Traffic Flow Prediction of Expressway Section Based on RBF Neural Network Model

Qun Liu, Zhuocheng Yang, and Lei Cai

Abstract In order to further improve the prediction accuracy of expressway traffic flow, this study proposed an RBF neural network model. Firstly, RBF is used to train the model by using ETC (Electronic Toll Collection) gantry historical data considering the time-varying characteristics of the flow, to ensure the similarity of the flow curve and robustness of the model. Then, taking three typical ETC gantries from thirty gantries of Beijing-Shanghai Expressway in Shandong province as an example, the accuracy of the model is verified by using the historical operation data of them during holidays. The results show that: (1) The flow of ETC gantry section in holidays predicted by RBF is closer to the actual value, and the prediction accuracy is significantly better than that of BP and ELMAN. (2) The MAE is within 75 veh/min, the RMSE is within 6veh/min, and the MAPE is less than 4.5%.

1 Introduction

ETC gantry data are traffic flow section data, accurately record different types of vehicles and key traffic characteristics such as speed and flow. The accuracy of data is more than 99%, compared to other data sources such as traffic survey data, RTMS (Remote Traffic Microwave Sensor) data. The integrity, accuracy and authenticity of ETC gantry data are better than them. The data can help managers and decision makers to make more effective traffic flow prediction.

In the field of traffic flow prediction, scholars have been proposed parameter models, nonparametric models, hybrid models and machine learning/deep learning models [1]. Neural network models are a subclass of nonparametric model, which have widely interconnected structure and effective learning mechanism to simulate

Q. Liu
Shandong High-Speed Construction Management Group Co., Ltd, Jinan 250101, China

Z. Yang (✉) · L. Cai
Beijing GOTEC ITS Technology Co., Ltd, Beijing 100088, China
e-mail: yzc@itsc.cn

L. C. Jain et al. (eds.), *Advances in Wireless Communications and Applications*,
Smart Innovation, Systems and Technologies 299,
https://doi.org/10.1007/978-981-19-2255-8_19

the process of the human brain information processing. Based on a large amount of historical data for training, high prediction accuracy is achieved [2–5].

In order to improve the prediction accuracy and make full use of ETC gantry data information, this study used RBF neural network model to predict traffic flow based on historical data. At the same time, the accuracy of the prediction curve is guaranteed from the time-varying characteristics of the flow. The proposed method has high prediction accuracy and strong robustness, and provides an idea for the prediction of section traffic flow.

2 ETC Gantry System

In 2019, China vigorously promoted ETC technology on 143,000 km of expressways, cancelled 487 toll stations at provincial boundaries of expressways, built 24,588 sets of ETC gantry systems, reconstructed 48,211 ETC lanes. The number of ETC users reached 2.0400 million. Highway infrastructure and management level to achieve qualitative progress by leaps and bounds, obtain a series of surprising results.

According to the overall technical requirements of the expressway provincial boundary toll stations, ETC gantry systems should be set up between each interchange and entrance/exit of expressways. ETC vehicles and MTC (Manual Toll Collection) vehicles realized segmented tolling. Generating transaction flow (or pass certificate), ETC pass record and captured image information (including license plate number and license plate color, etc.) for ETC vehicles, and timely upload to provincial settlement center and ministry network center. For MTC vehicles, read vehicle information in CPC card (including license plate number, license plate color, model information, etc.), calculate the fee and write it into CPC card, form CPC line record, and upload the captured image information to provincial settlement center and ministry network center in time.

Through the detailed study on the content of the data of ETC gantry systems, it can be found that ETC gantry systems can obtain high-precision information such as traffic flow and interval average speed, and can evaluate the real-time traffic running state of expressways. Identify the traffic congestion near gantry or predict the coming traffic congestion is of great important.

3 Overview of RBF Neural Network

RBF (Radial Basis Function) neural network is a typical feedforward neural network. Its characteristic is that the radial basis function is used as the transformation function of the nodes in the hidden layer, so that the hidden layer can convert the low-dimensional input data into the high-dimensional space, and convert the linear non-separable problem in the low-dimensional space into the linearly separable problem in the high-dimensional space. RBF neural network is usually composed of input layer, hidden layer and output layer, and its structure is shown in the Fig. 1 [6, 7].

Vector $X(X = (x_t, x_{t-1}, \ldots, x_{t-n}))$ is the input variable of the network, vector $Y(Y = (\hat{x}_{t+1}, \ldots, \hat{x}_{t+d}))$ is the output variable of the network, vector W represents the connection weight matrix between the input layer and the hidden layer,.

The prediction process of RBF neural network can be expressed as:

$$w_j = \exp\left(-\frac{x_{t-j} - u_j^2}{2\sigma_j^2}\right) \tag{1}$$

$$\hat{x}_{t+d} = \sum w_{jd} w_j \tag{2}$$

In this formula, w_j is the output of the jth node in the hidden layer, x_{t-j} is the traffic flow observation value at $t - j$ moment, $\left\| x_{t-j} - u_j \right\|$ is the normal function, u_j is the center of the Gaussian function, and σ_j is the variance of the Gaussian function. n indicates the number of nodes of the input layer, m indicates the number of nodes of the hidden layer, d indicates the number of nodes of the output layer, w_{jd} indicates the connection weight between the output layer and the hidden layer, and $\hat{x}_{t+d}$ indicates the predicted traffic flow at $t + d$.

