Design and Research of Wuhan Urban Railway Network Simulation Calculation Information System Utilizing Big Data

Han Zou* and Shujun Liu
Hubei University of Technology, Wuhan, China
*Corresponding author: zouhangogo@qq.com

Abstract. Based on the big data-assisted algorithm, the paper uses the dynamic relationship data of railway passenger flow as a research case, adopts the urban network chain model, and uses the mobility and agglomeration of railway passenger flow between cities as the research entry point to analyse the network connection strength of provincial cities and spatial organization structure characteristics. At the same time, it puts forward the idea of using Wuhan's rich railway resources as the framework of urban railway network passenger flow intensity simulation calculation system, demonstrates the necessity and feasibility of its construction, and puts forward an implementation plan based on the characteristics of the railway network and urban development for the reference of relevant decision makers.

Keywords: Big data, spatial data model, simulation system, passenger flow intensity, railway network.

1. Introduction
As a large central city in central China, Wuhan has four transportation elements due to its unique geographical environment: railway, water transportation, expressway and aviation. It is especially famous for its highly developed railways, which play an important role in connecting the east, west, north and south, maintaining the layout of the national railway network and other directions. It is the central city of China's railway network, one of the nine railway hub cities, and has the largest railway marshalling yard in Asia [1]. It is the transportation center of Beijing-Guangzhou, Xiang-Han, Han-Jiu (River) Railway and the intersection of the Yangtze River and the Hanshui River. It is known as the "thoroughfares of nine provinces".

At the same time, with the development and wide application of network communication technology, positioning and navigation technology and big data, it becomes easier and easier to obtain massive spatiotemporal data, which records the spatiotemporal changes of moving objects (people, cars, etc.). The generation of big data with spatiotemporal attributes makes it possible to track the spatiotemporal movement characteristics of research objects for a long time, with high accuracy and efficiency [2].

Therefore, under the background of the rise of central China, the macro strategy of the Yangtze River Economic Belt and other countries and economic globalization, Wuhan, as an important hub of China's high-speed railway passenger dedicated line network, combines big data technology, sensor technology, monitoring technology, pattern recognition technology and other advanced technologies.
Building a railway network simulation computing system based on big data can provide methods and decision-making guidance for researches in Wuhan city planning and management, traffic monitoring and forecasting, tourism monitoring and analysis and many other fields.

2. Overview of Wuhan urban railway network

2.1. Introduction to Wuhan Railway Hub

At present, Beijing-Guangzhou Railway and Beijing-Guangzhou high speed railway run through the north and South in Wuhan. Wuhan-Jiujigang Railway, Hefei-Wuhan Railway, Macheng-Wuhan connecting line, Wuhan-Ankang Railway and Wuhan-Yichang Railway are respectively introduced in the east-west direction. They are network hubs connecting the six railway trunk lines. Major large passenger stations (Wuhan Station, Hankou Station, Wuchang Station), marshalling stations (Wuhan North Station) and freight facilities are concentrated on both sides of the Beijing-Guangzhou Corridor, and a passenger and cargo transportation pattern has been formed with Beijing-Guangzhou Railway, Beijing-Guangzhou High Speed Railway, Tianxingzhou Bridge and the First Bridge of the Yangtze River as the main axis. According to the plan, Wuhan railway hub will be upgraded to the Shanghai-Wuhan-Chengdu Expressway in the east-west direction, and the Beijing-Kowloon high-speed railway in the north-south direction will be added with a short connecting line to connect Wuhan. Efforts will be made to build an oblique road network expressway connecting Xi'an, Qingdao, Nanning, Kunming, Xiamen, Ningbo, etc., and finally form a railway network that can reach the major urban agglomerations in China. It is a comprehensive transportation hub connecting "two longitudinal, two transverse and two oblique" 12 direction radial high-speed railway or intercity railway [3]. The current layout of Wuhan railway hub is shown in Figure 1.

