

Received January 19, 2020, accepted February 1, 2020, date of publication February 6, 2020, date of current version March 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2972130

# Neural Network-Based Prediction Model for Passenger Flow in a Large Passenger Station: An Exploratory Study

ZHUCUI JING<sup>1</sup> AND XIAOLI YIN<sup>2</sup>

<sup>1</sup>School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China

<sup>2</sup>Business College of Shanxi University, Taiyuan 030031, China

Corresponding author: Zhucui Jing (zcjing@bjtu.edu.cn)

This work was supported by the Fundamental Research Funds for the Central Universities under Grant 2016JBWZ007.

**ABSTRACT** As the hub and carrier to transfer the passengers, the railway station is an important factor that affects the rail passenger transportation because the normal operation of the station without load redundancy is determined by the moderate passenger flow. It means reasonable and accurate prediction of passengers entering and leaving the station can provide the basis and guarantee for the station security, the resources allocation and the personnel deployment. Since the neural network model is good at processing the common regular data changes through training the network and adjusting the weight value based on a large number of training samples, the neural network model is used in processing the short-term irregular data to predict the passenger flow at the railway station which is susceptible to the constantly changing external factors. In this paper, the neural network is used to predict the passenger flow. First, the key factors affecting the change of the passenger flow are selected and analyzed as the input of the neural network. Second, the learning and the rate updating of variable step size are adopted to estimate the number people entering the station during a certain time interval, which is then weighted with the historical data to derive the prediction of the passenger flow during the next time interval. The simulation results show that the experiment results show that the method proposed in this paper can better track and predict the sudden changes in the passenger flow caused by emergencies. Meanwhile, it can be found that in the process of forecasting abnormal passenger flow, the most critical link is to summarize and summarize the characteristics of railway station passenger flow, clarify the type and time distribution of passenger flow at each station, and analyze the factors that cause abnormal changes in passenger flow.

**INDEX TERMS** Railway station, passenger flow, neural network, prediction.

## I. INTRODUCTION

In Asia, especially in China, rail transit is favored by more people for it can solve the growing contradiction between the supply and the demand of transportations in large and medium cities with its technical advantages of large capacity, fast speed, little pollution, high safety and good punctuality. In recent years, with the development of the urban rail transit system, the high-speed rail, the bullet train, the inter-city train and the ordinary train have become more and more popular and multi-functional, bringing great convenience to the people's travelling [1]. Now, the trains are becoming the

first choice for more and more people to travel for its steadily increased speed, high safety and excellent comfort.

As the hub and carrier to transfer the passengers, the railway station is an important factor that affects the rail passenger transportation because the normal operation of the station without load redundancy is determined by the moderate passenger flow [2], [3]. It means reasonable and accurate prediction of passengers entering and leaving the station can provide not only the basis for the personnel deployment and the resources allocation but also the strong support for the security work.

The passenger flow volume not only changes with time, but also is the function of time which is of high nonlinearity since it's affected by multiple external factors. Compared with other similar application scenarios (shopping malls, highway,

The associate editor coordinating the review of this manuscript and approving it for publication was Sabah Mohammed<sup>1</sup>.

scenic spots), the passengers entering the railway station are of more obvious characteristics such as obvious periodicity which is measured by days. For example, first, the passenger flow at the railway station is susceptible to the weather. But the overall trend nearly remains the same despite the local fluctuations caused by the weather. Second, at certain time intervals, the number of passengers entering the station at each time interval is closely connected to the number of trains at the next time interval: the fluctuation in the passengers' number is particularly evident during the winter and summer vacation and other statutory holidays.

This paper is going to estimate the number of passengers entering the railway station at different time intervals by using the neural network for its great advantage in data prediction [4] and the historical data [5]. Since the neural network model is good at processing the common regular data changes through training the network and adjusting the weight value based on a large number of training samples, the neural network model is used in processing the short-term irregular data or data that will abruptly change to predict the passenger flow at the railway station which is susceptible to the constantly changing external factors.

The rest of this paper is organized as follows: we first conduct a comprehensive review that forms the theoretical foundation of this study. Section 3 discusses applications of prediction model. Section 4 discusses the improved prediction model for the prediction of the passenger flow. Section 4 presents the experiment. Section 5 concludes.

## II. LITERATURE REVIEW

In 1984, Okutani and Stephanedes [6] applied Kalman filtering theory to the prediction of dynamic short-term passenger flow. Friedman *et al.* [7] first proposed the theory of combining the neighborhood prediction method with the time series method in 1977. Davis and Nihan first introduced the nearest neighbor prediction method and time series method to short-term passenger flow prediction in 1991 [8]. Since the k-nearest neighbor algorithm does not require parameter assumption and parameter adjustment process, the prediction effect is better. The research of Smith *et al.* [9] Mainly focused on how to reduce the time-consuming problem of the algorithm. Some scholars studied the prediction of traffic passenger flow data using passenger flow data of neighboring time observation points [10], [11].

Setyawati and Creese [12] applied a differential autoregressive integrated moving average (ARIMA) prediction model to the short-term passenger flow prediction of railway station. The algorithm considered that the model could not solve the influencing factors of emergency events and achieved good results Effect. In 1997, Arem *et al.* [13] summarized the existing short-term traffic flow prediction theory at that time, pointed out the necessity of short-term traffic flow prediction, and the accuracy of short-term traffic flow prediction technology needs to be improved.

