Frontiers in Traffic and Transportation Engineering (2025)
Clausius Scientific Press, Canada
DOI: 10.23977/ftte.2025.050107
ISSN 2616-2334 Vol. 5 Num. 1

Research on Data Compilation of Parallel Scheduling System for Railway Technical Station

Yutian Jiao*, Junting Liu, Junwei Wang

College of Rail Transit, Shandong Jiaotong University, Jinan, 250357, China
**Corresponding author*

Keywords: ACP Approach, Railroad Technical Operation Station, Intelligent Processing, Data Compilation, Parallel Scheduling System

Abstract: The railway technical operation station is a huge and complex operation system, and its scheduling management is faced with a variety of random disturbance factors, which increases the complexity of scheduling. In order to improve the scheduling of the technical operation station, a basic data processing method is provided. Based on the ACP theory, this method collects the physical scheduling data of the actual scheduling system of the technical operation station, and inputs the actual scheduling data into the artificial scheduling system of the technical operation station for computational experiment scheduling. Combined with the optimization model of the technical operation station, the optimal scheme data and operation planning data are obtained. The physical adjustment system is run according to the optimal scheme data and operation plan data, and the parallel execution between the actual scheduling system and the artificial scheduling system is realized through the real-time feedback mechanism. It provides a more efficient and intelligent command and scheduling solution for the actual scheduling system, and improves the intelligence and integrated level of the technical operation station. It strengthens the scheduling and management ability of the railway technical operation station in the face of a large number of complex operation plans, and improves the quality and efficiency of railway freight transportation and production.

1. Introduction

Railway transport is an important mode of transportation in China. "Traffic development, Railway first" has become synonymous with Chinese railway. By the end of 2023, the mileage of China's railways in operation had reached 159000 km, of which 45000 km was high-speed railway, ranking first in the world [1]. The national railway double-track rate and electrification rate reached 59.6% and 73.8% respectively, ranking second and first in the world . With the development of economy, China has begun a new journey of building a modern socialist country in an all-round way, which puts forward higher requirements for the production efficiency of goods transportation.

Technical station refers to the general name of marshalling station and section station. Equipped with relatively perfect shunting equipment and locomotive, vehicles and other equipment. Its main production activities are to supply locomotives and handle the disintegration and marshalling of freight trains for adjacent railway sections, as well as technical operations for arriving disintegrated trains, initial marshalling trains and non-adapted transit trains [2]. How to do well the command and control of the railway technical operation station and improve the transportation scheduling technology is very important to improve the freight transport productivity. At present, the operation of command and scheduling mainly depends on the operator being familiar with the operation plan, strengthening the shift management of key trains, the operation language of receiving and departing trains, filling in the busy lines, strengthening the construction skylight repair operation organization and so on, mainly using train scheduling commands system (TDCS), railway transport scheduling management system (TDMS) and centralized traffic control system (CTC) [3]. In order to solve the problems of complicated technical operation flow, many interference factors and relying on manual scheduling in railway technical operation station, the intelligent parallel scheduling system based on ACP theory has become a new choice.

ACP theory, that is, the organic combination of artificial system, calculation experiment and parallel execution, can effectively solve the analysis and control problems of complex systems. By using the parallel system method, we can easily establish the relationship between the actual system and the artificial system, through a variety of experiments on the artificial system to adjust the management and control strategy of the actual system, accordingly, through the actual system to modify the model of the artificial system, so that it may become the "standby" system of the actual system. The most remarkable feature of parallel execution is to transform the conventional passive computer simulation into a parallel active manual system, and to make its role from passive to active, static to dynamic, offline to online, and finally from subordinate status to equal status, giving full play to the role of artificial system in the management and control of real systems [4]. Based on the multi-agent framework, an artificial system is established to replicate with the actual hub. The artificial hub system is combined with the actual hub to carry out parallel execution. In abnormal cases, the system can be backed up until normal operation is restored, minimizing the loss [5]. Parallel execution can be used to undertake a variety of tasks, which can be independent or related to each other. At present, ACP theory has been successfully applied in many fields, such as transportation, electric power, business administration, transportation and so on. The basic theoretical framework of ACP theory in railway technical operation station is shown in Figure 1:

