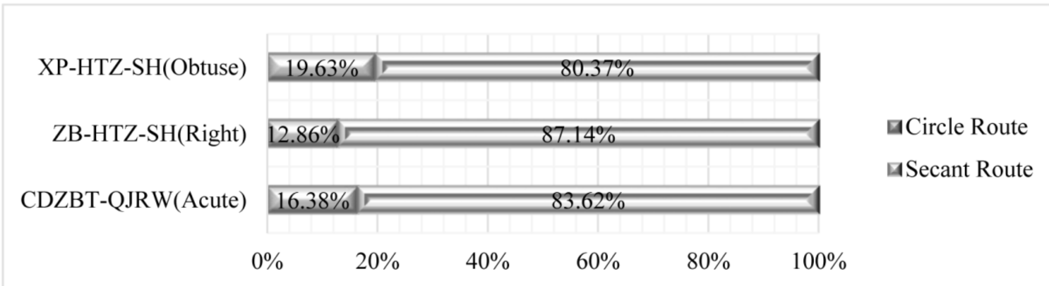
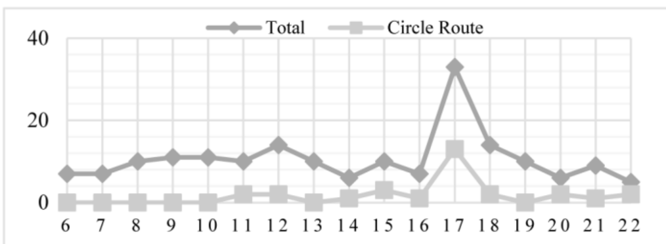
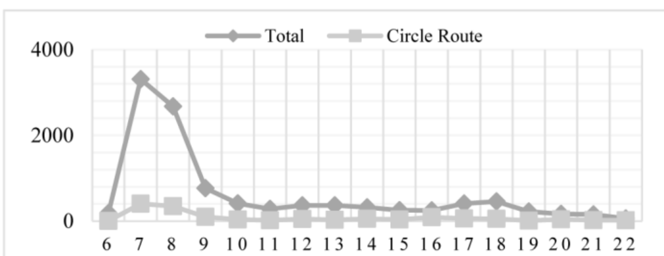


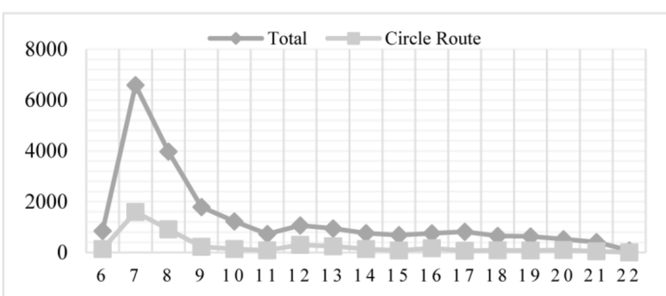
Figure 7. Route choice in transfer sub-networks



(1) Acute-triangle



(2) Right-angle



(3) Obtuse-triangle

Figure 8. Temporal distribution of passenger flow on a typical weekday

5.2 Analysis of passenger flow distribution

(1) Ridership distribution over time

The ridership distribution over time for all passengers and passengers choosing circle route on a typical weekday is shown in Figure 8. In the acute-triangle sub-network, passenger flow is small and PM peak (i.e. 16:30-19:00) appears obvious. In two vertical sub-networks, passenger flows are big and AM peak (i.e. 7:00-10:00) is highly significant. Comparing to the overall passenger flow, the passenger flow of circle route is flat and peak periods are not obvious, which implies that passengers prefer the circle route with high quality of service mainly at off-peak hours.

(2) Ridership distribution over stations in obtuse-triangle sub-network

Figure 9 shows the proportion of passenger choosing the circle route on several typical weekdays between Xipu and 9 stations from the north of High-Tech Zone to the south in the obtuse triangle sub-network. As shown in Figure 9, the attraction to circle route for passengers does not grow along with the extension of travel distance. However, the proportion of passengers choosing circle route from Hi-Tech Zone is nearly double than that to Hi-Tech Zone. Considering Figure 8 (3), Xipu to Hi-Tech direction is more crowded and passengers do not expect to be dwelled at transfer stations during AM peak. The secant route with only one transfer is thus preferred over circle route. As the circle route is estimated to be time-saving and less crowded, passengers (mainly after work) travelling from Hi-Tech to Xipu prefer choosing it to obtain a higher quality of service.

