参赛队员姓名(Name of contestant): Ge, Jingyi

中学(Middle School): The High School Affiliated to Renmin University of China

省份(Province): Beijing

国家/地区(Country/Region): China

指导教师姓名(Name of tutor): Zhang,Qi

指导教师单位(Tutor unit): China Academy of Railway Sciences Corporation Limited, Signal and Communication Research Institute

论文题目(Paper Title)： A Delay Prediction Algorithm based on Train Time-efficient Driving Strategy and Critical Node Identification

A Delay Prediction Algorithm based on Train Time-efficient Driving Strategy and Critical Node Identification

Ge Jingyi

The High School Affiliated to Renmin University of China

Abstract

Chinese high-speed railway has entered a phase of large-scale networked operation. It travels across the area with complex environment, e.g., cold, wet and hot, strong wind, earthquake zone, etc. The railway network is full of nodes, and external disturbances or disruptions will likely lead to different degrees of train delay and propagation, which significantly influences high-speed railway operations and passenger services. Thus, it is urgent to solve the problem with scientific and effective methods. Based on the above issues, this paper firstly studies the railway network with massive nodes and multilayer nested features using the complex network method. The operation relationship within and between trains is analyzed. A multilayer network in the train timetable is established. A node centrality measurement algorithm of the improved PageRank is proposed to identify the key disturbance node of delay propagation in the train timetable network. Then, different scenarios of the arrival delay or departure delay are installed at the critical disturbance node. With the Spatio-temporal characteristics of train operation and dynamics, the kinematic parameters of train operation are calculated, such as acceleration, speed, and passing moment. A delay prediction algorithm based on train time-efficient driving strategy and critical node identification is proposed to predict when trains arrive at the subsequent stations and acceleration and deceleration in subsequent intervals, which provides a reliable basis for rescheduling strategies. Finally, simulation experiments are conducted on the Beijing-Shanghai high-speed railway to verify that the improved PageRank algorithm can effectively identify the critical disturbance node in the delay propagation of the railway network. The proposed delay prediction algorithm can predict the delayed recovery of trains at the subsequent stations for dispatchers.

Keywords: Railway Timetable, Delay Propagation, Complex Network, PageRank, Delay Prediction, Train Dynamics

Content

1 Introduction

High-speed railway (HSR) is an important driving force to promote economic and social development, and has become the main way for passengers to choose medium and long-distance travel in China [1]. Nowadays, China has ranked the highest in the world in terms of road network complexity, running speed, driving density, working conditions and passenger transport volume.

However, the HSR network is full of massive station nodes. If the train is delayed at a node of an intermediate station or a running interval due to disturbance, the delay will quickly spread to the adjacent trains and station nodes. This result shows domino effect of delay propagation, resulting in the delay of this line and cross-line trains, and even more serious train delay and line paralysis. For example, the Beijing-Shanghai HSR with 350 km/h speed has 426 trains running every day. The minimum train tracking time is less than 4 min. Under such high-density driving conditions, the line operation capacity has reached saturation. Train delays are easy to occur and causes great economic losses. For example, at 23:04 on August 12, 2018, due to the impact of high winds on colored steel plates, G40 train broke down near Wanzhuang Station and stopped urgently. This caused more than 40 Beijing-Shanghai high-speed trains to be delayed, some trains were delayed for as much as 5 hours. A large number of passengers were stranded at stations along the HSR way. It has become a practical problem to be solved urgently in the HSR field to restrain or reduce the degree of delay generation or transmission by effective means. "Revealing the mechanism of train delay propagation" has been listed as a scientific problem to be solved urgently by the Ministry of Science and Technology and other ministries and commissions of China. The research of train delay prediction method based on critical node identification of HSR network is an important basis to solve this problem.

At present, China's HSR dispatching command system adopts the organization mode of "driving according to the map", that is, supervising and commanding the safe and effective of trains according to a pre-established daily plan. Train timetable is the representation mode of train operation plan, which stipulates the arrival and departure times, departure sequences and routes at various stations along the way. It is the formal representation of HSR network and the abstraction of train operation space-time track, and contains discrete space-time nodes, structural characteristics of road network, dynamic and static traffic flow characteristics and delay propagation mechanism. The delay propagation network in the train timetable is established and evaluated. The critical disturbance node of train delay propagation is identified and the temporal and spatial characteristics of train operation are analyzed. The arrival time and train delay recovery at the subsequent stations are predicted. The train rescheduling strategy that can suppress delay propagation is generated, which is of great significance to eliminate train delay propagation and restore the train normal order. Some studies show that increasing the buffer time of trains in critical sections and stations to reduce and eliminate delays of compiling the train timetable has a significant role in identifying the critical disturbance node for train delay propagation and improving the robustness and stability in the train network [2-4]. However, the above research only evaluates the current nodes, ignoring the correlation and coupling characteristics of different nodes in terms of absorption delay.

HSR has the multi-layer coupling network characteristics of "station-line-road network". Complex network theory and delay time prediction method have been widely studied in the field of HSR [1, 5-6]. Some scholars apply different centrality measurement methods to search for critical nodes of the train timetable [7-9]. Based on various measures and performance indicators, the coupling relationship between stations and interval is analyzed to realize the identification of the critical disturbance node in the HSR network [10-11]. The delay prediction method is used to analyze the whole life cycle of train delay from generation, propagation, reduction to elimination, which is divided into data-driven method and model prediction method. The former method studies the coupling relationship between discrete Spatio-temporal characteristics of train operation and before the equipment is closed by deep learning, Bayesian network, random forest and other machine learning methods. This method predicts the correlation relationship between initial delay, secondary delay and buffer time, and predicts the Spatio-temporal evolution information of trains at the subsequent stations and intervals for train dispatchers. Based on the train dynamics model and kinematics characteristics, the latter method analyzes the evolution trend among train kinematics parameters, delay and dynamic performance index of train timetable. This method realizes microscopic prediction of train delay change. Compared with the data-driven method, the model-driven method has higher prediction accuracy and stronger feasibility. However, the existing research mainly analyzes the topological characteristics of HSR static network based on complex network, but has not combined the dynamic coupling network of railway network with the delay prediction method. The research seldom analyzes the strong correlation and dynamic time-varying characteristics between delay propagation network and train stage plan. It is difficult to accurately and reasonably identify the critical disturbance node of delay propagation in the train timetable. It is also difficult to block the secondary propagation of train delay in advance, thus affecting the recovery of train operation order as soon as possible.

In this paper, the coupling relationship is analyzed in the train timetable among trains, lines and networks. Multi-layer network model is established. An improved PageRank node center measure algorithm is developed to identify the key disturbance node in the Spatio-temporal network of the train timetable under disturbances or disruptions. Then, the scenarios of arrival delay and departure delay are installed in the critical disturbance node. The kinematic parameters such as train running acceleration, speed and passing time are calculated considering the train traction and braking characteristics. The delay prediction algorithm is proposed, and the delay change and train recovery analyzed accordingly. Finally, taking the Beijing-Shanghai HSR as an example, the simulation experiment is executed to verify the effectiveness and feasibility about the delay prediction algorithm based on train time-efficient driving strategy and critical node identification.

2 Algorithm

2.1 Identification algorithm of the critical disturbance node in the train timetable

When the HSR network is affected by emergencies, the train delay spreads rapidly among the massive coupled station resource nodes. Real-time and accurate identification algorithm of the critical disturbance node in the train timetable is the critical basis for predicting train delay and suppressing delay propagation.

