Citation: JIANG Yin, ZHENG Haixing, YU Shiyuan, TANG Xiao. Relationship Between Job-Housing Spatial Distribution and Rail Transit Network in Tianjin: An Analysis Based on Cellular Data [J], Urban Transport of China, 2018 (06): 26–35.

Relationship Between Job-Housing Spatial Distribution and Rail Transit Network in Tianjin: An Analysis Based on Cellular Data

JIANG Yin¹, ZHENG Haixing², YU Shiyuan³, TANG Xiao⁴

1. Tianjin Municipal Engineering Design & Research Institute, Tianjin 300201, China;

2. Tianjin Urban Planning and Design Institute, Tianjin 300201, China;

3. Tianjin Traffic Management Bureau, Tianjin 300201, China;

4. World Bank Financed Project Management Office of Tianjin Urban and Rural Construction Commission, Tianjin 300201,

China

Abstract: In order to better explain the low passenger volume of rail transit network in Tianjin and promote the coordinated development of urban transportation and land use, this paper proposes the techniques for analyzing the spatial distribution of job-housing based on cellular data. Through discussing the relationship between rail transit network and urban spatial structure, the paper analyzes the urban spatial structure in Tianjin in three aspects: commuting circle, Central Business District, and commuting travel linkage. The results show that the coordination between the existing or recently-constructed rail transit network and urban spatial structure in Tianjin is generally not very good, and there is a large room for improvement. The paper emphasizes that a regular assessment should be carried out on the relationship among the job-housing spatial distribution, urban land use development and rail transit network based on cellular data, which can provide scientific basis for the rail transit network development and the adjustment of urban land use development. **DOI:** 10.13813/j.cn11-5141/u.2018.0604-en

Keywords: rail transit; urban spatial structure; cellular data; job-housing distribution; relationship; Tianjin City

0 Introduction

As of the end of 2017, the permanent resident population of Tianjin was 15 568 700, the operating mileage of rail transit about 166 km, the daily ridership about 970 000, and the ridership intensity was 6 000 passengers $d^{-1} \cdot km^{-1}$. Compared with cities with similar population such as Guangzhou and Chongqing, Tianjin has relatively low ridership and ridership intensity on its rail transit network ^[1]. Whether the rail transit network and the urban spatial structure are interconnected is an important factor affecting the rail transit ridership. It is also a key indicator to measure whether the development of urban transportation and land use is going well. Relevant literature studies believe that rail transit plays an important role in the reconstruction and regeneration of metropolitan space, while the spatial scale and structure of metropolitan areas determine the scale and shape of the rail transit network ^[2]. Therefore, the study of the relationship between urban spatial structure and rail transit network in Tianjin is an extremely important topic in the context of large-scale rail transit construction.

As big data are constantly applied and popularized in urban transportation, the study on the relationship between job-housing spatial distribution and urban spatial structure based on cellular data has achieved rich research results. Reference [3] explored the reliability of different sampling methods of cellular data and their impacts on the identification of home and work locations and the analysis of commuting characteristics. Reference [4] used cellular data to analyze the job-housing spatial distribution of users in Beijing. Reference [5] studied the relationship between job-housing conditions and transportation systems in the main urban area of Chongqing. References [6-7] used cellular data to look into the relationship between commuting area and job-housing spatial distribution in Shanghai. Reference [8] analyzed the spatial characteristics of the city cluster of Pearl River Delta based on cellular data. Reference [9] estimated the Origin-Destination (OD) distribution of trips among districts in San Francisco, USA, based on cellular data.

Based on the experience of analysis algorithms of cellular data in the world, this paper obtains job-housing spatial distribution relationship by analyzing the cellular data in

Received: 2018-06-06

1

Supported by: World Bank technical assistance project (the third batch of loan) for Tianjin "Development Strategies of Green Transportation in Central Urban Area of Tianjin (Phase I)" (P148129)

First author: JIANG Yin (1977–), male, from Huaihua, Hunan Province, Master's degree, senior engineer, deputy dean of Integrated Transportation Development Research Institute, is mainly engaged in the research on transportation planning, transportation big data and model, and transportation policy. E-mail: jyin88@126.com

Tianjin, and further proposes the idea of analyzing the urban spatial structure in three aspects: commuting area, Central Business District (CBD) and commuting travel connection. This idea is applied in Tianjin to explore the relationship between the existing and recently-constructed rail transit network and urban spatial structure, which provides a reference for the optimization of rail transit network and the adjustment of urban spatial structure.

1 Application of cellular data

1) Users' job-housing and commuting analysis based on the identification of the residence and workplace

With the advantages of high sampling rate, wide coverage and high accuracy, cellular data provide an important opportunity and data foundation to track each individual's spatial-temporal behaviors. In recent years, it has become a hot topic and has been widely used in the field of urban and transportation planning to establish models to identify the residence and workplace of urban residents based on cellular data to study the characteristics of job-housing spatial distribution.