In 1943, biologists and mathematical physiologists jointly demonstrated the mathematical model of neurons, which

laid a solid foundation for future study and research [9]. The main achievements of foreign scholars using neural networks for prediction are: In 2001, Dia [14] and other researchers proposed a goal-oriented neural network prediction model to predict short-term passenger flows, and obtained good results. Chen and Muller [15] studied the short-term traffic flow prediction method of neural network model based on dynamic ranking learning with great advantages; Vlahogianni *et al.* [16] optimized the neural network prediction model, Structural optimization theory is applied to road traffic flow prediction. Through simulation experiments, multi-section prediction is achieved and the results show that the optimization model effectively reflects the random variability of road traffic flow. In 2012, Zheng *et al.* [17] combined Bayes rule and conditional probability theory to optimize different neural networks, and applied the optimized model to the prediction of short-term traffic flow on expressways. The example verification results show that the optimized prediction results are more accurate than the prediction results of a single neural network model.

In 2010, Sugiyama *et al.* [18] proposed a real-time forecasting method for passenger flow in railway station; Ozerova [19] used the bilinear correlation method to study the influencing factors of intercity commuter passenger flow, and then use linear regression method to predict commuter passenger flow; Hrushevskaya [20] discussed the passenger flow rule from the train schedule and other aspects, and based on this, the suburb The daily passenger flow of the railway station is forecasted.

In recent years, the combined forecasting model has become the mainstream of passenger flow forecasting. The combined forecasting model combines several algorithms that perform well to effectively overcome the limitations of a single algorithm in forecasting and improve forecast accuracy. Hao *et al.* [21] analyzed the short-term passenger flow changes in railway station, and analyzed the short-term historical passenger flow changes in detail. Guo *et al.* [22] studied the short-term prediction method of interval passenger flow, and selected the BP neural network based on the characteristics of interval passenger flow itself; the interval passenger flow forecasting was realized with Matlab, and a comparative analysis was performed, empirical research shows that the characteristics of interval passenger flow data and neural network have good prediction accuracy. George *et al.* [23] proposed that passenger flow forecasting has important strategic significance in the management of transportation systems. They adopted a new data source, social media, to address this challenge, and developed a systematic approach to check social media activity and perceived events. Preliminary analysis shows that there is a moderately positive correlation between passenger flow and the ratio of social media posts. This discovery prompted researchers to develop a new method for improving traffic prediction. Asmer *et al.* [24] conducted a feasibility study on the introduction of security checks at train stations. A simulation model for Brunswick, Germany train station has been

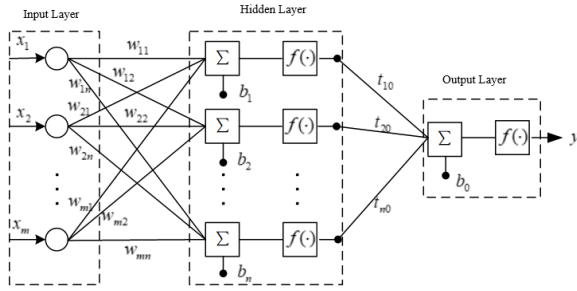


FIGURE 1. BP neural network model.

created and used to investigate the impact of additional security checks on the passenger flow.

It's found that there's a close correlation between the data at adjacent time intervals, so this paper decides to make the prediction of the data at the next time interval according to the data at the previous time interval. In this way, the short-term fluctuation of the data can be better tracked and predicted. Meanwhile, the prediction results by the model and the historical data were weighted to obtain the final estimation result of the passenger flow volume which has been proven to be of higher accuracy by the simulation results [25], [26].

### III. THE PASSENGER FLOW PREDICTION MODEL

The passenger flow prediction model is established mainly through the BP neural network, in which the input is the major factors influencing the change of the passenger flow, and the output is the prediction results of the passenger flow. The method of rate updating of variable step size is adopted to avoid oscillation and falling into local optimal value. Since the station schedule is designed in a day cycle and the passenger flow is of distinct periodicity without considering the interference of external factors, it is of higher accuracy to make the weighted prediction by combining with the historical data.

#### A. CONSTRUCTION OF THE NEURAL NETWORK FOR THE PASSENGER FLOW PREDICTION IN THE PASSENGER STATION

Here, a three-layer BP neural network is used as a model for predicting the number of people entering the station. The three-layer neural network mainly includes the input layer, the hidden layer and the output layer. There is no connection within the layer, but there are full connections between the layers. The specific network structure is shown in the figure 1.

Among them,  $x_i, i = 1, 2, \dots, m$  represent the key factors affecting the number of people entering the station.