2.1.1 Train timetable network modeling

The train timetable network is constructed as a multi-layer network with N nodes, M layers and K edges. It is expressed as a directed graph $G = (V, E, W)$. The connection relation of graph G is described as adjacency matrix $A = [a_{ij}]_{N \times N}$. The element a_{ij} equals 1 indicates that the current node i points to the node j. The nodes can be expressed as the arrival, departure and passing operations of trains at stations. A layer in the network represents a train. The edges between nodes of the same layer are used to describe the interval running additional time or stop additional time of a train. The edges between nodes in different layers represent the buffer time along HSR lines. Additional and buffer times are the extra time set to improve the robustness and stability of the train timetable, which are the main factors affecting the delay propagation. Therefore, the additional time of trains in intervals and stations is expressed by the connection of nodes in different stations. The buffer time in intervals and stations is described by establishing the link relationship between the single network layer and multi-layer network.

In order to explain and study the problem, the variables are listed in Table 1.

Table 1 Symbolic description

Variable	Description	Unit
N	Number of nodes	-
J	Number of stations	-
V	Node set	-
E	Edge set	-
W	Edge weight set	-
w_{ij}	Edge weight between node i and node j	-
o_{ij}^{s}	Arrival time of train i at station j	-
o_{ij}^{e}	Departure time of train i at station j	-
d_{ij}	Minimum dwell time of train i at station j	min
r_{ik}	Minimum running time of train i in interval k	min
H_k	Minimum tracking interval between two successive trains in interval k in the same direction	min
G	Number of trains	-
I	Number of stations	-
J	Position index of the front dwell station	m
g	Train number, $g \in \{1, 2, \cdots, G\}$	-
i	Station number, $i \in \{1, 2, \cdots, I\}$	-

j	Position index, $j \in \{1, 2, \cdots, J\}$	-
F	Train traction force	N
F_{max}	Train maximum traction force	N
B	Train braking force	N
B_{max}	Train maximum braking force	N
W	Train running resistance	N
W_0	Train basic resistance	N
n_1, n_2	Coefficient of train operation conditions	-
b_1, b_2, b_3	Constants related to the train basic resistance	-
F_c	Train resultant force	N
$a_{g,j}$	Acceleration of train g at position point j	m/s^3
$v_{g,j}$	Speed of train g at position point j	m/s^2
$t_{g,j}$	Passing time of train g at position point j	s
$v^k_{g,j}$	Speed restriction value of train g in the k th temporary speed restriction section	m/s

A train timetable network is established with a three-layer network (three trains), as shown in Fig. 1. Similar to the train timetable, nodes with equal heights indicate that trains are at the same station. The first and last node of the same level indicate the departure and arrival station of the train. Among the other nodes, provided that the station only has one node, it means that a train passes through the current station. If the station has two nodes, they respectively indicate the train departs and arrives at the current station.

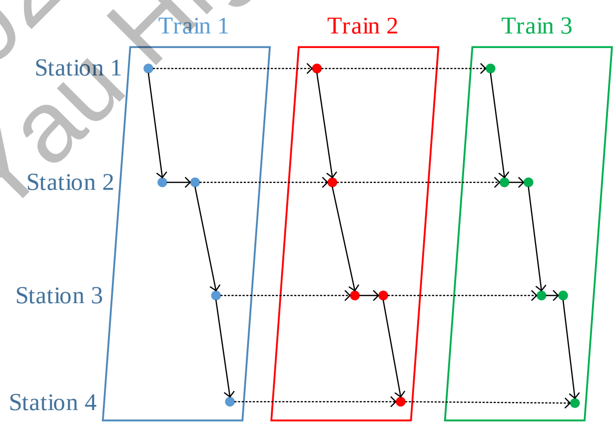

Fig. 1　A three-layer train timetable network. A layer represents a train. A node is a train operation. Each side represents additional time (edge within the layer, solid line) or buffer time (edge between layers, dashed line)

The train timetable network is established in the following two steps:

Step 1: Establish nodes and edges for each layer. A node is defined as a triple (arrival and departure times, station and train). Each layer represents a train in the train timetable, because the first node shows that the train departs from the current station. Fig. 2 shows a formal representation of the nodes and edges of train 1 from Station J - 1 to Station $J+1$. The train operation of receiving, departing, and passing are listed in Fig.2(a) and Fig.2(b) respectively. The edge includes two types: interval additional time and dwell additional time.

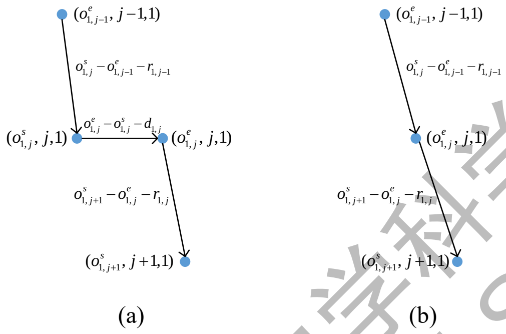

Fig. 2 Construct nodes and edges in a certain layer in the network. The types of train operations are: (a) receiving and departing operations; (b) Passing operation

Step 2: Create edges for different layers. The network layers are established for all trains. The nodes between different layers are connected to analyze the tracking situation of different trains in the train timetable. According to the different operation conditions of trains at stations, there are four kinds of connection conditions of inter-layer nodes, as shown in Fig. 3.

1. Pass-pass: If both trains pass through station j, the interval of tracking trains for arrival and departure operation is the same, so the minimum buffer time will be set as the edge weight. The calculation formula is $o_{2,j}^{s}-o_{1,j}^{s}-\max(H_{j-1},H_{j})$. When the train passes through, $o_{1,j}^{s}=o_{1,j}^{e}$.

2. Pass-arrival: When train node 1 is passing at station j and train node 2 is arrival and departure operation, the tracking interval corresponding to train arrival will be less than the tracking interval of train departure. Therefore, the tracking interval corresponding to train arrival is installed as edge weight. The calculation formula is $o_{2,j}^{s}-o_{1,j}^{s}-H_{j-1}$.

3. Departure-pass: When train node 1 is arrival and departure operation at station j and train node 2 is passing operation, the tracking interval corresponding to train

arrival will be longer than the tracking interval of train departure. Therefore, the tracking interval time corresponding to train departure is noted as edge weight. The calculation formula is $o_{2,j}^{e} - o_{1,j}^{e} - H_j$.

4. Arrival-arrival and departure-departure: When the two nodes are both arrival and departing operations at station *j*, the tracking intervals corresponding to train arrival and departure are noted as edge weights, which are $o_{2,j}^{s} - o_{1,j}^{s} - H_{j-1}$ and $o_{2,j}^{e} - o_{1,j}^{e} - H_j$ respectively.