2) Travel OD analysis based on the identification of trip stops

Urban residents' travel OD and its characteristics constitute the most important basic materials in transportation planning and management. They are also the focus of research and discussion in the field of transportation. Cellular data accumulate massive users' real-time location data, which provides rich resources to study residents' travel behaviors, monitor real-time traffic status and estimate travel demand. In recent years, the research on travel OD analysis based on cellular data has been constantly making progress. However, subject to the positioning accuracy of mobile base stations, the applications have been limited in aspects such as short-distance travel OD and the analysis of related travel characteristics.

2 Identification of the workplace and residence

The analysis of urban spatial structure mainly uses the job-housing distribution obtained based on cellular data.

2.1 Algorithm to identify the workplace and residence

The algorithms to identify the workplace and residence based on cellular data can be divided into three major categories, including the time threshold method, the relative dwell time method and the information entropy method. 1) Time threshold method. The location at which the dwell time over the night (e.g. 0:00-6:00) exceeds a certain threshold for more than a predefined number of times can be identified as the residence. The location at which the dwell time during the daytime (e.g. 9:00-17:00) exceeds a certain threshold for more than a predefined number of times can be identified as the workplace.

2) Relative dwell time method. This method calculates the ratio (*P*) of time staying in each space grid over the night (e.g. 0:00-6:00), and takes the grid with the maximum *P* value as the residence. This method calculates the ratio (*P*) of time staying in each space grid during the daytime (e.g. 9:00-17:00), and takes the grid with the maximum *P* value as the workplace.

3) Information entropy method. In order to distinguish users resting at night and users resting during the day, Reference [10] introduced the concept of information entropy to analyze users' activities during the study period, then effectively identified users' concentrated resting period, which effectively solved the problem of distinguishing users working during the day and users working at night.

The time threshold method and the relative dwell time method are relatively simple, which are commonly used by various operators and research institutions. The information entropy method is relatively complicated with few application cases, but it has effectively solved the problem of mis-identifying the workplace as the residence due to users' irregular working hours or night shift jobs.

A set of effective algorithms is developed in this research to identify the workplace and residence in Tianjin by referring to cellular data algorithms in the world. This set of algorithms is based on the actual situation of cellular data from three operators in Tianjin, i.e., China Unicom, China Mobile and China Telecom. It comprehensively considers the application of the time threshold method, the relative dwell time method and the information entropy method, and adopts the spatial-temporal double-layer clustering analysis. The temporal clustering refers to the merging and processing of consecutive records located in the same grid (300 m \times 300 m). The spatial clustering means when calculating the dwell time of a potential residence location (or workplace), the dwell time of the grids within a certain range around this location should also be included due to the "ping-pong" phenomenon of the cellular data. The details are as follows:

1) Data overview

The cellular data for one consecutive month are obtained from three operators in Tianjin, i.e., China Unicom, China Mobile and China Telecom. The data include user number, grid ID, the longitude and latitude of the grid center, start time, end time and other information. The grid size is approximately $300 \text{ m} \times 300 \text{ m}$.

2) Data pre-processing

First of all, for users whose mobile phones are only used

for surfing the Internet and users whose cell phone signals do not appear on more than 50% of the working days in a month, they are filtered out and are not included in the identification of the workplace and residence.

Second, the first record is added and the last record is improved. The grid number of the newly-added first record is set to be the user's last grid number for yesterday. The start time is set as 00:00:00, and the end time is set as the start time of the existing first record on the current day. For the last record, its grid number and start time remain unchanged and its end time is set as 23:59:59.

3) Identification of users' concentrated resting period and selection of users' work analysis period

The value of the information entropy reflects the user's activity degree during the study period. The larger value indicates the higher degree of activity. The information entropy value is 0 when the user is in a completely resting state.

A working day is divided into four periods, i.e., 0:00–6:00, 6:00–12:00, 12:00–18:00 and 18:00–24:00, and each period contains six hours. The dwell time of each user in each grid is calculated for each time period, and the corresponding information entropy value is calculated as follows.

$$H(X_i) = -\sum_{j=1}^{n} P_{ij} \log_2 P_{ij}$$

In this formula, P_{ij} is the ratio of the dwell time of user X_i at stay point *j* and $P_{ij} = T_{ij}/T$, where T_{ij} is the total dwell time (in hours) of user X_i at stay point *j* during the study period, and *T* is the total length of the study period (6 h).

The period with the minimum information entropy value in the four periods of the whole day is selected as the concentrated resting period of the user. If a user has multiple periods with the same minimum information entropy value, the period of 00:00–06:00 should have higher priority; otherwise, the concentrated resting period can be selected randomly from these periods.