Here, a hyperbolic tangent S-type function is used as the transfer function of the hidden layer and the output layer  $f(\cdot)$ , that is:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad (1)$$

The input vector is  $X = [x_1, x_2, \dots, x_m]^T$ . Each element represents a quantified factor affecting the passenger flow at

the station. These factors usually include weather, holidays, etc.  $W$  is the connection weight vector from the input layer to the hidden layer, which can be explicitly expressed as the following matrix.

$$W = \begin{bmatrix} w_{11} & w_{21} & \dots & w_{m1} \\ w_{12} & w_{22} & \dots & w_{m2} \\ \vdots & \vdots & \dots & \vdots \\ w_{1n} & w_{2n} & \dots & w_{mn} \end{bmatrix}, \quad (2)$$

In this paper, in the calculation of the output value of the hidden layer, the threshold vector  $B = [b_1, b_2, \dots, b_n]^T$  is regarded as a node weight value of constant 1 in the input layer. Therefore, the input vector  $X$  can be extended to  $X' = [X^T, 1]^T = [x_1, x_2, \dots, x_m, 1]^T$ ; correspondingly, the weight vector from the input layer to the hidden layer  $W$  can be extended to:

$$W' = \begin{bmatrix} w_{11} & w_{21} & \dots & w_{m1} & b_1 \\ w_{12} & w_{22} & \dots & w_{m2} & b_2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ w_{1n} & w_{2n} & \dots & w_{mn} & b_n \end{bmatrix}, \quad (3)$$

In this case, the output  $F$  can be expressed as follows:

$$F = f(W' \cdot X'), \quad (4)$$

Similarly, the weight vector from the hidden layer to the output layer  $T$  can be expanded to  $T' = [t_{10}, t_{20}, \dots, t_{n0}, b_0]$ , so the output of the output layer of the BP neural network is:

$$y = [F^T \quad 1] \cdot T'^T, \quad (5)$$

#### B. THE PREDICTION MODEL OF THE NUMBER OF PEOPLE ENTERING THE STATION

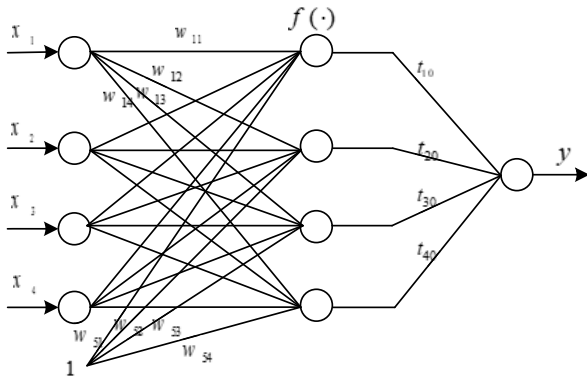
For time  $k$ , the input vector is  $X'(k) = [x_1(k), x_2(k), x_3(k), x_4(k), 1]$ ,  $x_1(k)$  represents factors affecting the number of people entering the station, which indicates the actual flow of people entering the station in the time interval  $((k - 1)T_0, kT_0]$ ,  $T_0$  represents the interval period of two adjacent time points, such as one hour or one day;  $x_2(k)$  is the weather condition at time  $KT_0$ , the value of  $x_2$  and its corresponding neural network input value constitute a hash table {sunny day = 1, rainy day = 2, snowy day = 3, gale = 4};  $x_3(k)$  represents the outdoor temperature  $P$  at time  $KT_0$ , the value of and its corresponding quantized neural network input value also constitute a hash table:

This is an example of an equation:

$$\{(P < 0^\circ) = 1, (0^\circ \leq P < 10^\circ) = 2, (10^\circ \leq P < 20^\circ) = 3, (20^\circ \leq P < 30^\circ) = 4, (30^\circ \leq P) = 5\}, \quad (6)$$

$x_4(k)$  represents holidays during time intervals, similarly, the value of  $x_4(k)$  and its corresponding neural network input value can also be expressed as a hash table {statutory holiday = 1, winter and summer vacation is 2, weekend is 3, else is 4}.

Based on the above definition of the input layer nodes and the transformation of the model calculation in



**FIGURE 2.** The neural network prediction model of the station passenger flow.

section 2.1, the structure of the prediction model of the number of people entering the station at different time intervals is shown in Figure 2. Among them, the node whose input layer is identified as 1 is used to transform the threshold vector into the weight vector, and the value is updated iteratively by the gradient descent method.

The output of the station passenger flow prediction model based on the neural network is as follows:

$$\tilde{x}(k+1) = \sum_{j=1}^4 \{t_{j0}(k-1)\} f\left(\sum_{i=1}^4 w_{ij}(k-1)x_i(k) + w_{5j}(k)\right) \quad (7)$$

$\{t_{j0}(k-1), w_{ij}(k-1)\}$ , the weight value of the trained neural network during the  $k-1$  time interval and  $X(k) = [x_1(k), x_2(k), x_3(k), x_4(k), 1]$ , the input value of the different conditions during the  $k$  time interval are used to predict  $\tilde{x}(k+1)$ , the station passenger flow volume during the  $k+1$  time interval. The initial weight  $\{t_{j0}(0), w_{ij}(0)\}$  is obtained by training historical data.

#### IV. THE IMPROVED PASSENGER FLOW PREDICTION MODEL

Due to the advantage of algorithm, it is developed to do research in many areas; while the network will be unable to achieve the desired results with a standard algorithm as demonstrated by many cases. The main reasons are the defects of the algorithm itself: (1) Easy to form the local minimum but not the global optimum; (2) Low learning efficiency and slow convergence rate because of massive training times; (3) The falsely saturated in learning process. We carried out the following improvements toward the defects.

##### A. WEIGHT VECTOR UPDATE

An error function is obtained by comparing the difference between the passenger flow obtained from the actual statistics and the passenger flow predicted by the neural network during the  $k$  time interval:

$$E(k) = \frac{1}{2}(x(k) - \tilde{x}(k))^2, \quad (8)$$

Among it,  $x(k)$  represents the actual statistics of the passenger flow during the  $k$  time interval,  $\tilde{x}(k)$  represents the passenger flow volume during the  $k$  time interval predicted by the neural network model.