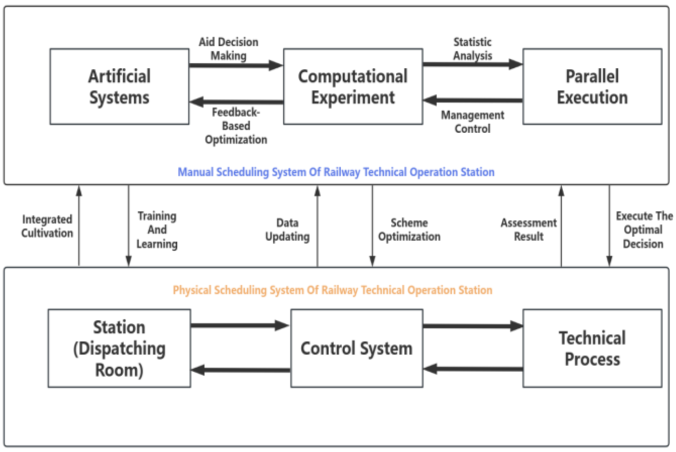

Figure 1 The basic theoretical framework of ACP theory in railway technical operation station

Prior to this, Yang Donghu and other researchers used advanced technologies such as artificial intelligence and ACP to build an intelligent parallel scheduling system for technical operation stations by establishing artificial scheduling systems, computing and experimental modules, and parallel execution modules.

The construction of the system needs to use the basic data of the technical station as the parameter support. This paper focuses on how to use the parallel system theory to compile the data of Xinfengzhen marshalling station.

2. Technical station business process

Technical station refers to the general name of marshalling station and section station. It is equipped with relatively perfect shunting equipment and equipment such as locomotives and vehicles. Its main production activities are to supply locomotives and handle the disintegration and marshalling of freight trains for adjacent railway sections, as well as technical operations for arriving disintegrated trains, initial marshalling trains and non-adapted transit trains.

2.1 Technical Station Business Process

1) Freight train arrival and departure operation;
2) Freight train shunting operation;
3) Freight train maintenance operation;
4) Locomotive locomotive operation.

2.2 Detailed Business Process

Train arrival and disintegration marshalling: when the freight train arrives at the technical station, the duty personnel of the technical station first carry out the pick-up operation to confirm the arrival information of the train, the type of goods and the number of carriages, and then disintegrate the arriving train. Each carriage is classified according to the target place and the type of goods, and temporarily parked in the shunting yard of the technical station.

Make the marshalling plan: the dispatcher collects the train scheduling plan, freight demand and train arrival information, determines the number and sequence of carriages that need to be grouped, and formulates the specific marshalling plan according to the scheduling instructions and the actual situation.

Departure operation: after the completion of the marshalling operation, the technician carries out the technical inspection of the carriage according to the technical inspection standard. After ensuring that the departure conditions are met, the dispatcher of the technical station confirms the completion of the marshalling, fills in the relevant marshalling records, and inputs the marshalling information into the scheduling system, notifies the relevant departments of the marshalling information, and the dispatcher of the technical station issues the train departure command according to the instructions of the CTC system. Coordinate the train driver to prepare for departure, the train leaves on time at the specified time, and run according to the scheduling plan. The technical station will feedback the train departure information to the CTC system to complete the whole marshalling operation process.

It is worth noting that after the completion of the train operation, the departure operation can be arranged directly according to the actual situation and specific scheduling requirements.

2.3 Basic Data Compilation Method

As a station for handling the technical operation of freight trains, the technical station is a relatively complex and huge railway system, with many complex data sets. According to the requirements of the intelligent parallel scheduling system, the method process is divided into five modules, namely, acquisition module, compilation module, simulation module, processing module and operation module.