The characteristics of concentrated resting and working periods of various users should be comprehensively considered, such as users who work in a three-shift schedule or in a two-shift schedule, and who commute during normal commuting hours. When the user's concentrated resting period is 0:00–6:00, the period of 10:00–20:00 should be selected as the work analysis period; otherwise, the period of 1:00–7:00 should be selected.

In order to better reflect the relationship between the characteristics of users' activities and the values of information entropy, the information entropy values of two users (named User 1 and User 2) are analyzed. These values are for one consecutive week with 0.5 h as the time interval. As shown in Figure 1, the abscissa represents different days in a week; the ordinate represents different statistical time intervals within a day; and the numerical value is the information entropy value of a user's activity in the corresponding time period on the corresponding day. The redder color indicates

the larger information entropy value. When the information entropy value is 0, the user is not active. Figure 1 shows that the activities of User 1 are mainly concentrated in the daytime, and the concentrated resting period is 00:00–06:00; the activities of User 2 are mainly concentrated at night, and the concentrated resting period is 12:00–18:00. Therefore, the calculation of information entropy values can effectively avoid the issue of mis-identifying the workplace as the residence, which is possible when a single concentrated resting period is selected and applied to all users.

4) Inference of user's residence and workplace

First, the dwell time $T_{i,j,k}$ should be calculated, which is the dwell time of user *i* in grid *k* during this user's concentrated resting period (or work analysis period) on working day *i*.

Second, a user's potential residences and workplaces for a working day should be identified. Centering on the grid with the maximum dwell time of the day, spatial clustering of the grids that appear in this user's concentrated resting period (or work analysis period) is conducted. Considering that the service range of the base station is about 0.4 km–1.5 km, the clustering radius is set to be 1.5 km when inferencing the residence and set to be 0.6 km when inferencing the workplace. The total dwell time of each cluster is counted, and the users with total dwell time no less than ΔT are selected as the target users. ΔT is set to be 4 h and 2 h when inferencing the residence and the workplace respectively. For each target user, the top three grids by dwell time are selected as the potential residences (or workplaces) for the studied working day.

As shown in Figure 2, user *i* appears in grids 1, 2 and 3 during the concentrated resting period (6 h) on working day *j*, and $T_{i,j,1} \ge T_{i,j,2} \ge T_{i,j,3}$. Since the user's location should be stable during the concentrated resting period, grid 1 is taken as the center to calculate the total dwell time $T_{i,j} = \sum_{k=1}^{3} T_{i,j,k}$ of all grids within a radius of 1.5 km. If $T_{i,j}$ is no less than 4 h, the potential residences of the user on the studied working day are identified to be grid 1, grid 2, and grid 3.

Finally, all potential residence (workplace) grids in which a user appears in a month are ranked in the descending order, based on conditions such as the cumulative number of days and the cumulative dwell time. If the first potential residence (workplace) grid appears on more than half of the working days in the month, it is identified as the user's residence (workplace) grid.

In addition, users whose residence and workplace grids are the same are treated as special users, such as boarding students or stay-at-home elderly people.