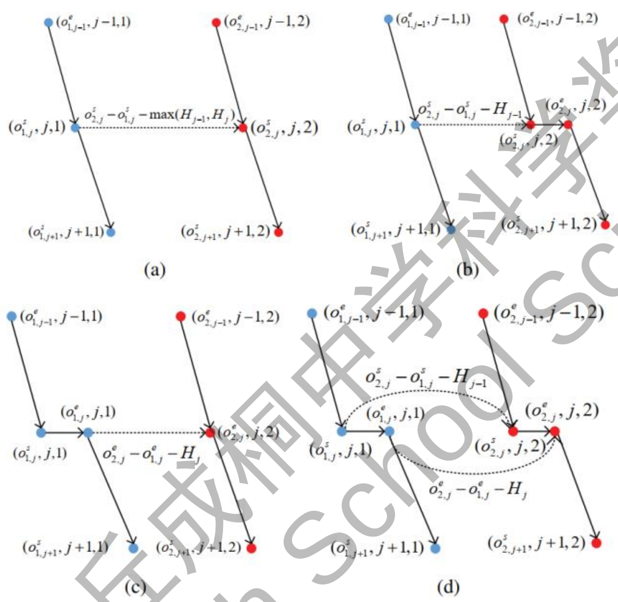

Fig. 3 Train 1 and train 2 connect nodes between different layers at Station *J*. The types of train operations are: (a) pass-arrival; (b) departure-pass; (c) departure-pass; (d) arrival-arrival and departure-departure

2.1.2 Efficient measure for complex network of train timetable--an improved PageRank algorithm

In the previous research, a new effective measurement method of complex network in the train timetable, improved PageRank algorithm [11], has been proposed and verified. This algorithm is used to evaluate the influence degree of station nodes in the process of delay propagation in the train timetable network. It is also used to output the station nodes with the greatest disturbance for the delay prediction algorithm.

The core thought of the algorithm is that the importance of nodes is not only related to the neighbors (the degree of nodes), but also related to the importance of neighbors. To identify the critical disturbance node in the train timetable network, it is necessary to make clear how the edges and nodes influence HSR management. Because the edge weight is determined by the additional time or buffer time that is suppressing train delays, the higher the weight, the stronger the delay absorption capacity. With the influence of initial delay, the train may have associated delay at the subsequent stations and sections, so the importance of nodes is only related to the importance of the subsequent nodes and the edge weight pointing to subsequent nodes. There is a positive correlation between the value of the subsequent and current node. A negative correlation between the value of the edge (i.e., the outgoing edge) in the current node is pointed to the subsequent neighbor node. The improved PageRank index IPR_i(k) of node i is related as follows:

$$IPR_i(k) = d\sum_{j=1}^{N} a_{ij} e^{-w_{ij}} IPR_j(k-1) + (1-d)\frac{1}{N} \quad (1)$$

where w_{ij} is the edge weight for node i and j, and $e^{-w_{ij}}$ is the natural index item of the edge weight. The variable d is the attenuation factor. The improved PageRank algorithm for identification of the critical disturbance node is shown in **Algorithm 1** [11].

Algorithm 1 Improved PageRank for identification of the critical disturbance node [11]

Input: Train timetable network diagram $G = (V, E, W)$, the node set V, the edge set E, the weight set W, the network size N, the Adjacency Matrix **A**; the weight Matrix **W**; attenuation factor d

Output: Improved PageRank value **IPR**

Set the **IPR** of nodes in the whole network to $1/N$;
for k = 1 to 100 **do**
    t = **IPR**;
    for i = 1 to N **do**
        $IPR_i = d\sum_{j=1}^{N} a_{ij} e^{-w_{ij}} IPR_j + (1-d)\frac{1}{N}$;
    end for
    if $\|\mathbf{t} - \mathbf{IPR}\|_\infty < 1e^{-4}$ **then**
        break;
    end if
end for
return

2.2 Delay prediction algorithm

According to the identification algorithm, the critical disturbance node in the train timetable is determined. The scenario of arrival and departure delay is installed at the station node. The delay prediction algorithm is proposed to predict the arrival time and train delay recovery at the subsequent stations, and assist dispatchers to adjust the train timetable.

2.2.1 Train dynamics modeling

Based on the train movement process in the interval, the train delay at the following station is calculated by using the thought of model prediction. Considering the running process of trains in stations and intervals, the train force is analyzed based on the train dynamic characteristics. When the train runs on a straight track, it is affected by gravity, train traction force F, train braking force B and train running resistance W, as shown in Fig. 4.

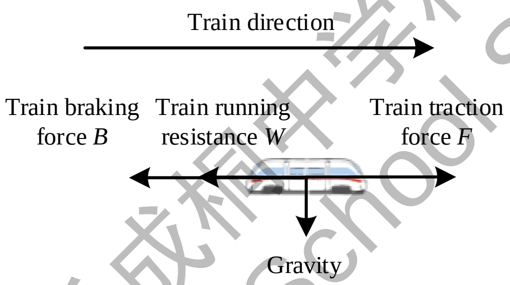

Fig. 4 Schematic diagram of train force

The traction force F and braking force B can be obtained from traction or braking characteristic curves according to linear interpolation method or curve fitting method. For the train running resistance W, the basic train running resistance W_0 is mainly considered. It can be calculated according to the empirical formula applied in railway engineering [12], that is

$$W_0 = b_1 + b_2 \cdot v_{g,j} + b_3 \cdot \left(v_{g,j}\right)^2 \tag{2}$$

where the constants b_1, b_2, b_3 are related to the train type [12].

According to the different resultant force of the train, the train running conditions are divided into traction, cruising, coast and braking. The formula of calculating Fc under different working conditions is as follows

$$F_{合} = \begin{cases} F - W_0 - W_1, \text{traction} \\ 0, \text{cruise} \\ -W_0 - W_1, \text{coast} \\ -B - W_0 - W_1, \text{braking} \end{cases} \tag{3}$$

where braking needs to pass through coast to transition to traction and cruise [12].

Considering the dynamic characteristics and kinematics equations of the train, F_c is assumed as constant in the same distance interval Δj. The acceleration, velocity and passing time of the train at each position are updated and calculated by using the distance step method. Then, the train arrival time at the subsequent station is predicted. The train kinematics parameters are computed is as follows.

(1) Acceleration

$$a_{g,j} = \min\left\{\frac{F_c}{m}, a_{\max}, \delta_{\max} \cdot \Delta t_{g,j-1,j} + a_{g,j-1}\right\} \tag{4}$$

where the acceleration $a_{g,j}$ of train g at position j is limited by the maximum acceleration $a_{\max}$. The maximum impact rate $\delta_{\max}$ is to ensure passenger comfort.

(2) Speed

$$v_{g,j} = \min\left\{v_{g,j}^{k}, \sqrt{\left(v_{g,j-1}\right)^2 + 2 \cdot a_{g,j} \cdot \Delta j}\right\} \tag{5}$$

Where the speed $v_{g,j}$ of train g at position point j is constrained by the speed restriction value $v_{g,j}^{k}$ of the k-th temporary speed restriction (TSR) section. When the train speed under traction condition is greater than $v_{g,j}^{k}$, only part of the traction force is applied to ensure that the current train speed is equal to $v_{g,j}^{k}$.

(3) Passing time

$$t_{g,j} = \Delta t_{g,j-1,j} = \frac{v_{g,j} - v_{g,j-1}}{a_{g,j}} \tag{6}$$

where the passing time $t_{g,j}$ of train g at any position j is approximately equal to the running time $\Delta t_{g,j-1,j}$ passing through the two consecutive position points $j-1$ and j.