Here, the gradient descent method is used to adjust the weight vector from the input layer to the middle layer  $W(k)$  and the weight vector from the middle layer to the output layer  $T(k)$  to make  $E(k) \rightarrow 0$ . The specific adjustment mechanism is as follows:

This is an example of an equation:

$$T(k) = T(k-1) + \delta \frac{\partial E(k-1)}{\partial T(k-1)}, \quad (9)$$

$$W(k) = W(k-1) + \eta(k-1) \frac{\partial E(k-1)}{\partial W(k-1)}, \quad (10)$$

$$\eta(k-1) = \eta(k-2) + r < \frac{\partial E(k-1)}{\partial W(k-1)}, \frac{\partial E(k-2)}{\partial W(k-2)} >, \quad (11)$$

where,

$$T(k) = [t_{10}(k), t_{20}(k), t_{30}(k), t_{40}(k)]^T,$$

$$W(k) = [w_{11}(k), w_{12}(k), \dots, w_{53}(k), w_{54}(k)]^T$$

And the weight vector  $T(k)$  is updated by step size  $\delta$  while the weight vector  $W(k)$  is updated by step size  $\eta(k-1)$ . The inner product term of the error gradient of the current sample and the previous sample is considered at the same time. When the inner product is positive, it indicates that the two adjustments have the same direction on the error surface, and the learning rate can be increased to quickly search for the global optimum; when the inner product is negative, it indicates that an inflection point is encountered on the error surface so the learning rate should be reduced to avoid oscillation.

##### B. THE WEIGHTED FUNCTION OF THE PASSENGER FLOW PREDICTION

Since the number of trains at each passenger station is in a day cycle, the fluctuation of the passenger flow in different time intervals of each day should have the same trend. Besides, there is a certain connection between the passenger flows of two adjacent days. Therefore, the following weighted prediction model of the passenger flow is considered:

$$\tilde{x}(k) = \alpha x'(k) + (1 - \alpha)\tilde{x}(k), \quad (12)$$

Among it,  $\tilde{x}(k)$  is the predicted value of the final passenger flow of the model in this paper during the  $k$  time interval;  $x'(k)$  is the real value of the passenger flow during the  $k$  time interval;  $\tilde{x}(k)$  is the predicted value of the neural network for the passenger flow during the  $k$  time interval;  $\alpha$  is an adjustable weighting coefficient.

Based on the calculation and transmission of the above data, the prediction model of the station passenger flow as shown in the figure below is obtained. Among them,  $X(k)$  is the input of the neural network and  $\tilde{x}(k+1)$  is the final output value of the prediction model.

**C. THE ERROR FUNCTION TO ACCELERATE THE LEARNING SPEED**

Input layer can be changed as

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \frac{\partial E \partial net_k}{\partial net_k \partial w_{jk}} \quad (13)$$

Hidden layer can be changed as

$$\Delta v_{ij} = -\eta \frac{\partial E}{\partial v_{ij}} = -\eta \frac{\partial E \partial net_j}{\partial net_j \partial v_{ij}} \quad (14)$$

Error signal  $\delta_k^o = -\eta \frac{\partial E}{\partial net_k}$  and  $\delta_j^y = -\eta \frac{\partial E}{\partial net_j}$ , so

$$\Delta w_{jk} = \eta \delta_k^o y_j = \eta (d_k - o_k) o_k (1 - o_k) y_j \quad (15)$$

$$\Delta v_{ij} = \eta \delta_j^y x_i = \eta \left( \sum_{k=1}^l \delta_k^o w_{jk} \right) y_j (1 - y_j) x_i \quad (16)$$

And

$$\delta_k^o = -\eta \frac{\partial E}{\partial net_k} = -\eta \frac{\partial E \partial o_k}{\partial o_k \partial net_k} = -\eta \frac{\partial E}{\partial o_k} f'(net_k) \quad (17)$$

$$\delta_j^y = -\eta \frac{\partial E}{\partial net_j} = -\eta \frac{\partial E \partial y_j}{\partial y_j \partial net_j} = -\eta \frac{\partial E}{\partial y_j} f'(net_j) \quad (18)$$

For output layer, do differentiation for

$$\frac{\partial E}{\partial o_k} = -(d_k - o_k) \quad (19)$$

For hidden layer, do differentiation for  $E$

$$\frac{\partial E}{\partial y_i} = - \sum_{k=1}^l (d_k - o_k) f'(net_k) w_{jk} \quad (20)$$

And  $f'(x) = f(x)[1 - f(x)]$ , so

$$\delta_k^o = (d_k - o_k) o_k (1 - o_k) \quad (21)$$

$$\begin{aligned} \delta_j^y &= \left[ \sum_{k=1}^l (d_k - o_k) f'(net_k) w_{jk} \right] f'(net_j) \\ &= \left( \sum_{k=1}^l \delta_k^o w_{jk} \right) y_j (1 - y_j) \end{aligned} \quad (22)$$

Therefore, connection weights are

$$\Delta w_{jk} = \eta \delta_k^o y_j = \eta (d_k - o_k) o_k (1 - o_k) y_j \quad (23)$$

$$\Delta v_{ij} = \eta \delta_j^y x_i = \eta \left( \sum_{k=1}^l \delta_k^o w_{jk} \right) y_j (1 - y_j) x_i \quad (24)$$