![Figure 1. Wuhan High Speed Rail Map (Image credit: chinadiscovery.com).](image)

2.2. Distribution of railway passenger stations

According to the overall planning of Wuhan hub and the planning of Wuhan metropolitan area intercity railway network, four passenger railway stations will be formed in Wuhan: Hankou station to the north of the Yangtze River, Wuchang station, Wuhan station and Liufang station to the south of the Yangtze River. Among them, Wuchang, Hankou and Wuhan railway stations have been completed and put into operation, and Liufang station started construction in 2009. These four railway passenger stations are the main distribution points of Wuhan urban passenger flow. As a comprehensive transportation hub, they integrate a variety of transportation modes, such as long-distance bus, bus, subway, train and so on.
2.3. Socio-economic factors
Taking the Beijing-Guangzhou railway line as an example, it extends southwest from Wuchang along the Beijing-Hong Kong-Macao Expressway and National Highway 107, passing through Huangjiahu University Town. In addition, there are mainly natural villages along the way, belonging to the agricultural and fishing areas [4]. Industry, commerce and services are relatively developed and densely populated. Along the way, there are many universities, science and technology industrial parks, high-tech enterprises, shopping centers and residential areas. The urbanization is developing rapidly and the economy is active. Wuhan City Circle to undertake "Pearl River Delta" industrial transfer and Changsha-Zhuzhou-Xiangtan City Circle contact the main channel.

3. The establishment of railway network strength simulation calculation system framework

3.1. Spatial data system architecture
Spatial data system can realize the storage of large-scale spatial data and provide users with the function of spatial data analysis and mining. Important spatial data queries include spatial region queries, spatial association queries and spatial k-nearest neighbor queries. A spatial scope query queries a spatial object that is located within the query region. For example, find a railway station in an area on a map. A spatial associative query is a query that combines two spatial data sets and a predicate condition. The predicate condition can be the distance, such as: query is located in Wuhan urban area, less than 2 kilometers apart from the railway stations.

3.2. Building the basic idea
An ideal road network simulation system should have these characteristics: 1) The simulation model should cover different simulation objects in the road network (such as stations, roads, locomotives, vehicles, signals, etc.) and can accurately describe the behavior of different simulation objects (Such as locomotive operation, signal changes, etc.). 2) The simulation model should complete the simulation of different granularities to support the research needs of different levels (such as marshalling plan, operation diagram, etc.). 3) The simulation scale is variable, which can not only meet the needs of road network simulation, but also simulate a certain part of the road network. The mechanism of the simulation system has nothing to do with the simulation scale. 4) The core simulator in the simulation system is separated from the simulation results, that is, the simulation results can be expressed in different forms. The paper decomposes the entire simulation system into three relatively independent parts: simulation object, simulation environment, simulation result and its expression (Figure 2). On the one hand, the simulation result can be expressed in different forms. On the other hand, although the simulation object can be different the levels are abstract, but the simulation environment remains relatively unchanged.

![Figure 2. Basic relationship diagram of the simulation system (Image credit: drawed by author).](image)
As shown in Figure 2, the simulation task of this railway transportation simulation system is decomposed into multiple subtasks, including n simulation objects and k expression modes. Relative independence between subtasks [5]. The content of the simulation object is not fixed. For example, in a certain operation of the system, the simulation object 1 may represent a collection of several stations. In a certain operation after the simulation purpose is changed, the simulation object 1 can represent the collection of all stations and operating sections in a certain area of the road network. The same or different simulation results can be expressed in many different ways. Different simulation schemes applied to the system will produce different simulation results and different simulation task decomposition schemes.

3.3. The structure of road network simulation calculation system based on big data
According to a comprehensive analysis, the simulation environment in the road network simulation system must not only adapt to the requirements of simulation objects of different granularities, but also meet the needs of cooperative communication of various simulation objects, which plays an important role in the entire system. To build this environment, you can learn from HLA’s ideas (high-level architecture) and introduce the technology of soft bus software components, see Figure 3.

3.4. The basic function analysis of analog computing system
The application program interface in Figure 4 is the interface between the Agent and the simulation member, and the completed tasks include: the simulation member uploads registration information, current status information, interactive data declaration, simulation data transmission, clock advance request, etc. Agent conveys experiment control instructions and experiment configuration information to the simulation members, and sends registration results, task distribution information, clock advance instructions, etc. The external interface is the interface to the manager, which conveys experimental control instructions and configuration information downwards.
Figure 4. Basic functional block diagram of local simulation environment (Image credit: drawed by author).