2.4 Acquisition Module

The acquisition module is used to obtain the physical scheduling data of the physical scheduling system of the technical operation station. In the intelligent parallel scheduling system of railway technical operation station, the stage of data acquisition is a crucial step, which provides original information for subsequent data processing and decision support. The sources can be stations under the jurisdiction of major railway bureaus, transportation universities directly under the Ministry of Railways, or related journals. The main types of information obtained include but are not limited to station information, train information, line information, equipment information and operation plan information.

2.4.1 Station Information

The station information mainly includes the name, geographical location, number of stations, signal equipment, communication equipment configuration information. Take Xinfeng Town marshalling Station as an example, Xinfeng Town marshalling Station is located in Lintong District, Xi'an City, Shaanxi Province. It belongs to China Railway Xi'an Bureau Group Co., Ltd., its east end upstream connects Longhai, Ningxi and Baoxi, and the western end downlink Longhai, Beihuan and Xikang. It is the connecting station of Xinzhang Section of Baoxi Line, Northern Ring Line, Xikang Line, Ningxi Line and Longhai Line. It is an important railway hub connecting Northwest and Southwest China, and it is a large network freight marshalling station. As of October 2018, Xinfeng Town Station Yard adopts two-way, three-level and seven longitudinal station type and main line outsourcing layout, with a total of four train inspection sites, namely, the arrival of the uplink system, the departure yard and the arrival and departure yard of the downlink system; an exchange yard is set up between the up and down marshalling yards; Xinfeng Town locomotive depot is set up between the descending arrival yard and the uplink departure yard; there is a vehicle depot in the west of the station; and a station in the south of the station [6-7]. As of April 2010, there are 48 shunting lines in the upstream yard of Xinfeng Town Station, 12 shunting lanes in the arrival yard of the uplink system, 3 shunting lines reserved in the upstream system shunting yard, 34 shunting lines in the upstream system shunting yard, 15 shunting lanes reserved in the upstream system departure yard, 48 shunting lines in the downlink yard, 12 shunting lanes in the downlink system arrival yard, 33 shunting lines in the downlink system shunting yard, and 16 shunting lines in the downlink system departure yard. There are six exchange lines in the exchange field. As of April 2010, the construction area of work and equipment in Xinfeng Town Station is 78702.29 square meters, with a total of 82 buildings. In addition, Xinfeng Town Station also adopts a new generation of marshalling yard integrated automation (SAM) system, equipped with 11 shunting machines, including 5 uplink systems and 6 downlink systems [8].

2.4.2 Train Information

Common train information includes train type; train formation information; train operation plan and train status. Specifically, the collection of train information at a certain time period at a station mainly includes the train number (both passenger and freight trains have corresponding train numbers); the type of train, such as passenger, freight or special vehicles; the train formation information, such as the number of vehicles, vehicle types, and load capacity (among which, for freight purposes, mainly collect open-top cars, flat cars, covered cars, tank cars and insulated cars, etc.; for passenger purposes, mainly collect hard-seat cars, hard-sleeper cars, soft-sleeper cars, dining cars and luggage cars, etc.); the train operation plan, such as the arrival and departure time of the train, the stop time and the running route; the train status, such as whether it is on time, whether it is delayed and the reason for the delay, and whether there is a fault, etc.

Table 1, 2, 3 shows the station information of Xinfeng marshalling Station.