2.2 Calculation results

By analyzing the one-month cellular data of China Unicom, China Mobile and China Telecom, the workplaces and residences of Tianjin residents are identified, and the job-housing spatial distribution is analyzed.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday		Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	S
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00	1.13	0.00	0.00	0.90	0.00	- ì
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.5	0.00	0.99	0.00	0.00	0.61	0.00	
1.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.0	0.65	1.27	0.00	0.00	0.00	1.22	
1.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.5	0.00	0.00	0.00	0.36	0.91	0.65	
2.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.0	0.00	0.00	0.00	0.00	0.00	0.00	
2.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.0	0.00	0.98	0.00	0.00	0.95	0.00	
3.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.0	0.00	0.69	0.00	0.00	0.93	0.00	
3.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.5	0.55	1.56	0.00	0.00	0.00	1.12	
															1
4.0 4.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.0 4.5	0.00	0.91 0.16	0.83	0.00 0.00	0.00	0.04	
								4.5							
5.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	1.00	0.89	0.28	0.84	0.89	
5.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.5	0.00	0.99	0.74	0.95	0.88	0.84	
6.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.0	0.00	0.99	0.00	0.00	1.40	0.00	
6.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.5	0.94	0.50	0.00	0.46	0.00	0.00	
7.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.0	0.00	0.00	1.08	0.00	0.00	0.00	
7.5	1.00	0.93	1.27	0.00	0.00	0.00	0.00	7.5	0.00	0.19	0.97	0.00	0.75	0.40	
8.0	0.00	0.94	0.00	0.00	0.00	0.00	0.00	8.0	0.00	1.57	0.00	0.37	0.24	1.18	
8.5	0.00	0.00	1.00	0.00	0.00	0.00	0.00	8.5	0.79	0.00	1.00	1.00	0.38	0.00	
9.0	0.00	0.00	0.00	1.76	0.00	0.00	0.00	9.0	0.83	0.78	0.00	0.00	0.00	1.00	
9.5	1.73	1.20	0.00	0.00	0.76	0.00	0.21	9.5	0.62	0.00	0.00	0.00	0.00	0.00	
10.0	0.59	0.00	1.82	0.64	1.73	1.75	1.19	10.0	0.83	0.00	0.77	0.98	0.78	0.98	
10.5	1.34	0.00	2.32	0.00	0.27	1.36	1.45	10.5	0.75	0.00	0.00	0.00	0.00	0.00	
11.0	0.00	1.02	1.68	0.00	0.87	0.00	0.00	pe 11.0 11.5	0.00	0.00	0.00	0.24	0.00	0.88	
11.5	1.34	1.07	0.88	0.94	0.78	0.00	0.00	la 11.5	0.00	0.00	0.00	0.00	0.00	0.00	
12.0	0.00	0.40	1.00	1.25	0.73	0.63	0.00	12.0 12.5 13.0	0.96	0.00	0.00	0.00	0.00	0.00	
12.5	0.00	0.00	0.00	1.73	1.93	0.00	0.00	雲 12.5	0.00	0.00	0.00	0.00	0.00	0.00	
13.0	0.00	0.75	0.00	2.03	0.90	0.00	0.00		0.00	0.00	0.35	0.00	0.00	0.00	
13.5	0.00	0.00	1.32	0.08	1.23	0.00	0.00	13.5	0.00	0.00	0.00	0.00	0.00	0.00	
14.0	0.96	0.00	0.98	0.00	1.36	0.00	0.00	14.0	0.00	0.00	0.00	0.00	0.00	0.00	
14.5	0.91	1.03	0.00	0.00	0.00	0.60	0.00	14.5	0.00	0.98	0.00	0.00	0.00	0.00	
15.0	0.00	1.54	0.07	0.98	0.00	0.00	0.00	15.0	0.00	0.00	0.94	0.00	0.00	0.00	
15.5	0.74	0.00	0.00	0.00	0.00	0.00	0.00	15.5	0.00	0.00	0.00	0.00	0.00	0.00	
16.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.0	0.00	0.00	0.00	0.00	0.00	0.00	
16.5	0.00	1.42	0.00	0.00	0.00	0.00	0.00	16.5	0.00	0.00	0.00	0.00	0.00	0.64	
17.0	0.55	1.53	1.33	0.00	1.30	0.00	0.00	17.0	0.00	0.00	0.00	0.00	0.00	0.00	
17.5	1.73	2.25	1.54	1.02	1.84	0.00	0.00	17.5	0.00	0.00	0.00	0.00	0.00	0.00	
18.0	0.90	0.00	0.13	0.00	1.39	0.00	0.00	18.0	0.00	0.00	0.00	0.00	0.00	0.00	
18.5	1.21	0.00	1.40	0.00	1.71	0.00	0.00	18.5	0.00	0.00	0.00	0.39	0.00	0.00	
19.0	0.51	0.00	0.49	0.00	1.22	0.83	0.84	19.0	0.38	0.76	0.96	0.95	0.00	0.00	
19.5	0.00	0.00	0.00	0.00	0.00	1.01	0.00	19.5	1.31	0.83	1.00	0.46	0.26	0.90	
20.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	20.0	0.00	0.79	0.00	0.63	1.56	0.00	
20.5	0.00	0.00	0.98	0.00	0.00	0.00	0.00	20.5	0.99	1.52	0.00	0.69	0.58	0.00	
21.0	0.00	0.00	0.53	0.00	0.00	0.00	0.00	21.0	0.94	0.77	0.00	0.90	0.43	0.05	
21.5	0.00	0.00	0.00	0.00	0.00	0.00	0.14	21.5	1.42	0.00	0.83	0.71	0.69	0.28	
22.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	22.0	1.00	1.53	0.00	0.82	0.00	0.33	
22.5	0.00	1.17	0.00	0.00	0.00	0.00	0.00	22.5	0.20	0.53	0.00	0.00	0.28	0.97	
23.0	0.00	0.00	0.00	0.00	1.17	0.00	0.00	23.0	0.39	0.00	0.00	0.47	0.00	0.00	
23.5	0.00	0.00	0.00	0.00	1.33	0.00	0.00	23.5	0.00	0.00	0.70	0.00	0.00	0.48	
			Working day	s		Week	cends				Working day	s		Wee	ken