(4) Predicted arrival time

$$t_{g,i} = \Delta t_{g,j_i-\Delta j} + \Delta t_{g,j_i-\Delta j,j_i} = \Delta t_{g,j_i-\Delta j} + \frac{v_{g,j_i} - v_{g,j_i-\Delta j}}{a_{g,j_i}} \tag{7}$$

Where $t_{g,i}$ is the predicted arrival time of train g at station i. The speed $v_{g,j_i-\Delta j}$ and v_{g,j_i} of train g at position $j_i-\Delta j$ and j_i can be obtained from the predicted target speed curve. The running acceleration a_{g,j_i} is calculated by Equation (5). The variable $t_{g,i}$ is equal to the margin between the predicted arrival time $t_{g,i}$ and the planned arrival time.

2.2.2 Delay prediction algorithm

Because the optimization goal of high-speed railway company is the shortest train total delay, an optimization algorithm of the train time-efficient driving strategy is proposed by combining train kinematics model and kinematics parameter calculation formula. The scenarios of the arrival delay and departure delay are set at the critical disturbance node. The arrival time, delay and its recovery of trains at the subsequent stations at the critical disturbance node are predicted based on the train time-efficient driving strategy.

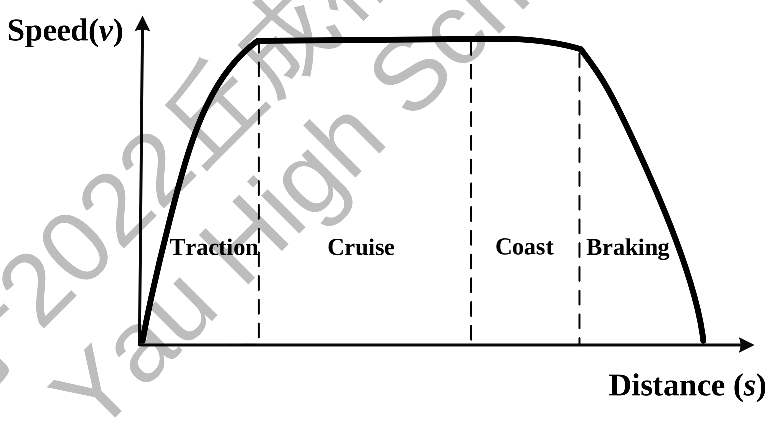

Fig. 5 Running conditions under the train time-efficient driving strategy

The train driving strategy directly affects the accuracy and efficiency of delay prediction. Note that the goal of this paper is to calculate the driving strategy of "rushing to run" after the critical disturbance node and predict the delay recovery of trains at the subsequent stations. So the train time-efficient driving strategy is adopted under the requirement of the shortest train total delay. Researchers have proved theoretically that the train time-efficient driving strategy consists of the running conditions of "Maximum traction-Cruise-Coast-Maximum braking" [13], as shown in Fig. 5. According to the difference of TSR in station and interval, the train running section is divided into different speed restriction sections, including in-station speed restriction section, unlimited speed restriction section and TSR section. The schematic diagram and flow chart of the delay prediction algorithm based on train time-efficient driving strategy and critical node identification are listed in Fig. 6 and Fig. 7 respectively. The calculation process are described as the following steps.

Step 1: Calculated the critical disturbance node in the HSR dispatching command system based on the critical disturbance node identification algorithm in **Section 2.1**. The departure delay scenario is installed in the critical station node. The arrival delay scenario is set under TSR in a certain section after the critical station node.

Step 2: The on-board equipment safety computer calculates the train time-efficient driving strategy based on the train dynamics model of **Section 2.2.1**. The arrival time and delay recovery are computed of the train at the subsequent stations at the critical disturbance node.

Step 2-1: Starting from the left boundary point of each TSR section, calculate the train speed at each position under the "maximum traction-cruise" curve from left to right by applying **Equations (3) to (5) of Section 2.2.1**. If the train running speed under the maximum traction curve is greater than the TSR value in the current section, only a part of traction force is applied to make the train speed at the next position equal to the TSR value.

Step 2-2: Starting from the right boundary point of each speed restriction section, apply **Equations (3) to (5) of Section 2.2.1** to solve the train running speed at each position under the "maximum braking-cruising curve" from right to left. If the train running speed under the maximum braking curve is greater than the TSR value in the current section, only a part of braking force is applied to make the train speed at the next position equal to the TSR value.

Steps 2-3: The train actual speed using the delay prediction method is equal to the minimum value under the maximum traction curve in **Step 1** and the maximum braking curve in **Step 2**.

Step 2-4: Calculate the acceleration, speed, passing time and predicted the train arrival time according to **Equations (4) to (7) of section 2.2.1**. Predict the train delay and its recovery at the subsequent stations.

Step 3: The train delay recovery at the subsequent stations in **Step 2** is sent to the HSR dispatching command system to assist dispatchers to adjust the train timetable.