So error gradient expression:

$$\frac{\partial E}{\partial w_{ik}} = -\delta_k^o y_j \quad (\delta_k^o \text{ is the error signal}) \quad (25)$$

Minimal error gradient means  $\delta_k^o$  is closed to zero. And

$$\delta_k^o = (d_k - o_k) o_k (1 - o_k) \quad (26)$$

So one new error function is introduced:

$$E_{\Sigma} = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^L \left[ \left( 1 + d_k^p \right) \ln \frac{1 + d_k^p}{1 + o_k^p} + \left( 1 - d_k^p \right) \ln \frac{1 - d_k^p}{1 - o_k^p} \right] \quad (27)$$

Activation function of the neuron is the hyperbolic tangent function:

$$f(a) = \tanh(a) = \frac{1 - e^{-2a}}{1 + e^{-2a}} = \frac{e^a - e^{-a}}{e^a + e^{-a}} \quad (28)$$

where,  $a$  is the input of activation function.

The output range is  $(-1,1)$  for the hyperbolic tangent function, thus reducing the training time.

Because,

$$f^2(a) = \tanh^2(a) = \frac{e^{2a} - 2e^a e^{-a} + e^{-2a}}{(e^a + e^{-a})^2} = \frac{e^{2a} - e^{-2a} - 2}{(e^a + e^{-a})^2} \quad (29)$$

$$f'(a) = \tanh'(a) = \frac{(e^a + e^{-a})(e^a + e^{-a}) - (e^a - e^{-a})(e^a - e^{-a})}{(e^a + e^{-a})^2} \quad (30)$$

$$= \frac{(e^{2a} + e^{-2a} + 2) - (e^{2a} + e^{-2a} - 2)}{(e^a + e^{-a})^2} = 1 - f^2(a) \quad (31)$$

Take output layer for example, so

$$\delta_k^o = o_k (d_k - o_k) \quad (32)$$

Compared with the error signal  $\delta = f'(net_k)(d_k - o_k)$ ,  $f'(net_k)$  disappears. So that, it can accelerate learning rate.

**V. EXPERIMENT AND RESULTS**

The annual passenger flow data of Beijing from June 1, 2017 to May 31, 2018 are shown in Table 1:

From Table 1, it can be seen that the passenger flow of the Beijing has a certain periodicity. Among them, the passenger flow in February 2018 has a large fluctuation. The main reason for this phenomenon is that the most important traditional festival in China is the Spring Festival. In January, Beijing was a first-tier city dominated by immigrants. At this time of the year, a large number of immigrants left Beijing to return to their hometowns, causing a sharp increase in passenger traffic. In addition, another peak period of passenger traffic growth is in July of each year. This period is the summer vacation period of schools in China. As a cultural, political, and economic center of China, Beijing has also become an important place for many students to travel. As a result, the passenger flow of the Beijing railway station passenger flow has increased.

In the simulation part, the time interval is divided by hours and  $T_0 = 1$ , so there are 24 time intervals in each day. The data of the 32-day passenger flow at the simulated station is taken as the sample set, in which the first 30 days are taken as the training samples, the 31st and 32nd days are taken as the test samples respectively, and the coefficient in the weighted prediction function  $\alpha$  is taken as 0.4. The MATLAB simulation is used to compare the performance of the weighted neural network prediction model proposed in this paper and the general basic neural network prediction model.

TABLE 1. Experiment data.

Data	Daily passenger traffic/Ten thousand people	Data	Daily passenger traffic/Ten thousand people
2017/6/1	132.5	2017/10/1	134.2
2017/6/2	113.5	2017/10/2	106.5
2017/6/3	136.5	2017/10/3	117.5
2017/6/4	96.7	2017/10/4	86.8
2017/6/5	124.7	2017/10/5	117.6
2017/6/6	94.2	2017/10/6	93.9
2017/6/7	88.3	2017/10/7	94.8
2017/6/8	85.2	2017/10/8	128.7
2017/6/9	95.3	2017/10/9	116.2
2017/6/10	105.8	2017/10/10	124.4
2017/6/11	111.7	2017/10/11	128.4
2017/6/12	125.7	2017/10/12	97.2
2017/6/13	93.7	2017/10/13	138.2
2017/6/14	105.5	2017/10/14	137.1
2017/6/15	94.6	2017/10/15	121.6
2018/1/31	85.2	2018/2/14	72.9

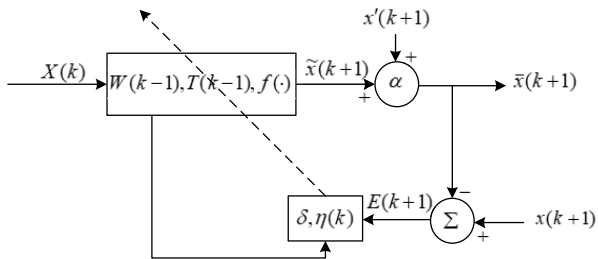


FIGURE 3. The prediction model of the station passenger flow.

The basic neural network takes the data of different time intervals in the first 30 days in the sample set as the training set for batch input training, and finally obtain the weight vector satisfying the training accuracy and the corresponding output results according to the input test vector. The weighted neural network proposed in this paper repeatedly trained the first 30 samples in the sample set and adjusted the weight vector one by one. After the training, the test vector is input to obtain the prediction result. At the same time, the weight vector is further adjusted according to the error function for the prediction at the next moment.