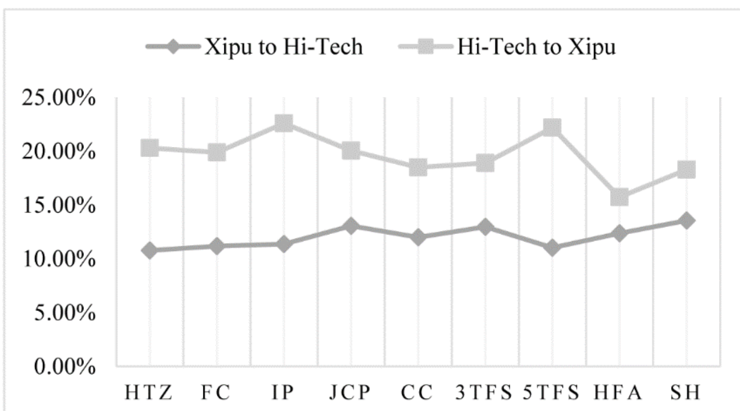


Figure 9. Spatial distribution of circle route in obtuse-triangle sub-network

6 CONCLUSIONS

The loop line has the function on enhancing the connectivity of existing metro networks, especially on improving transfer convenience and relieving congestion in central area. Generally, the loop line and secant line generate two types of transfer sub-networks: parallel and vertical, in which passenger route choice is affected by spatio-temporal information obtained, e.g. metro map and navigation. A spatio-temporal route match model defines temporal constraints for potential routes and simulates the effect of spatial factors i.e. number of stations.

In fact, passengers care the increase of transfer times for circle route, and the unpopularity of circle route in the transfer sub-networks are unavoidable. It is found that the angle of sub-network and travel distance are not related to the preference of circle route. It is also observed that the passenger flow of circle route over time is relatively flat and passengers choosing circle route is double on CBD-residential area direction. All imply that passengers prefer secant route in condition of large passenger volume. During off-peak hours, circle route, which might be time-saving and more convenient, has more attraction to passengers.

7 RECOMMENDATIONS FOR FUTURE RESEARCH

As an extension of this study, the proposed model could be generalized by introducing some methods such as machine learning, in order to simulate the route choice behavior and estimate the passenger transfer flow, especially in case that new secant lines intersect with the loop line.

ACKNOWLEDGEMENT

This research was supported by the Open Research Fund for National Engineering Laboratory of Integrated Transportation Big Data Application Technology (NO. CTBDAT201910). The authors appreciate the data support from Chengdu Metro.

REFERENCES

CAMET (China Associations of Metros). (2019). "Statistical Analysis Report of China Urban Rail Transit Lines in 2018", http://www.camet.org.cn/index.php?m=content&c=index&a=show&catid=18&id=16219 (Apr.27, 2019)

CHENG J. (2014). Study on the passenger routing in loop line of urban rail transit system, Chang'an University, Xi'an, China

GUO Z. (2011). Mind the map! The impact of transit maps on path choice in public transit, *Transportation Research Part A*, 45(2011), 625-639

QU H., X.X., C.I.S. (2019), Analysis of passenger load and wait time for a metro system with automatic fare collection data, *Transportation Research Board*, the 98th Annual Meeting, Washington D.C.

Rohde M., "Metro Rings and Loops", https://mic-ro.com/metro/metrorings.html (Aug.6, 2017)

SONG X. (2016). Discussion on setting causes and transfer effects of Beijing Rail Transit Loop Line, *Science and Technology*, 2016-06, 130-132

ZHAO Y., T.X., C.F., C.S., Z.Q. (2018), Analysis of passenger flow characteristics and operation model for Chengdu rail transit loop line, *Railway Standard Design*, 1004-2954(2018)05-0014-06

ZHOU L. (2014), Verification and Optimization of Network Clearing Model, *Urban Mass Transit*, 2014.11.105