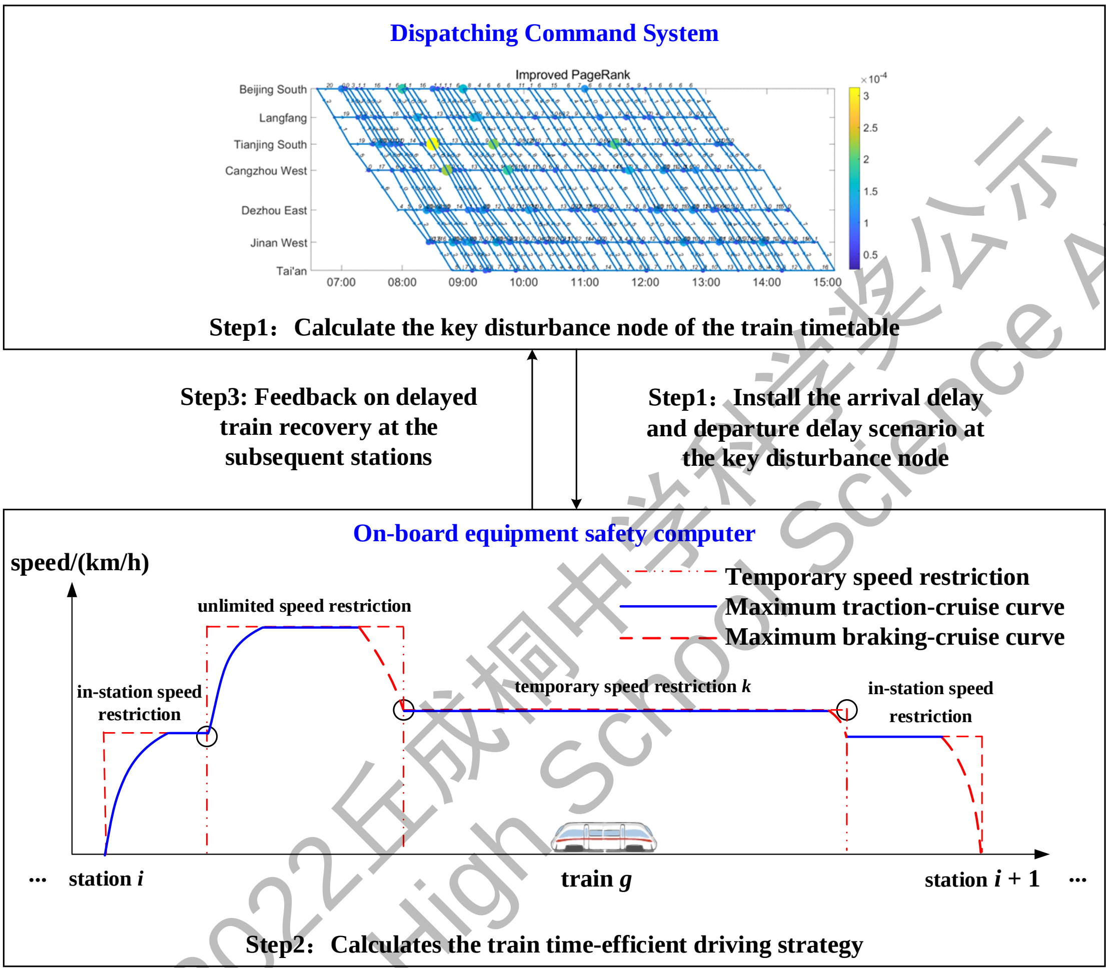

Fig. 6 Schematic diagram of the delay prediction algorithm

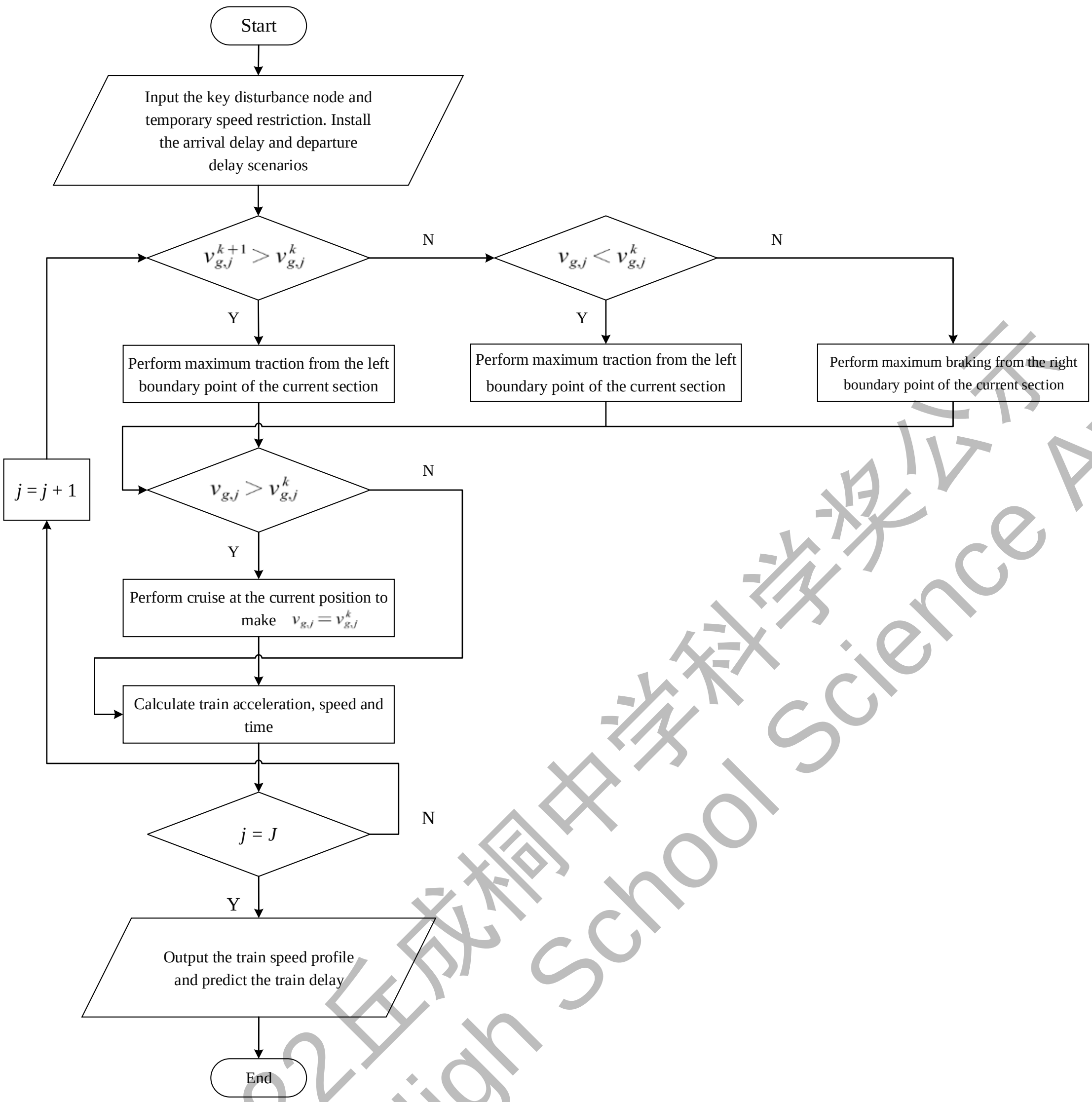

Fig. 7 Flow chart of the delay prediction algorithm

The delay prediction algorithm proposed is written in C + +. The simulation environment is Visual Studio 2019. The code interface of the delay prediction algorithm is shown in Fig. 8. It is divided into four code blocks, which are:

(1) Train traction force is obtained according to the real-time train speed-traction curve, in which traction energy consumption is generated.

(2) The train braking force is acquired according to the real-time train speed-braking curve, in which traction energy consumption is not produced.

(3) Calculate the train basic resistance according to the real-time speed.

(4) Predict the train delay under the train time-efficient strategy.

```
/*Code block 1: Calculate the train traction force*/
double CTrain::getTractionForce(double speed){ ... }

/*Code block 2: Calculate the train braking force*/
double CTrain::getBrakingForce(double speed){ ... }

/*Code block 3: Calculate the train basic resistance */
double CTrain::getBasicResistance(double speed){ ... }

/*Code block 4: Predict the train delay based on the time-efficient strategy*/
void CTrain::calMinTmSpeed2(int* lmtPos, int* lmtSpeed, int lmtNum, int* block_length, int* sectionIndex){ ... }
```

Fig. 8 Code interface of the delay prediction algorithm

3 Simulation experiment

To verify the feasibility and effectiveness of the proposed delay prediction algorithm, the planned train timetable of Beijing-Shanghai HSR is considered for simulation experiment. The critical disturbance node in the delay propagation process are identified. The arrival and departure scenarios are installed at the station nodes. The delay recovery of trains at the subsequent stations is predicted based on the train time-efficient driving strategy.

3.1 Scenarios setup

This paper takes Beijing South to Tai'an of Beijing-Shanghai HSR as an example and considers 40 trains of down direction. The train departure time is from 6:30 to 13:00. The corresponding train timetable is drawn in Fig. 9. The minimum dwell time and train tracking interval are set to 2 min and 4 min, respectively. The minimum running time between stations is listed in Table 2. When the distance step Δj is set to 10 m, the calculation efficiency and accuracy of train kinematics parameters can be guaranteed at the same time.