Table 1 is the overview of upfield data of Xinfeng Town Technical Station:

Table 1 Upfield data of Xinfeng town Station Line information

Designation	Quantity
Upfield Shunting Line	48
Arrive At The Yard Lane	12(reserved 3 lanes)
Shunting Line In The Yard	34(reserved 14 lines)
Departure Lane	15(reserved 3 lanes)

Table 2 shows the basic data of downlink field of Xinfengzhen Station:

Table 2 Data survey of descending field at Xinfeng Town Station

Designation	Quantity
Downfield Shunting Line	48
Arrive At The Yard Lane	12
Shunting Line In The Yard	33
Departure Lane	16
Exchange Field	6

And Table 3 is the overall station information overview of Xinfeng Town Station:

Table 3 Station information of Xinfeng Town Station

Designation	Quantity
Automatic Hump	2
Freight Yard	2
Gantry Crane	1
Station Track	224
Total Length(km)	192
Turnout	663
Annunciator	793
Dowty Retarder	14044
Flip Line	4

Line information includes technical parameters such as railway line name, mileage, speed limit, line occupation status, construction and maintenance information; specifically, collecting line information, mainly collecting railway line name, mileage and speed limit connected by technical stations; line occupation and so on.

2.5 Compilation Module

The compilation module is used for compiling the data based on the physical scheduling data to obtain the database structure of the manual scheduling system of the matching technical operation station.

The data compilation process is shown in Figure 2:

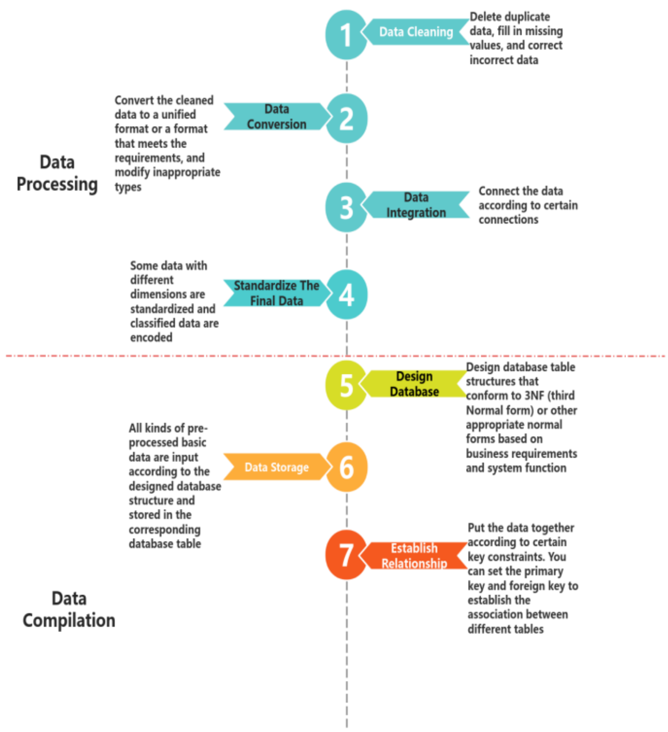

Figure 2 The data compilation process

2.5.1 Data Processing

Data processing involves the following steps:

• Data integration, connecting the data according to a certain degree of connection. Then model the system based on ACP theory; among them, association integration includes associating each independent data set to form a complete business view based on train number, train number, time stamp and other key field.

• The standardization of the final data mainly includes the standardization of some data with different dimensions, such as the unit unification of distance, speed, time and other data; Encode the categorical data, such as using One-hot Encoding or Label Encoding; In addition, it also includes data normalization or standardization, the specific step is to normalize continuous variables (Min-Max Scaling) or standardize (Z-score standardization), so that different indicators are comparable, which is conducive to the training and optimization of machine learning algorithms [9].

Z-score standardization is a commonly used data standardization method, which can compare and analyze values between different data sets. In statistics and data analysis, Z-score standardization can help us standardize data processing, making different data sets have comparability, so as to better data analysis and mining [10].

The mean is 0 and the standard deviation is 1.