Figure 1 Information entropy distribution by time

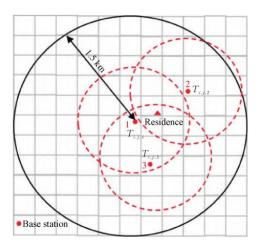


Figure 2 Recognition of potential residence place

1) Overview of the identification of the workplace and residence

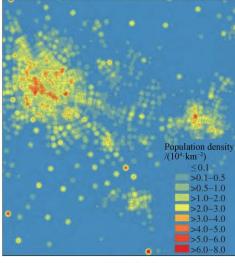
The residences of about 8.65 million Tianjin residents are identified, which is equivalent to sampling 55% of the permanent residents in Tianjin. The workplaces of about 6.97 million residents are identified, and there are still 2.89 million residents after filtering out special users (such as boarding students, community business owners, and stay-at-home elderly people), among which about 10% are identified as working at night (shown in Table 1).

2) Population density and employment density in twin cities of Tianjin

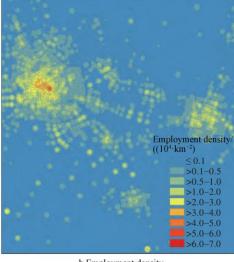
As shown in Figure 3, in Tianjin's central urban area, residents are mostly distributed within the fast ring road and jobs are highly concentrated along the Nanjing Road. In the core area of Binhai New Area, residents are mostly distributed in Tanggu Old Town and jobs are concentrated in areas such as the development zone and the waterfront area, which are basically consistent with the actual situation.

Table 1 Overview of work and housing locations of Tianjin residents $\!\!\!/10^4$

Operator	Number of users whose residences are identified	Number of users whose workplaces are identified	Number of users whose workplace and residence are located in different grids
China Unicom	222.83	231.42	86.45
China Mobile	531.99	386.44	168.06
China Telecom	110.08	79.27	34.57
Total	864.90	697.13	289.08



a Population density



b Employment density

Figure 3 Population density and employment density in twin cities of Tianjin

3 Main ideas for analyzing urban spatial structure based on job-housing distribution

The urban spatial structure connects various elements (such as economic activities, land use, buildings, and social groups) within the city, urban forms, and interactions between cities in a certain organizational way to form a complete system^[11]. The use of cellular data to analyze the urban spatial structure mainly reflects the law of human activities, which indirectly reflects the urban spatial form, so it is more inclined to analyze the urban spatial structure from the perspective of geography.

The analysis of the urban spatial structure based on the characteristics of job-housing distribution, which are obtained through cellular data, mainly focuses on three aspects (shown in Figure 4), including the scope of urban spatial expansion, the size, structure and influence range of the urban CBD, and the strength of connections among areas.

1) The scope of urban spatial expansion is the extent of a metropolitan area. Employment Interchange Measure (EIM) is the core indicator used to define metropolitan areas in foreign countries. Therefore, EIM of the central urban areas and peripheral area clusters can be obtained through cellular data, and the scope of the commuting area can be studied to define the main scope of urban spatial expansion.

2) Urban form is mainly analyzed for the urban CBD. Through cellular data, the population density and employment density can be obtained, and the scale and extent of the CBD can be defined. The distribution law of workers and residents in the CBD can then be further analyzed to study the influence range of the CBD.

3) Spatial connection among areas mainly analyzes the strength of commuting travel connection. The job-housing distribution relationship can be obtained through cellular data, and can then be used to study the strength of commuting travel connection among areas.

4 Relationship between job-housing spatial distribution and rail transit network

4.1 Commuting area analysis

Commuting areas are also called metropolitan areas. In general, EIM of the peripheral areas and the central urban area is used as the main indicator to express the spatial hierarchy structure of a metropolitan area. EIM is the sum of the percentage of workers living in the peripheral areas who commute to the central urban area and the percentage of employment in the peripheral areas that is accounted for by workers who reside in the central urban area. The criterion for metropolitan areas proposed by the U.S. in 1990 is that at least 15% of non-agricultural workers commute to the central urban area or the EIM (total of in- and out-commuting) reaches more than 20%. Japan defines metropolitan areas with the criterion of 5% or 10% EIM ^[12].

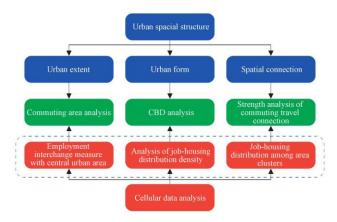


Figure 4 Techniques of urban spatial structure based on job-housing distribution

Tianjin presents a spatial structure model with the central urban area and the core area of Binhai New Area as its dual cores. Therefore, the analyses of commuting area are conducted for the central urban area and the core area of Binhai New Area respectively. The peripheral areas are divided into area clusters and their EIM with the twin cities are calculated. For each EIM criterion of 5%, 10%, 15% and 20%, the proportions of workers' residences and residents' workplaces covered by the area clusters are calculated. The EIM criterion of 10% is finally selected as the threshold to define commuting areas (shown in Figure 5, Table 2 and Table 3), so that the proportion of residents' workplaces covered by the commuting area is over 96%.