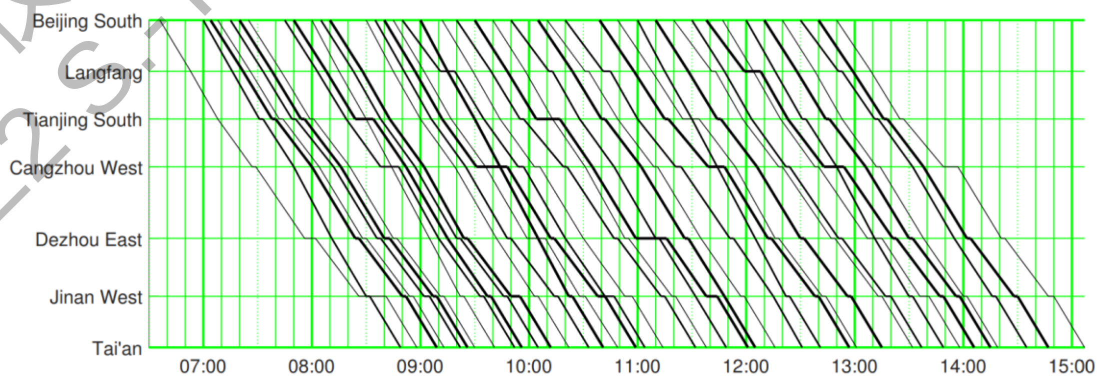

Figure 9 The planned timetable of Beijing-Shanghai HSR, with a total of 40 down

direction trains

Table 2 Minimum running time between two stations

Index	Interval	Time (minutes)
1	Beijing South-Langfang	15
2	Langfang-Tianjin South	14
3	Tianjin South-Cangzhou West	14
4	Cangzhou West-Dezhou East	21
5	Dezhou East-Jinan West	17
6	Jinan West-Tai'an	15

The train model of simulation experiment is CR400AF. The related parameters such as mass, maximum running speed and maximum acceleration are shown in Table 3. There are three scenarios of arrival delay under TSR from 171 km to 291 km behind Tianjin South Station of Beijing-Shanghai HSR, with the TSR value of 300, 200 and 120 respectively (in km/h).

The identification algorithm of the critical disturbance node in the train timetable is carried out on Intel Core i5-8265U CPU 1.60 GHz, 8 GB memory, Windows 10, and simulated in Matlab R2018b. The delay prediction model and algorithm are compiled based on C + + language, and simulated on Intel Xeon Gold 5218 CPU @ 2.30 GHz, 32.0 GB RAM computer.

Table 3 Train parameters

Parameter	Value	Unit
Quality	501000	kg
Maximum speed	350	km/h
Maximum acceleration	1	m/s^2
Maximum impact rate	0.5	m/s^3
Constant b_1	0.399	N/kg
Constant b_2	0.00127	N s/m kg
Constant b_3	0.0001092	N s^2/m^2 kg

3.2 Result analysis

3.2.1 Identification of the critical disturbance node in train timetable network

According to the previous research [11], there is a positive correlation between the improved PageRank and the node delay for random arrival delay or departure delay of 1 to 10 min. The improved PageRank values of nodes in the HSR network and the train total delay can be seen in Fig. 10 and Fig. 11, where the colors in the figure and the size of nodes represent the size of the values. As can be seen from Fig. 10 and Fig. 11, the node (8: 30, 3, 10) corresponds to the node where the 10th train passes at 8:30 at Tianjin South Station (the third station). The nodes (8: 44, 4, 10), (9: 30, 3, 18) and (11: 30, 3, 29) also have great influence on train delay, which shows that Tianjin South Station is the critical disturbance node in the HSR network.

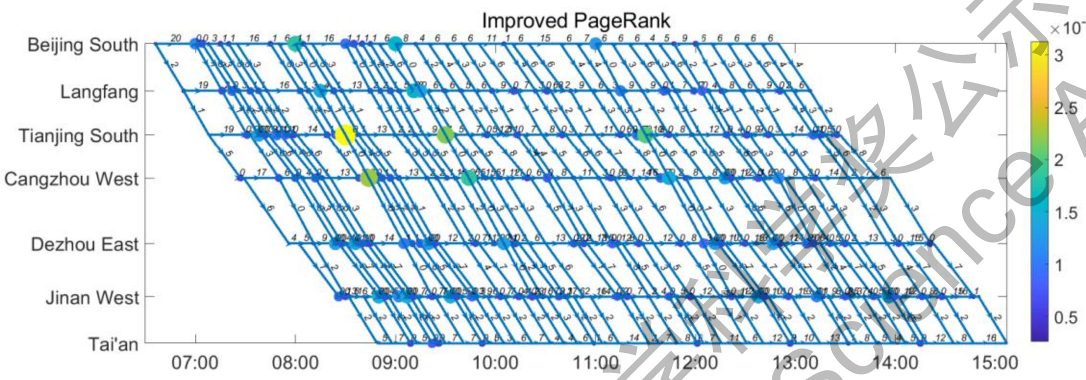

Fig. 10 Schematic diagram of the improved PageRank in the HSR network

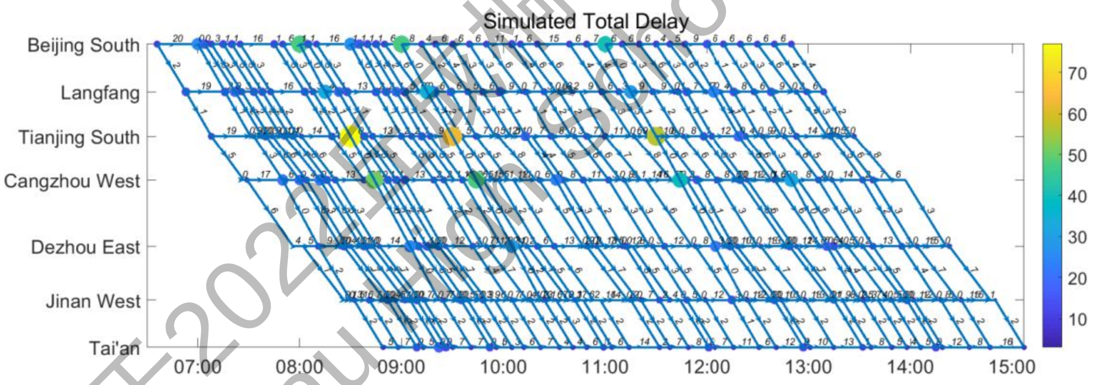

Fig. 11 Schematic diagram of the train total delay in the HSR network

Table 4 gives the timetable and the corresponding node parameters of train 10, including the train total delay, the node value and the node name. The bold parts in the table are the nodes whose total delay exceeds 50 min under simulation and the nodes whose improved PageRank value exceeds 2.00 E-4. The node value simulated by improved PageRank algorithm is consistent with the total delay obtained by simulation. Tianjin South Station can be effectively identified as the critical disturbance node in the timetable by the improved PageRank algorithm.

Table 4 Timetable and node parameters of train 10

Index	Station	Arrival time	Departure time	Train total delay	Improved PageRank	Node
1	Beijing South	-	8: 00	47 min	1.80E-4	(8: 00, 1, 10)
2	Langfang	8: 15	8: 15	33 min	1.41E-4	(8: 15, 2, 10)
3	Tianjin South	8: 30	8: 30	**77 min**	**3.13E-4**	(8: 30, 3, 10)
4	Cangzhou West	8: 44	8: 44	**51 min**	**2.27E-4**	(8: 44, 4, 10)
5	Dezhou East	9: 05	9: 05	24 min	1.00E-4	(9: 05, 5, 10)
6	Jinan West	9: 22	9: 24	3 min	5.55E-5	(9: 22, 6, 10)
		9: 22	9: 24	6 min	2.96E-5	(9: 24, 6, 10)
7	Tai'an	9: 41	-	6 min	2.61E-5	(9: 41, 7, 10)

3.2.2 Analysis of delay prediction and recovery

According to the identification results of the critical disturbance node in the above train timetable network, Tianjin South Station is the critical disturbance node, so the following two types of delay scenarios are set in Tianjin South Station.