For A matrix A:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

We have:

$$b_{ij} = \frac{a_{ij} - a_{\bullet j}}{S_j}, i = 1,2,3,\ldots,m, j = 1,2,3,\ldots,n \ (1)$$

$$\bar{a}_{\bullet j} = \frac{1}{\mathbf{m}}\sum_{\mathbf{i=1}}^{m} a_{ij} \ (2)$$

$$S_j = \sqrt{\frac{1}{\text{m -1}}\sum_{i=1}^{m}\left(a_{ij} - \bar{a}_{\bullet j}\right)^2} \ (3)$$

$$\bar{b}_{\bullet j} = \frac{1}{\mathbf{m}}\sum_{\mathbf{i=1}}^{m} b_{ij} = \frac{1}{\mathbf{m}}\sum_{\mathbf{i=1}}^{m}\frac{a_{ij} - \bar{a}_{\bullet j}}{S_j} = \frac{1}{\mathbf{m}}\frac{\left(\sum_{i=1}^{m} a_{ij}\right) - m\bar{a}_{\bullet j}}{S_j} = 0$$

$$\mathrm{D}_{\mathrm{j}}^2 = \frac{1}{\mathrm{m}-1}\sum_{\mathrm{i=1}}^{m}\left(b_{ij} - \bar{b}_{\bullet j}\right)^2 = \frac{1}{\mathrm{m}-1}\sum_{\mathrm{i=1}}^{m}\left(b_{ij}\right)^2 = \frac{1}{\mathrm{m}-1}\sum_{\mathrm{i=1}}^{m}\left(\frac{a_{ij} - \bar{a}_{\bullet j}}{S_j}\right)^2$$

$$= \frac{1}{S_j^2}\bullet\frac{1}{\text{m -1}}\sum_{\mathrm{i=1}}^{m}\left(a_{ij} - \bar{a}_{\bullet j}\right)^2 = 1$$

As an example, we take the daily disassembly train and daily train formation from 2017 to 2023 in Table 4:

Table 4 Daily disassembly train and daily train formation from 2017 to 2023

Year	Daily disassembly train	Daily train formation
2017	177.3	155.8
2018	178.5	161.5
2019	179.6	162.0
2020	179.4	162.4
2021	181.7	164.2
2022	184.1	179.4
2023	188.0	162.0

Then we organize the above data into a matrix form:

$$A=\begin{pmatrix} 177.3 & 155.8 \\ 178.5 & 161.5 \\ 179.6 & 162.0 \\ 179.4 & 162.4 \\ 181.7 & 164.2 \\ 184.1 & 179.4 \\ 188.0 & 162.0 \end{pmatrix}$$

We can use $\bar{A}_{\cdot \mathrm{j}} = \frac{1}{\mathrm{M}}\sum_{\mathrm{I=1}}^{\mathrm{M}} \mathrm{A}_{\mathrm{Ij}}$ To determine $\bar{A}_{\cdot \mathrm{j}}$:

$\bar{A}_{\cdot 1}$=179.4; $\bar{a}_{\cdot 2}$=163.9

Similarly, we can also determine S_J And B_{Ij}:

S_1=6.69; S_2=7.27

B_{Ij}:

	b_{1j}	b_{2j}
j=1	-0.58	-1.11
j=2	-0.40	-0.33
j=3	-0.24	-0.26
j=4	-0.27	-0.20
j=5	0.07	0.04
j=6	0.43	2.13
j=7	1.01	-0.26

And then we can produce the normalized matrix B:

$$B=\begin{pmatrix} -0.58 & -1.11 \\ -0.40 & -0.33 \\ -0.24 & -0.26 \\ -0.27 & -0.20 \\ 0.07 & 0.04 \\ 0.43 & 2.13 \\ 1.01 & -0.26 \end{pmatrix}$$

Now we're going to solve it in python from Figure 3:

We can conclude that the mean of both the horizontal and vertical coordinates is 0, and the variance of these data on the projection is 1.

Through the above data collation and preprocessing steps, the original disorganized and uneven data can be transformed into high-quality basic data suitable for intelligent parallel scheduling system.