1) The communication management module, according to the relevant information provided by the upper layer, uses the services provided by network communication to manage the establishment of communication connections and data transmission processes between simulation members. 2) Simulation task management module, for simulation tasks, taking into account the specific computer load capacity and total task volume, achieve the maximum load uniformity, and transfer tasks appropriately when the network is interrupted. 3) The data distribution management module needs to send control instruction data to the node where the relevant simulation member is located, establish interactive data matching for the interaction of simulation data, and be able to judge the flow direction and communication channel of the simulation data and call the data sending function. 4) The simulation clock management module receives the simulation clock advance request of each simulation node, and realizes the clock synchronization control of each simulation node based on the regional distribution.

4. Overview of the study area, data sources and research methods

4.1. Data source

This study collected daily train population migration data between 17 prefecture-level cities in Hubei Province from October 21, 2020 to November 20, 2020 (one month before and after) through the Tencent location big data platform. The data includes the starting point, the total number of population migrations between the city and the destination city. According to statistics, during this period, the total railway flow between cities in Hubei province reached 13.5 million person-times. This was used as the basic data to study the network relationship and spatial organization structure between cities in Hubei Province, and the central location of each city was set as each in the centre of the city, each prefecture-level city is used as the basic space unit to collect and sort out passenger flow data generated by moving in and moving out.

4.2. Measurement method

Urban network research from the perspective of flow space mainly uses the "chain network model" proposed by British geographer Taylor to calculate the sum of the inbound and outbound railway passenger flows between two cities to measure the strength of the relationship between cities. The specific method is:

4.2.1. Contact strength \( \left( S_{ij} \right) \). Assuming that each city in the railway network has train connections with at most other n-1 cities, the corresponding expression is:
\[ (S_y) = S_{i,j} + S_{j,i} \]  

(1)

Where: \( S_{i,j} \) is the railway passenger volume from city \( i \) to city \( j \). \( S_{j,i} \) is the passenger traffic volume from city \( j \) to city \( i \). \( S_y \) is the connection strength of the two cities, reflecting the mobility of the city's passenger flow in the regional network.

4.2.2.  Agglomeration intensity \( S_m \). Let \( S_m \) be the sum of the connection intensity between city \( m \) and other cities, \( S_m \) is the sum of the connection intensity between city \( i \) and other cities, and \( S_m \) is the sum of the connection intensity between city \( m \) and city \( i \), reflecting a certain city the ability to gather passenger flow in the regional association network. The corresponding expression format is:

\[ (S_m) = \sum_{m=1}^{n} (S_m + S_m) \]  

(2)

In the formula: \( S_m \) is the sum of the connection strength between city \( m \) and other cities. \( S_m \) is the sum of the strength of connection between city \( i \) and other cities. \( S_m \) is the sum of the connection strength between city \( m \) and city \( i \), reflecting the passenger flow gathering ability of a certain city in the regional association network.

5.  Data analysis of passenger flow intensity based on railway network simulation system

According to the above railway network simulation calculation system, it can be seen from Table 1 that the central city Wuhan has the largest correlation with the passenger flow in the eastern Hubei, followed by the passenger flow in the central Hubei, while the correlation intensity between the passenger flow and the western Hubei presents a low level and equilibrium. The average railway passenger flow from eastern Hubei to Wuhan reached 509,000 person-times, and the average railway passenger flow from Wuhan to eastern Hubei reached 480,700 person-times, both of which were higher than that from Wuhan, central Hubei and western Hubei. As shown in Table 1, this is 1.8 times of the average passenger flow intensity of the Wuhan-Western Hubei railway.