(1) Departure delay scenario: The random departure delay is set of 1 to 10 min at Tianjin South Station. The serial numbers of departure delay scenarios are 1 to 10 respectively. For example, the train departs 1 min later than scheduled at Tianjin South Station, indicating the delayed departure scenario 1.

(2) Delay arrival scenario: Affected by the three arrival delay scenarios under TSR (the values of TSR are 300 km/h, 200 km/h and 120 km/h respectively), the train arrives later than scheduled at Jinan West Station.

Based on the delay prediction algorithm, the train speed profile in two types of delay scenarios is solved. Taking the arrival delay scenario with the TSR of 300 km/h as an example, the computed train speed profile is drawn in Fig. 12. According to the train speed profile, the train dispatcher can know the train motion process in the TSR interval and the passing time at the subsequent stations. It can be seen that the delay prediction algorithm can assist the train dispatcher to adjust the timetable.

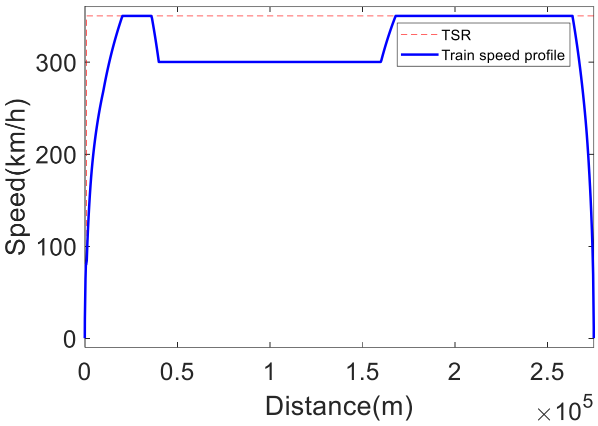

Fig. 12 Train speed profile under the TSR value of 300 km/h

According to the delay prediction algorithm, the train delay of the subsequent stations behind Tianjin South Station is calculated under the departure delay scenario, as shown in Table 5. When the train delay at Tianjin South Station is less than 2 minutes, the train will run quickly between Tianjin South and Cangzhou West. Finally, it resumes punctuality at Cangzhou West Station.

When the train departure delay at Tianjin South Station is more than 3 min but less than 6 min, the train can not return to punctuality at Dezhou East Station even if it accelerates using buffer time, but it can completely eliminate the train delay at Jinan West Station. Finally, if the train departs later at Tianjin South Station for more than 6 min, the train will not arrive on time at Jinan West Station. According to the predicted results of the train delay mentioned above, the train dispatcher can be directly assisted to automatically adjust the train timetable.

Table 5 Train arrival delay at the subsequent stations of Tianjin South Station under the departure delay scenario

Index of the departure delay scenario	Cangzhou West	Dezhou East	Jinan West
1	On time	On time	On time
2	0.62 min	On time	On time
3	1.62 min	On time	On time
4	2.62 min	0.14 min	On time
5	3.62 min	1.14 min	On time
6	4.62 min	2.14 min	0.31 min
7	5.62 min	3.14 min	1.31 min
8	6.62 min	4.14 min	2.31 min
9	7.62 min	5.14 min	3.31 min
10	8.62 min	6.14 min	4.31 min

The train delay at the subsequent stations behind Tianjin South Station is calculated under the three types of arrival delay scenarios, as shown in Fig. 13.

(1) As for the arrival delay scenario with a TSR of 300 km/h, although the train has a delay of 0.28 min at Cangzhou West Station, the train will return to punctuality at Dezhou East Station by using the buffer time between Cangzhou West and Dezhou East.

(2) For the arrival delay scenario with the TSR of 200 km/h, the train will have a delay of 5.87 min, 13.30 min and 11.46 min at the following three stations behind Tianjin South Station, respectively. This shows that the arrival delay scenario has a great impact on train operation. Although the buffer time between Dezhou East and Jinan West can reduce some train delay, it is difficult to completely eliminate the train delay through the buffer time. Using the delay prediction algorithm, dispatchers can further reduce and eliminate the train delay in adjusting train departure sequences at the subsequent stations.

(3) According to the delay prediction algorithm, the train will have a delay time of 16.26 min, 38.62 min and 36.78 min in Cangzhou West, Dezhou East and Jinan West, respectively. The train dispatcher needs to take necessary adjustment strategies to further reduce and eliminate the train delay.

(4) As for all the above-mentioned arrival delay and departure delay scenarios, the proposed delay prediction algorithm can give the predicted delay within 22 seconds, which meets the real-time requirements of HSR management.

It can be seen that the delay prediction algorithm based on the critical disturbance node identification can accurately forecast the predicted delay and its recovery at the subsequent stations behind Tianjin South Station in real time. The algorithm provides train dispatchers with acceleration and deceleration of trains in future running intervals. It can assist train dispatchers in adjusting stage plans in real time, reduce the operation times of train dispatchers manually adjusting timetable, reduce the work intensity of dispatchers and improve the efficiency of HSR.

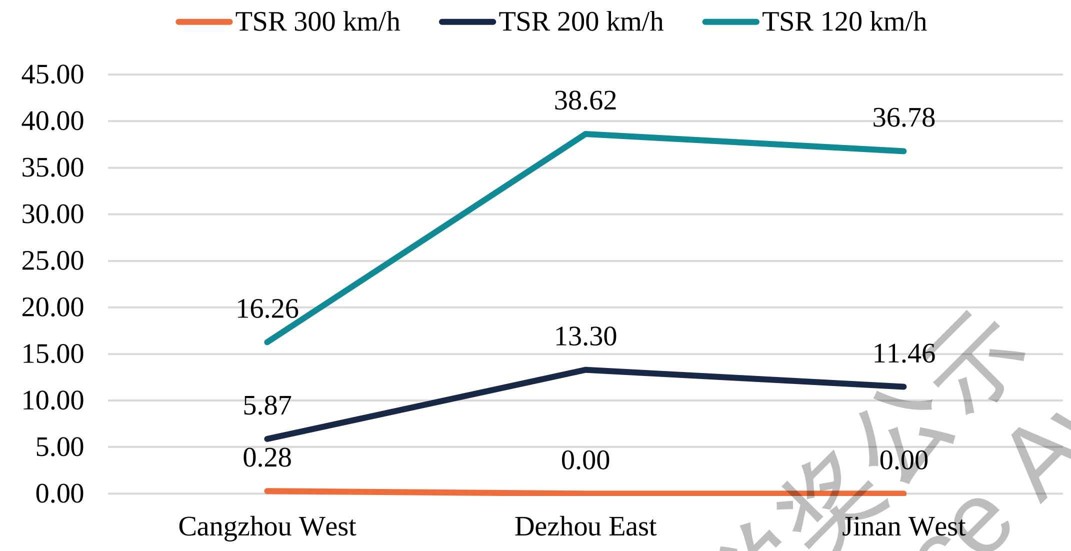

Fig. 13 Train delay at the subsequent stations behind Tianjin South Station under the arrival delay scenario

4 Summary and prospect

Based on China's HSR network and the timetable characteristics, this paper constructs a multi-layer network model of HSR timetable by using complex network theory. A delay prediction algorithm based on train time-efficient driving strategy and critical node identification is proposed. The algorithm can identify the critical disturbance node, predict the arrival time and delay recovery of trains at the subsequent stations, and assist dispatchers to adjust the train timetable in real time.