The analyses show that 1) in terms of the area, the commuting area of the central urban area is much larger than that of the core area of Binhai New Area, indicating that the central urban area is the primary core of the dual-core structure, and the core area of Binhai New Area is the secondary core; 2) in terms of the commuting radius, the average commuting radius of the central urban area is about 27 km, while that of the core area of Binhai New Area is about 20 km, both are relatively small; 3) in terms of the commuting direction, whether from the aspect of commuting range or commuting volume, commuting from inside to outside is greater than commuting from outside to inside for both cores. The ratio of inward to outward commuting for the central urban area is 1:1.87, while that for the core area of Binhai New Area is 1:1.08 (shown in Table 4). These ratios indicate that since Tianjin is an industrial city at this stage, the commuting volume to the secondary industry jobs in the peripheral area is greater than that to the tertiary industry jobs in the twin cities of Tianjin.

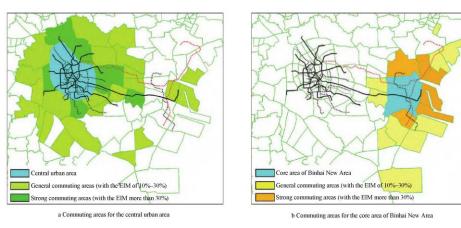


Figure 5 Commuting loops in twin cities of Tianjin Table 2 Commuting loops in urban central area

Employment Interchange Measure	Area/km ²	For workers who work in the central urban area, the proportion of residences covered by the commuting areas/%	For workers who live in the central urban area, the proportion of workplaces covered by the commuting areas/%
Area clusters with more than 5% of workers working in the central urban area or with more than 5% of jobs taken by workers living in the central urban area	4 161	99	98
Area clusters with more than 10% of workers working in the central urban area or with more than 10% of jobs taken by workers living in the central urban area	2 306	98	97
Area clusters with more than 15% of workers working in the central urban area or with more than 15% of jobs taken by workers living in the central urban area	1 805	98	96
Area clusters with more than 20% of workers working in the central urban area or with more than 20% of jobs taken by workers living in the central urban area	1 609	98	94

 Table 3
 Commuting loop in core area of Tianjin Binhai New Area

Employment Interchange Measure	Area/km ²	For workers who work in the core area of Binhai New Area, the proportion of residences covered by the commuting areas/%	For workers who live in the core area of Binhai New Area, the proportion of workplaces covered by the commuting areas%
Area clusters with more than 5% of workers working in the core area of Binhai New Area or with more than 5% of jobs taken by workers living in the core area of Binhai New Area	1 470	87	96
Area clusters with more than 10% of workers working in the core area of Binhai New Area or with more than 10% of jobs taken by workers living in the core area of Binhai New Area	1 275	85	96
Area clusters with more than 15% of workers working in the core area of Binhai New Area or with more than 15% of jobs taken by workers living in the core area of Binhai New Area	976	85	94
Area clusters with more than 20% of workers working in the core area of Binhai New Area or with more than 20% of jobs taken by workers living in the core area of Binhai New Area	876	85	94

Table 4 Commuting directions in twin cities of Tianjin

Commuting directions	Central urban area	Core area of Binhai New Area
The number of samples living in the central urban area and working in the peripheral area (commuting from inside to outside)	190 101	48 086
The number of samples living in the peripheral area and working in the central urban area (commuting from outside to inside)	101 392	44 257
The ratio of inward to outward commuting (number of samples commuting from outside to inside/ number of samples commuting from inside to outside)	1:1.87	1:1.08

According to the *Tianjin Urban Rail Transit Short-term Construction Plan*^[13], on the basis of existing Rail Transit Lines 1, 2, 3, and 9 and Rail Transits Lines 5 and 6 that are about to open, Rail Transit Lines 4, 7, 8 and 11 will be built in the central urban area of Tianjin. In the Binhai New Area, Rail Transit Line B1 and Inter-city Rail Transit Lines Z2 and Z4 will be built.

From the perspective of the coupling relation between the commuting area and the rail transit network (shown in Figure 6), rail transit covers most area clusters with strong commuting connection between the peripheral areas and the twin cities (the EIM is over 30%), but there are also deficiencies. 1) The coverage of some area clusters with strong commuting connection is insufficient, such as the western area cluster of the central urban area. 2) The construction timing and functional positioning of inter-city rail transit lines need further study. For example, Line Z4 runs south-north with Han'gu at its north end, which has weak commuting connection with the core area of Binhai New Area. The construction timing of this section of rail transit line is open to question. Line Z2 runs east-west and passes through the northern periphery of the twin cities, with the major function to serve the commute between the industrial zones of the twin cities. However, passengers on Line Z2 need to transfer to access the central areas. Therefore, the function of Line Z2 is different from the typical function of inter-city rail transit, which is to mainly serve the commute between central areas, and the functional positioning of Line Z2 needs to be studied further.