Table 1. Inter-regional city passenger flow intensity.

| Departure                  | Wuhan | Central Hubei cities (except Wuhan) | Edong city | City of Western Hubei |
|----------------------------|-------|-------------------------------------|------------|-----------------------|
| Wuhan                      | -     | 144.22                              | 171.75     | 116.03                |
| Central Hubei cities (except Wuhan) | 152.94 | -                                   | 6.03       | 9.05                  |
| Edong city                 | 183.54| 4.71                                | -          | 39.66                 |
| City of Western Hubei      | 128.29| 7.27                                | 24.00      | -                     |

Central Hubei and East Hubei are the three regions that have the strongest correlation with Wuhan passenger flow. The three cities with the strongest passenger flow in Wuhan are Xiaogan, Huanggang and Xianning. The flow of passengers moving in and out is relatively balanced, which is in line with the objective fact that passengers are traveling-returning. Further analysis of the proportion of railway passenger flow moving into Wuhan from cities in the three major regions to the total passenger flow out of the city, the study found that the passenger flow moving into Wuhan from cities in the central Hubei region accounted for 73.35% of the total passenger flow out of the region. Among them, the railway passenger flow that Xiaogan moved into Wuhan accounted for more than 80%. Similarly, the railway passenger flow moving into Wuhan from the eastern Hubei region accounted for 60.37% of the total passenger flow from the region, while the western Hubei region was only 51.75%.
6. Conclusion
Based on the present situation of Wuhan railway network planning, this paper systematically introduces the preliminary system framework, key links and operating simulation environment of the railway transportation distributed simulation system, aiming at the realization of large-scale railway transportation network simulation calculation. Finally, the operability of the simulation system was verified by taking the intensity of inter-regional passenger flow as an example, that is, the data of railway passenger flow in Wuhan and its surrounding cities were quantified and analyzed. Through the analysis results, the correlation degree between the intensity of passenger flow distribution in Wuhan railway network and the inter-city railway network was known. Through the analysis, calculation and prediction of the intensity of the passenger flow in the railway network, the design can provide reference for the optimization design of the railway network in the future. It is suitable for the development of the current railway transportation and has good reusability and expansibility.

References
[1] He Wenbiao. Planning and design of Wuhan railway hub. Railway standard design, 2009 (01): 1-4.
[2] Gao Qiang, Zhang Fengli, Wang Ruijin, Zhou fan. Trajectory big data: review of key technologies of data processing. Acta Sinica Sinica, 2017, 28 (04): 959-992.
[3] Zhou Tianjie. Research on Wuhan Railway Hub Planning. Transportation science and technology, 2009 (03): 108-111.
[4] Zhou Tianjie. Rethinking the general layout of Wuhan railway hub. Railway standard design, 2014, 58 (03): 1-4 + 9.
[5] Li Wendong. Research and implementation of big data mining technology based on spark. Shandong University, 2015.
[6] Zhou fan. Research on intercity railway network planning in Wuhan metropolitan area. Southwest Jiaotong University, 2014.
[7] Yin Qiaoxuan, Duan bin, Kang Canping, Li Hui. Co simulation design of energy system and information communication system based on HLA / agent. Power system automation, 2016, 40 (17): 22-29 + 58.
[8] Deng Junlei. Study on connection scheme and selection between Wuhan hub railway and urban rail transit. Beijing Jiaotong University, 2012.
[9] Hu Xiaohong, Fang Zhicheng. Research on GSM-R wireless network planning of Wuhan railway hub. Railway technology innovation, 2011 (02): 26-29 + 89.
[10] Lu zongfei. Design and implementation of spatial data platform system based on spark. Shandong University, 2017.
[11] Le Peng, Wu Zhaoyan, Shangguan Boyi. Design and implementation of distributed spatial data storage structure based on spark. Journal of Wuhan University (Information Science Edition), 2018, 43 (12): 2295-2302.
[12] Van Gulijk, C., Hughes, P., Figueres-Esteban, M., El-Rashidy, R., & Bearfield, G. The case for IT transformation and big data for safety risk management on the GB railways. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 232 (2) (2018) 151-163.
[13] Wen, C., Li, Z., Lessan, J., Fu, L., Huang, P., & Jiang, C. Statistical investigation on train primary delay based on real records: evidence from Wuhan–Guangzhou HSR. International Journal of Rail Transportation, 5 (3) (2017) 170-189.