The Percentage Relative Error (PRE) is used to compare the prediction accuracy:

$$PRE = \frac{|\bar{x}(k) - x(k)|}{x(k)} \times 100\% \quad (33)$$

Figure 4 shows the prediction results and error percentage curves (interpolation fitting) of the two neural networks on the 31st day. The passenger flow on Day 31 has no obvious characteristics compared with the previous 30 days and

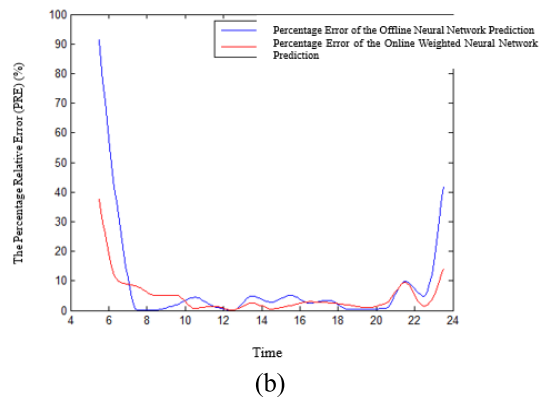
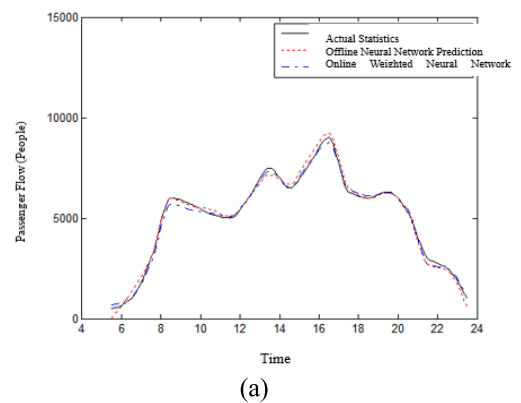
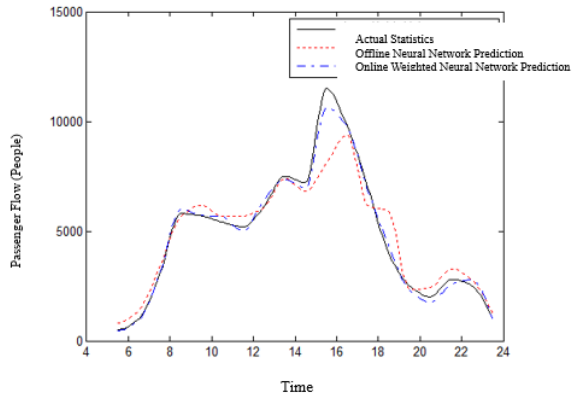
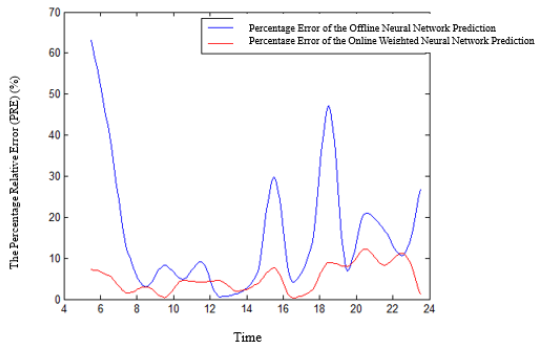


FIGURE 4. The prediction results and percentage error curves on day 31.

the overall trend of the passenger flow remain consistent. It can be seen from the simulation results that both the basic neural network and the weighted neural network have good



(a)



(b)

FIGURE 5. The prediction results and percentage error curves on day 32.

performance in the prediction of the passenger flow on Day 31, with the percentage error basically below 10%. The weighted neural network proposed in this paper is better than the general basic neural network in accuracy as a whole.

Figure 5 shows the prediction results and error percentage curves (interpolation fitting) of the two neural networks on the 32nd day. The simulated data of the 32nd day is the passenger flow affected by the short-term continuous rainfall at 3 p.m. It can be seen from the curve of the actual statistical results that the passenger flow increased rapidly during the period after the rain at 3 p.m. but declined rapidly after a sustained period of the time. Normally, the data in the abrupt situations cannot be well predicted by the basic neural network because the basic neural network relies too much on the historical data whose weight value won't be updated after the training. However, the basic neural network proposed in this paper takes the environmental factors into consideration and weights the historical data according to these factors, so the weight vector can be constantly updated according to the error value between the prediction results and the actual statistics in each time interval for the prediction during the next time interval. It can be seen from the prediction results that the weighted neural network can quickly track and predict the data the moment the data abruptly change; and based on the comparison results of the percentage error, the prediction performance of the weighted neural network is better than that of the basic neural network.

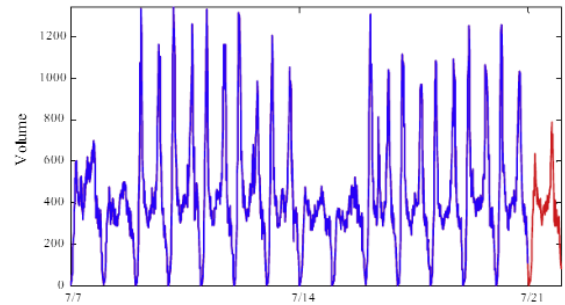


FIGURE 6. Alighting passenger flow at Beijing station.