The proposed delay prediction algorithm in this paper can be used for the HSR dispatching command system. It can be designed as a tool to evaluate the stability and robustness in the HSR network, assists train dispatchers to make more scientific and reasonable dispatching strategies, and effectively cut off the train delay propagation chain as well. In the future, the effectiveness and accuracy of the delay prediction algorithm will be further improved. The critical inducements causing the delay of the critical disturbance node will be effectively identified. The train dispatchers will be assisted to make more targeted delay elimination strategies.

References

[1] Y. Wang, Q. Lu, X. Cao, X. Zhou, V. Latora, L. C. Tong, and W. Du, “Travel time analysis in the Chinese coupled aviation and high-speed rail network,” Chaos, Solitons & Fractals, vol. 139, p. 109973, 2020.

[2] E. V. Andersson, A. Peterson, and J. T. Krasemann, “Quantifying railway timetable robustness in critical points,” Journal of Rail Transport Planning & Management, vol. 3, no. 3, pp. 95 – 110, 2013.

[3] E. V. Andersson, A. Peterson, and J. T. Krasemann, “Reduced railway traffic delays using a MILP approach to increase Robustness in Critical Points,” Journal of Rail Transport Planning & Management, vol. 5, no. 3, pp. 110 – 127, 2015.

[4] E. Solinen, G. Nicholson, and A. Peterson, “A microscopic evaluation of railway timetable robustness and critical points,” Journal of Rail Transport Planning & Management, vol. 7, no. 4, pp. 207 – 223, 2017.

[5] H. Zhang, H. Cui, W. Wang, and W. Song, “Properties of Chinese railway network: Multilayer structures based on timetable data,” Physica A: Statistical Mechanics and Its Applications, vol. 560, p. 125184, 2020.

[6] W. Wang, K. Cai, W. Du, X. Wu, L. C. Tong, X. Zhu, and X. Cao, “Analysis of the Chinese railway system as a complex network,” Chaos, Solitons & Fractals, vol. 130, p. 109408, 2020.

[7] A. Saxena and S. Iyengar, “Centrality measures in complex networks: A survey,” arXiv preprint arXiv:2011.07190, 2020.

[8] X. Meng, L. Jia, and W. Xiang, “Complex network model for railway timetable stability optimisation,” IET Intelligent Transport Systems, vol. 12, no. 10, pp. 1369 – 1377, 2018.

[9] L. C. Freeman, “A set of measures of centrality based on betweenness,” Sociometry, vol. 40, no. 1, pp. 35 – 41, 1977.

[10] S. Brin and L. Page, “The anatomy of a large-scale hypertextual web search engine,” Computer Networks and ISDN Systems, vol. 30, no. 1-7, pp. 107 – 117, 1998.

[11] J. Ge, Z. Yuan, Q. Zhang, “An improved PageRank algorithm for critical nodes identification in railway timetable with disturbances by multilayer complex networks”, 2021 China Automation Congress (CAC).

[12] Peng Junbin, Liu Gang. Traction and braking of EMU. Beijing: China Railway Publishing House, 2009 (in Chinese).

[13] P. Howlett. An optimal strategy for the control of a train[J]. The Journal of the Australian Mathematical Society Series B Applied Mathematics, 1990, 31(4): 454-471.

Acknowledgement

This project was carefully guided by Prof. Zhang Qi, the chief researcher of China Academy of Railway Sciences (CARS). With independence and research basis from what I participated in the competition last year, I propose a train delay prediction method based on train time-efficient driving strategy and critical node identification. During the scientific research practice in CARS, I also received the guidance and help from Prof. Yuan Zhiming and Dr. Wang Rongsheng. I would like to express my sincerely thanks here! I would also like to thank Fan Keke, the teacher of "Scientific Research Practice" in the Youth Science Club of the High School Attached to the National People's Congress, for his guidance and the help of Zhang Si, the computer teacher, and Wang Chen, the class teacher.

1. The source and research background of the thesis

By the end of 2021, the operating mileage of China's high-speed railway has exceeded 40, 000 kilometers, accounting for more than 2/3 of the total mileage in the world. China has become the country with the longest operating mileage, the highest transportation density and the most complex network operation scenario in the world. However, under the huge passenger flow pressure and changeable operation scenarios, trains often have delays, which brings great pressure to the punctual operation of trains, and also has a great impact on the railway transportation organization and operation service quality. Therefore, it is necessary to study the mechanism and elimination methods of high-speed train delays in order to further ensure the punctual operation.

In the "Transportation Science Volume of 10,000 Scientific Problems" (2018) jointly sponsored by the Ministry of Education, the Ministry of Science and Technology, the Chinese Academy of Sciences and the National Natural Science Foundation of China, the scientific problems hold that "revealing the mechanism of train delay propagation" is one of the urgent problems to be solved in rail transit transportation organizations. The critical to revealing the mechanism of train delay propagation is to analyze and evaluate the critical disturbance node of high-speed train schedule delay and accurately predict the delay time. Although some institutions and scholars at home and abroad have carried out some research in this field, there are still some deficiencies in the micro-mechanism and cause analysis and evaluation methods of train delay process, and there is still much room for improvement.

Based on the National Natural Science Foundation of China-High-speed Railway Joint Fund Project "Collaborative Commanding-based Intelligent Dispatching for High-Speed Railways" (No. U1834211), and based on the actual needs of high-speed railway transportation organization, this project has carried out in-depth research on the delay prediction method of train driving strategy based on the identification of the critical disturbance node of high-speed train delay.

2. The work and contribution of each team member in writing the paper

This project is completed independently by myself under the guidance of Prof. Zhang.

3. The relationship between the tutor and student, the role of the tutor in the process of writing the paper, and whether the tutor are paid for.

I began to take the course "Scientific Research Practice" of Youth Science Club of the High School Affiliated to Renmin University of China in my freshman year. Then, I run the experiment in the Laboratory of High-speed Railway Operation and Dispatching Simulation, CARS, according to the course requirements. My tutor was Prof. Zhang Qi, chief researcher of CARS. The topic selection and writing of the research report are independently completed by myself under the guidance of Prof. Zhang. There is no paid guidance.

4. Research results assisted by others

The work of this paper is completed independently under the guidance of my tutor.

Brief introduction of my tutor

Zhang Qi is currently a chief researcher of CARS. His research interests include railway signal and communication, automatic train operation, train operation control, intelligent dispatching, and cooperative control of multiple trains. He has published more than 80 papers in well-known academic journals at home and abroad (including 57 papers retrieved by SCI and EI). He has directed many national scientific projects in China and has a lot of achievements. One of his research achievements, the Centralized Traffic Control (CTC) system, has become one of the most important technical equipment in Chinese railway. Prof. Zhang was a recipient of the Youth Award, the Contribution Award, and the Achievement Award of Zhan Tianyou Railway Science and Technology. As the first finisher, he won the first prize of scientific and technological progress of China Railway Society in 2017, the first prize of scientific and technological progress of China Intelligent Transportation Association in 2018 and the second prize of scientific and technological progress of Beijing in 2019.