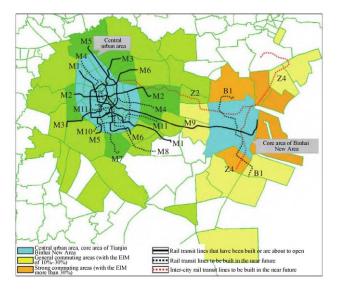


Figure 6 Relationship between commuting loop and recently-constructed rail transit network in twin cities of Tianjin

4.2 CBD Analysis

The distribution of employment density clearly shows the extent of the twin cities' CBD areas. Since the energy level of the CBD in the core area of Binhai New Area is low at this stage, the CBD analysis focuses on the CBD of the central urban area and its extent is shown in Figure 7. The area of this CBD is about 10.4 km², accounting for 0.5% of the commuting area. The resident population is about 345 000, accounting for 4.0% of the population living in the commuting area, among which about 180 000 are commuters. The number of jobs is about 465 000, accounting for 8.9% of jobs in the commuting area. In general, the energy level of this CBD is relatively small, and its aggregation effect must be further strengthened. Therefore, from the perspective of exerting the CBD effect, it is still necessary to gather rather than disperse at this stage.

In terms of the CBD's job-housing distribution, the CBD residents mostly work in the CBD and its immediate surrounding areas (shown in Figure 8a), which accounts for 61.2%; and workers working in the CBD mostly live in the areas within the outer ring road (shown in Figure 8b). The influence range of the CBD is small, with 82.4% of the CBD jobs taken by workers living outside.

From the perspective of the relationship between the residence distribution of the workers working in the CBD and the radial rail transit lines, the support of the rail transit network for the CBD needs to be strengthened, which is mainly reflected in two aspects: 1) some main commuting corridors lack rail transit coverage, or their main directions have not been covered (the commuting corridors are indicated by the red arrows in Figure 9a); and 2) the distance between stations on some rail transit lines is too large, leading to insufficient coverage of CBD's rail transit stations within the 500-m radius, which is only about 70% (shown in Figure 9b).

4.3 Strength analysis of commuting travel connection

Spatial connection can be divided into spatial connection of people flow and goods flow. The spatial connection of people flow can be further divided into commuting travel connection and non-commuting travel connection. The job-housing spatial distribution can be obtained through cellular data, so as to analyze the strength of commuting travel connection.

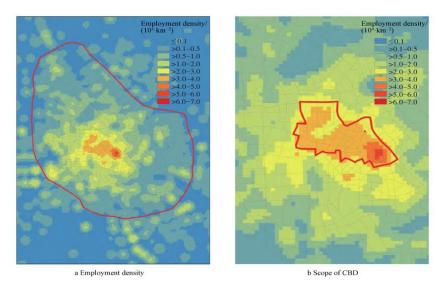
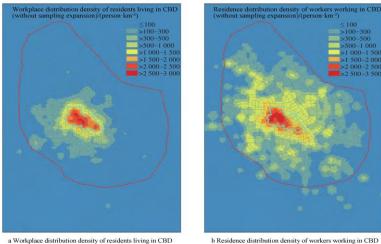


Figure 7 Employment density and scope of CBD in urban central area



a Workplace distribution density of residents living in CBD

Figure 8 Job-housing distribution in CBD

Figure 10 shows that 1) commuting travel is generally centered on the dual cores, and the overall commuting connection between the dual cores is not large; 2) the commuting connection between adjacent areas is stronger; and 3) the CBD in the central urban area has an obvious characteristic of outward radial travel.

From the relationship between the loop line composed of Rail Transit Lines 5 and 6 in the central urban area and commuting travel (shown in Figure 11), the rail transit loop line is consistent with the commuting travel loop outside the CBD. The location of the loop line is basically reasonable, but the layout pattern is debatable. The combination loop line occupies the location of four radial lines, and three of the four radial corridors have strong connection with the CBD (which are M6 North, M6 South, and M5 South respectively). Therefore, rail transit passengers going to the CBD through these three corridors need to transfer at least once. Since important radial corridors are invaluable channel resources,

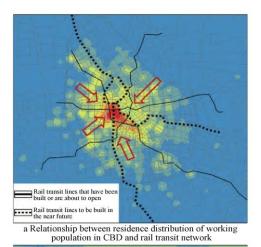
from the perspective of passenger travel, the layout mode of the combined rail transit loop line in the central urban area of Tianjin needs to be discussed, and the complete loop mode is relatively reasonable.