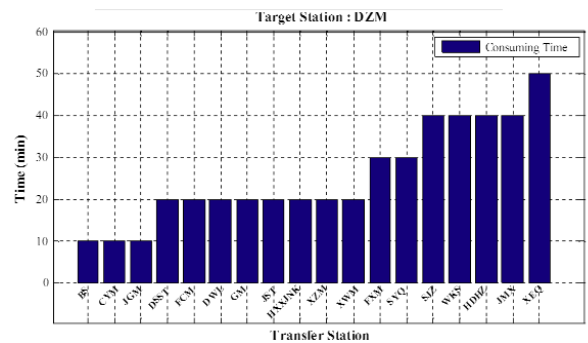


FIGURE 7. Elapsed time from transit station to Beijing target station.

The Beijing station as the prediction target station is considered to be an unknown structure (black box) dynamic system. The goal is to build a mathematical model that can be used to predict the downstream passenger flow at Beijing Station. The railway system is considered an input and single output system, while the drop-off passenger flow (circled in red) at Beijing Station is the system output. The boarding passenger flow of the 18 transfer stations (circled in other colors) in Figure 6 is the system input. Smart card data is recorded from 4:45 am to 23:15 am with a sampling rate of 15 minutes. Therefore, obtaining 76 data points per day results in a total of 32 days of observations and  $15 \times 76 \times 1140$  data points, as shown in Figure 7. The first 1064 data points are used for model identification, and the remaining 76 data points are used for model prediction. The goal is to use the proposed improved neural network modeling method for iterative multi-step prediction.

In order to facilitate the description of the drop-off passenger flow at the Beijing target station, the boarding passenger flow is instructed from the 18 original transit stations of the Beijing station. The next step is to determine the input lag shown for the longest time. Suppose the passenger chooses these routes with the least number of transfers, and estimates the commute time between any two stations by calculating the in-train travel time and transfer time. For example, it takes 30 minutes from the Beijing station to the other stations, which is equivalent to two time lags considering the 15-minute sampling rate. In other words, the boarding passenger flow at other stations 30 minutes ago was not enough

to help passengers getting off at Beijing station and should not be included in the prediction model. However, passengers boarding from other stations may take longer. 30 minutes to Beijing Station, as the actual travel time between each station for each passenger may exceed the minimum commute time (ie 30 minutes). In order to delay arrival, a 30-minute buffer is added during the shortest commute time, which is equivalent to two time intervals. Therefore, the maximum time lag for boarding passenger traffic from other stations can be calculated as the shortest calculation time (30 minutes) plus the buffer time (30 minutes) divided by the sampling rate (15 minutes). Based on similar logic, the maximum time lag for the remaining 17 stations can be calculated. The results show that each station will provide three time-delayed boarding passenger flows to the passenger flow at Beijing Station. Therefore, a total of 55 candidate variables were selected.

## VI. CONCLUSION

In this paper, the neural network is used to predict the passenger flow to prove that such prediction can bring great convenience to the station security, the resources allocation and the personnel deployment. The key factors affecting the passenger flow constitute the input of the neural network and the prediction results constitute the output. Besides, the prediction method is used to predict the passenger flow during the next time interval based on that during the previous time interval by weighing the prediction results with the history data of the passenger flow during the next time interval. Then the weight vector is modified according to the error between the weighted prediction value and the actual statistical value. Meanwhile, an improved step size updating method is used to prevent the error function from oscillating or falling into the local minimum during the updating of the weight value. The simulation experiment compares the performance of the weighted neural network prediction method proposed in this paper and the traditional basic neural network prediction method. The experiment results show that the method proposed in this paper can better track and predict the sudden changes in the passenger flow caused by emergencies. Meanwhile, it can be found that in the process of forecasting abnormal passenger flow, the most critical link is to summarize and summarize the characteristics of railway station passenger flow, clarify the type and time distribution of passenger flow at each station, and analyze the factors that cause abnormal changes in passenger flow. Classification, and finally, mathematical statistical methods are used to establish a functional relationship between various parameters, so as to obtain an effective plan for abnormal passenger flow prediction.