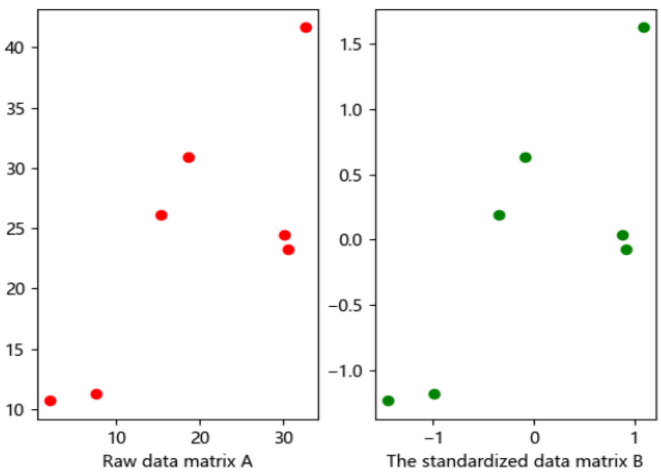

Figure 3 Results in Python

2.5.2 Data Compilation

To compile the sorted and preprocessed data, the main task is to integrate the sorted and preprocessed data and build an initial database structure suitable for the system based on ACP theory, including the following steps:

• According to the embodiment of the station, it is necessary to design the database table structure that conforms to 3nf (third normal form) or other appropriate paradigm according to the business requirements and system functions, including but not limited to station information table, train information table, line information table, equipment and facilities table, operation plan table, etc.

- Data is first inputted and stored. Subsequently, the processed data is inserted into the database in strict accordance with predefined standards. More specifically, various preprocessed basic data elements are fed into their respective tables, following the designed structural schema, and then stored therein.
- Database administrators simultaneously maintain data consistency, completeness, and accuracy, which is of utmost importance. For instance, they employ appropriate operational mechanisms to safeguard the security of data updates. This ensures that any modifications to the data do not compromise its integrity, and that the database remains a reliable repository of information for further analysis and research.
- Database administrators establish the connection of data association and integrate the data according to certain key constraints. They establish the association relationship between different tables by setting primary keys and foreign keys—for example, between train and train times, station and line, or equipment and maintenance records. Additionally, they create views or intermediate tables to simplify complex queries and support data presentation and analysis requirements for specific business scenarios.

2.6 Simulation Module

The manual scheduling system of technical operation station is based on ACP theory, introduces the physical data, models the data, mechanism and experience through expert experience and multi-module modeling method, and constructs a manual scheduling system that simulates the technical operation process in the actual scheduling system. That is, manual scheduling system (Artificial scheduling System). With the help of parallel execution and the dynamic interaction mechanism between the physical world and ADS, the ADS is updated and calibrated in order to update the corresponding operation plan in time when the actual situation changes, and reduce the difficulty of dispatcher scheduling.

2.7 Operation Module

The operation module is used for running the physical scheduling system of the technical operation station based on the optimal scheme data and the operation plan data. With the help of machine learning, deep learning, multi-agent and other technologies, the behavior rules and interaction mechanism of each experimental object in the artificial system are modeled, and through the integration and training of the feedback and evaluation of the actual scheduling system, the ADS is generated, which can deduce and simulate various operating conditions of the actual system. Compared with the actual scheduling system, the execution, interaction and deduction of technical operations in ADS are more complete, which makes up for the shortcomings of the actual scheduling system in extreme environment data and abnormal scene data, and provides indispensable support for the optimization of the actual system [11].

3. Data Update and Maintenance

There is a real-time feedback mechanism between the technical operation station manual scheduling system and the technical operation station physical scheduling system, and the real-time feedback mechanism is established based on cloud edge cooperation and model adaptive matching technology. After inputting the actual physical data to the manual scheduling system, the manual scheduling system can simulate the corresponding physical environment, deduce it in the system, and feedback the results to the actual system in real time, thus forming a situation of manual-actual interaction. In order to make the manual scheduling system more accurate, timely and real, it is necessary to constantly obtain brand-new, dynamic and real-time data from the actual system. And then we should continue to repeat the steps of obtaining-compiling-simulating-processing-running, and input data to the manual scheduling system through continuous in-depth learning. The system promotes the simulated environment with richer data, more complex scenarios, and additional constraints, and continuously feeds the results back to the actual system.