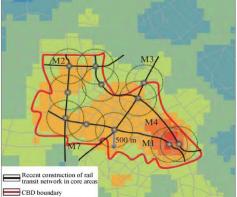
5 Conclusion

The synergy relationship between existing and recentlyconstructed rail transit network and the current urban spatial structure is ordinary and there is room for improvement. 1) In terms of commuting area, the coverage of rail transit network in the peripheral area clusters with strong commuting connection is insufficient. The construction timing and functional positioning of some inter-city rail transit lines in Binhai New Area need to be further studied. 2) In terms of the CBD, some main commuting corridors lack rail transit coverage, or their main directions have not been covered. The coverage of

CBD's rail transit stations within the 500-m radius is insufficient, only about 70%. 3) In terms of the strength of commuting travel connection, the layout mode of the combined loop line in the central urban area needs to be discussed.

The above-mentioned conclusions are mainly based on the relationship of current job-housing distribution characteristics and urban spatial structure with rail transit network, but rail transit construction has a strong guiding effect on urban land use and spatial layout. With the improvement of the rail transit network, urban spatial structure will also change accordingly. Therefore, it is recommended to conduct regular





b Coverage of CBD's rail transit stations within the 500-m radius

Figure 9 Relationship between CBD and radial rail transit network

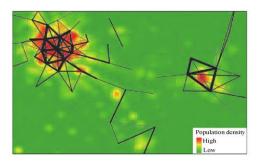


Figure 10 Degree of commuting travel connecting

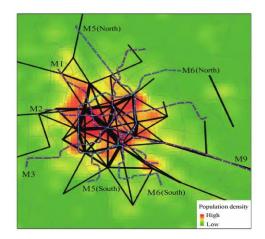


Figure 11 Relationship between loop lines of rail transit in urban central area and commuting travel

evaluation on the relationship between urban spatial structure and rail transit network based on cellular data, providing timely scientific basis for the construction of rail transit network and the adjustment of urban spatial structure.

References

- [1] Tianjin Urban Planning and Design Institute. 2017 年天津市年度交通发 展报告 [R]. Tianjin: Tianjin Planning Bureau, 2017 (in Chinese).
- [2] Wang Wenjing, Lu Huapu. Research on the Coupling of Metropolis Circle Rail Transit and the Spatial System [J]. Urban Development Studies, 2013, 20 (4): 112–118.
- [3] Xu Ning. Research of Distribution Features of Home-Work Location Based on Mobile Phone Positioning Data in Shenzhen [D]. Changsha: Central South University, 2014.
- [4] Lai Jianhui. Research on Data Mining and Analysis in Transportation Based on Mobile Communication Location [D]. Beijing: Beijing University of Technology, 2014.
- [5] Leng Bingrong, Yi Zheng, Cao Chunxia, et al. 基于手机信令数据的重 庆主城区职住关系识别及其交通响应 [C] // Urban Planning Society of China, 60 Years of Planning: Achievements and Challenges – Proceedings of Annual National Planning Conference 2016, Shenyang: China Architecture & Building Press, 2016: 1095–1109 (in Chinese).
- [6] Ding Liang, Niu Xinyi, Song Xiaodong. Identifying the Commuting Area of Shanghai Central City Using Mobile Phone Data [J]. City Planning Review, 2015, 39 (9): 100–106.
- [7] Zhang Tianran. Job-housing Spatial Distribution Analysis in Shanghai Metropolitan Area Based on Cellular Signaling Data [J]. Urban Transport of China, 2016, 14 (1): 15–23.
- [8] Zhou Yongjie, Liu Jiezhen, Zhu Jinfeng, et al. Research on the Spatial Characteristics of Pearl River Delta Urban Agglomeration Based on Mobile Signaling Data [J]. Planners, 2018, 34 (1): 113–119.
- [9] Gong Hang, Sun Li, Wang Pu. Using Mobile Phone Data to Estimate Trip Distribution of Urban Mega Traffic Analysis Zones: A Case Study in San Francisco [J]. Urban Transport of China, 2016, 14 (1): 37–42+50.
- [10] Song Shaofei, Li Weifeng, Yang Dongyuan.Research on the Methods of Home Identification Based on Mobile Phone Data [J]. China Transportation Review, 2015, 37 (12): 72–76.
- [11] Bourne L S. Internal Structure of the City: Readings on Urban Form, Growth and Policy [J]. Historian, 1982, 26 (1): 1–18.
- [12] Hong Shijian, Huang Xiaofen. Study on the Concept and Definition of Metropolitan Areas [J]. Urban Planning International, 2007, 22 (5): 50–57.
- [13] Third Railway Survey and Design Institute Group. 天津市城市轨道交通建设规划 (2015–2020年) [R]. Tianjin: Tianjin Rail Transit Group Co., Ltd., 2014 (in Chinese).