## REFERENCES

- [1] X. Wan, J. Yuan, and Q. Li, "An integrated framework for exploring the vulnerability due to passengers' aberrant behaviors in an urban metro system," in *Proc. ICCREM*, Sep. 2015, pp. 1124–1135.
- [2] X. Chu, J. Liu, D. Gong, and R. Wang, "Preserving location privacy in spatial crowdsourcing under quality control," *IEEE Access*, vol. 7, pp. 155851–155859, 2019.
- [3] D. Gong, M. Tang, G. Xue, H. Zhang, and B. Buchmeister, "Multi-objective cluster intelligent algorithms for railway door-to-door transportation routing design," *IEEE Access*, vol. 7, pp. 146702–146723, 2019.
- [4] W. Wu, L. Huang, and R. Du, "Simultaneous optimization of vehicle arrival time and signal timings within a connected vehicle environment," *Sensors*, vol. 20, no. 1, p. 191, Dec. 2019.
- [5] Y. Lu, X. Pei, C. Zhang, H. Luo, B. Liu, and Z. Ma, "Design of multi-modal transport path optimization model and dual pheromone hybrid algorithm," *J. Européen des Systèmes Automatisés*, vol. 52, no. 5, pp. 477–484, Nov. 2019.
- [6] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transp. Res. B, Methodol.*, vol. 18, no. 1, pp. 1–11, Feb. 1984.
- [7] J. H. Freidman, J. L. Bentley, and R. A. Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Trans. Math. Softw.*, vol. 3, no. 3, pp. 209–226, Sep. 1977.
- [8] G. A. Davis and N. L. Nihan, "Nonparametric regression and short term freeway traffic forecasting," *J. Transp. Eng.*, vol. 117, no. 2, pp. 178–188, Mar. 1991.
- [9] B. L. Smith, B. M. Williams, and R. K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transp. Res. C, Emerg. Technol.*, vol. 10, no. 4, pp. 303–321, Aug. 2002.
- [10] R. E. Turochy, "Enhancing short-term traffic forecasting with traffic condition information," *J. Transp. Eng.*, vol. 132, no. 6, pp. 469–474, Jun. 2006.
- [11] V. J. Hodge, R. Krishnan, J. Austin, J. Polak, and T. Jackson, "Short-term prediction of traffic flow using a binary neural network," *Neural Comput. Appl.*, vol. 25, nos. 7–8, pp. 1639–1655, Dec. 2014.
- [12] B. R. Setyawati and R. C. Creese, "Genetic algorithm for neural networks optimization," *Proc. SPIE*, vol. 5605, pp. 54–61, Nov. 2004.
- [13] B. Van Arem, H. R. Kirby, M. J. M. Van Der Vlist, and J. C. Whittaker, "Recent advances and applications in the field of short-term traffic forecasting," *Int. J. Forecasting*, vol. 13, no. 1, pp. 1–12, Mar. 1997.
- [14] H. Dia, "An object-oriented neural network approach to short-term traffic forecasting," *Eur. J. Oper. Res.*, vol. 131, no. 2, pp. 253–261, Jun. 2001.
- [15] H. Chen and S. Grant-Muller, "Use of sequential learning for short-term traffic flow forecasting," *Transp. Res. C, Emerg. Technol.*, vol. 9, no. 5, pp. 319–336, Oct. 2001.
- [16] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach," *Transp. Res. C, Emerg. Technol.*, vol. 13, no. 3, pp. 211–234, Jun. 2005.
- [17] W. Zheng, D.-H. Lee, and Q. Shi, "Short-term freeway traffic flow prediction: Bayesian combined neural network approach," *J. Transp. Eng.*, vol. 132, no. 2, pp. 114–121, Feb. 2006.
- [18] Y. Sugiyama, H. Matsubara, S. Myojo, K. Tamura, and N. Ozaki, "An approach for real-time estimation of railway passenger flow," *Quart. Rep. TRTI*, vol. 51, no. 2, pp. 82–88, 2010.
- [19] O. O. Ozerova, "Passenger flows prediction in major transportation hubs," *Sci. Transp. Progress. Bull. Dnipropetrovsk Nat. Univ. Railway Transp.*, vol. 6, no. 48, pp. 72–80, Apr. 2015.
- [20] T. M. Hrushevska, "Research of regularities of passenger flows in the rail suburban traffic," *Sci. Transp. Progress. Bull. Dnipropetrovsk Nat. Univ. Railway Transport*, vol. 5, no. 53, p. 39, Feb. 2015.
- [21] S. Hao, D.-H. Lee, and D. Zhao, "Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system," *Transp. Res. C, Emerg. Technol.*, vol. 107, pp. 287–300, Oct. 2019.
- [22] Z. Guo, X. Zhao, Y. Chen, W. Wu, and J. Yang, "Short-term passenger flow forecast of urban rail transit based on GPR and KRR," *IET Intell. Transp. Syst.*, vol. 13, no. 9, pp. 1374–1382, Sep. 2019.
- [23] E. P. B. George, M. J. Gwilym, C. R. Gregory, and M. L. Greta, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2015.
- [24] L. Asmer, A. Popa, T. Koch, A. Deutschmann, and M. Hellmann, "Secure rail station—Research on the effect of security checks on passenger flow," *J. Rail Transp. Planning Manage.*, vol. 10, pp. 9–22, Aug. 2019.
- [25] W. Wu, F. Zhang, W. Liu, and G. Lodewijks, "Modelling the traffic in a mixed network with autonomous-driving expressways and non-autonomous local streets," *Transp. Res. E, Logistics Transp. Rev.*, vol. 134, Feb. 2020, Art. no. 101855.
- [26] M. Guériaux, F. Armetta, S. Hassas, R. Billot, and N. E. El Faouzi, "Constructivist learning based on multiagent systems. An application to the complex problem of cooperative traffic regulation," *Revue d'Intell. Artificielle*, vol. 32, no. 2, pp. 249–277, 2018.





**ZHUCUI JING** received the Ph.D. degree in computational mathematics from the Chinese Academy of Sciences, China. She is currently an Associate Professor with the School of Economics and Management, Beijing Jiaotong University. Her current research interests include innovation, leadership, and health industry.



**XIAOLI YIN** received the master's degree in probability theory and mathematical statistics from Shanxi University, China. She is currently an Associate Professor with the Business College of Shanxi University. Her current research interests include intelligent transportation systems, wind power prediction, and natural language processing.

...