4. Conclusion

The parallel scheduling system built based on the basic data compilation method of ACP theory can meet the daily scheduling requirements of railway technical operation stations after verification. The manual system of the parallel scheduling system needs to use the actual physical data collected as data support, and in turn, the simulated data simulated by the artificial system will feed back to the actual system at the same time. So that the artificial system constantly makes dynamic adjustments, so that the simulated data is closer and closer to the real data, and finally predicts the future data trend. Just like the two people who were unrelated before, after constant testing, running-in, and adjustment, the two people become more and more tacit understanding, and finally reach the realm of "two people together". This method also provides a more efficient and intelligent command and dispatch solution for the actual dispatch system, improves the intelligence and integration level of the technical operation station, enhances the dispatch and management ability of the railway technical operation station in the face of a large number of complex operation plans, and improves the quality and efficiency of railway cargo transportation and production.

Acknowledgements

This work was supported in part by China State Railway Group Co., Ltd. (China Railway) under R&D project L2022x002; Chongqing Transportation Commission 2024 Transportation Science and Technology Project Guide 4 CQJT-CZKJ2024-04; Beijing Natural Science Foundation under grant L233005.

References

[1] X Yang, Analysis of Railway Operating Mileage, Railway Network Density, Duplication Rate, and Electrification Rate in China in 2023. Market Insights,2023.
[2] G Ma, Dictionary of Communication, Shanghai Jiaotong University Press, 2005.01.
[3] Y Li, Research and implementation of CTC3.0 station subsystem, Lanzhou Jiaotong University, 2018, DOI: 10. 27205/d. cnki. gltec.2018.000167.
[4] S Lun, Research on the classification of parallel execution modes of ACP theory, Acta Automatica Sinica, 2012, 38(10):1602-1608.
[5] G. Xiong, W. Zhang, X. Liu, Z. Yuan, F. Zhu and Z. Zhang, Framework and Key Technologies for High-speed Railway Comprehensive Transport Hub, 2023 IEEE 3rd International Conference on Digital Twins and Parallel Intelligence (DTPI), Orlando, FL, USA, 2023, pp. 4-5.
[6] Z Huang. Discussion on breaking through the transportation bottleneck of Xinfeng Town marshalling Station. Shandong Industrial Technology, 2018, (10):200-201. DOI: 10.16640/j.cnki.37-1222/t.2018.10.181.
[7] W Peng. Study on layout scheme of Xinfeng Town marshalling Station reconstruction and expansion. Journal of Railway Engineering Society, 2010, 27(4):112-116. DOI: 10.3969/j.issn.1006-2106.2010.04.024.
[8] J Zhao. Discussion on optimization countermeasures of transportation organization in Xinfeng Town marshalling Station. Railway Transport and Economy, 2019, 41(5):89-93. DOI: 10.16668/j.cnki.issn.1003-1421.2019.05.17.
[9] N Chen. SVM prediction of railway freight volume based on data processing. Shijiazhuang Tiedao University, 2019. DOI: 10.27334/d.cnki.gstdy.2019.000138.
[10] M. Islam, K. N. Khan and M. S. Khan, Evaluation of Preprocessing Techniques for U-Net Based Automated Liver

Segmentation, 2021 International Conference on Artificial Intelligence (ICAI), Islamabad, Pakistan, 2021, pp. 187-192, doi:10.1109/ICAI52203.2021.9445204.
[11] G Xiong, D Yang, W Xu,R Li, S Chen, B Song, X Dong, F Zhu, Architecture and Key Technologies of Parallel scheduling System for Railway Technical Operation Stations, 2023 International Annual Conference on Complex Systems and Intelligent Science (CSIS-IAC), Shenzhen, China, 2023, pp.1-2