The center vector of the radial basis function $u_j = \left[u_{j1}, u_{j2}, \ldots, u_{j(t-n)}\right]^T$. The kernel width σ_j and the connection weights w_{jd} of hidden layer and output layer are parameters of RBF neural network. u_j and σ_j can be determined by FCM clustering algorithm in Eqs. (3) and (4), and the parameter w_{jd} is obtained by gradient descent learning algorithm.

$$u_{jk} = \sum_{i=1}^{n} \mu_{ij} x_{ik} / \sum_{i=1}^{n} \mu_{ij} \tag{3}$$

$$\sigma_j = \sum_{i=1}^{n} \mu_{ij} x_i - u_j^2 / \sum_{i=1}^{n} \mu_{ij} \tag{4}$$

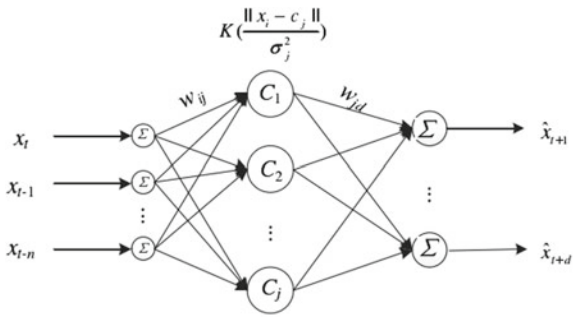


Fig. 1 Neural network prediction principle

In the formula, μ_{ij} represents the fuzzy membership degree of x_i of the sample obtained by FCM clustering algorithm for the jth class, and n represents the training sample size.

Let $\widetilde{x_j} = \varphi\left\|x_{t-j} - u_j\right\|$, $j = 1, 2, \ldots, m$, so

$$\tilde{x} = [\widetilde{x_1}, \widetilde{x_2} \ldots, \widetilde{x_m}]^T \tag{5}$$

The center u_j and the kernel width σ_j of the radial basis function obtained by Eq. (3) and (4) are substituted into Eq. (1) to realize the nonlinear mapping from the input layer to the hidden layer.

Then, start building the model as follows:

Step 1: According to formula (3) and (4), the values of u_j and σ_j are obtained, and the input model $\tilde{x}$ is also created according to formula (5).

Step 2: Introduce ε insensitive loss function.

ε insensitive loss function $L^{\varepsilon}(x, y, f)$ is defined as

$$L^{\varepsilon}(x, y, f) = |y - f(x)|_{\varepsilon} = \max(0, |y - f(x)|_{\varepsilon}) \tag{6}$$

In the formula, $x \in R^m$, $y \in R$.

For the linear model of formula (6), its corresponding ε insensitive loss function can be expressed as:

$$\sum_{j=1}^{n}\left|y_j^o - y_j\right|_{\varepsilon} = \sum_{j=1}^{n}\max\left(0, \left|y_j^o - y_j\right| - \varepsilon\right) = \sum_{j=1}^{n}\max\left(0, \left|p^T\widetilde{x_j} - y_j\right| - \varepsilon\right) \tag{7}$$

In the formula, y_j^o represents neural network output and y_j represents real output.

Step 3: Prediction.

$$y = p^T\varphi(\tilde{x}_{test}) = \lambda\sum_{j=1}^{n}\left(\alpha_j - \alpha_j^*\right)\varphi^T\left(\widetilde{x_j}\right)(\tilde{x}_{test}) = \lambda\sum_{j=1}^{n}\left(\alpha_j - \alpha_j^*\right)\tilde{K}\left(\widetilde{x_j}, \tilde{x}_{test}\right) \tag{8}$$

$$\begin{aligned} &y = [y_1 \ldots y_n]^T, \alpha = [\alpha_1 \ldots \alpha_n]^T, \alpha^* = \left[\alpha_1^* \ldots \alpha_n^*\right]^T, \\ &\tilde{K} = \left[\tilde{k}\left(\widetilde{x_j}, \widetilde{x_i}\right)\right] = \begin{bmatrix} K + \frac{\mu n}{\lambda}I & -K \\ -K & K\frac{\mu n}{\lambda}I \end{bmatrix} \end{aligned} \tag{9}$$

4 Experimental Verification

4.1 Basic Data Description

The data used in this study is the ETC gantry data of Shandong province. Firstly, the protocol and format of ETC gantry data are studied, and the data are initialized to determine the parameters of the model. On this basis, the prediction effect of the model was evaluated by root mean square error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE).

Basic data

In order to verify the validity and correctness of the model, three typical gantries of Beijing-Shanghai Expressway in Shandong Province were selected as the research objects, and the traffic flow during holidays was predicted. Among them, three gantry numbers are G000237011000320070, G000237011000420080 and G000237011000510040, all of which are uplink gantries. In this study, the historical gantry data of 8 days from September 29 to October 6, 2020 were selected to predict the traffic flow on October 7 and 8.

The ETC gantry data selected in this study includes ETC transaction flow (double-chip OBU), ETC traffic record (transaction failure), image flow record and CPC card record, among which the image flow record is the auxiliary data. Take ETC transaction flow (double-chip OBU) as an example, and its sample data are shown in Fig. 2.

ListNo	SerialNo	FlagNet…	FlagRoa…	FlagID	OBUType	OBUMa…	VehicleT…	VehPlate	VehColor	OpTime	VehStat…
G00023…	0137012…	G2	G00023…	G00023…	2	6867B296	13	冀D2E877	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B296	13	冀D2E877	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	蒙D44335	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JC7076	1	2020-10…	27
G00023…	0137012…	G2	G00023…	G00023…	2	6867B2…	13	鲁JD0562	1	2020-10…	27

Fig. 2 An example of ETC transaction flow (two-chip OBU) data

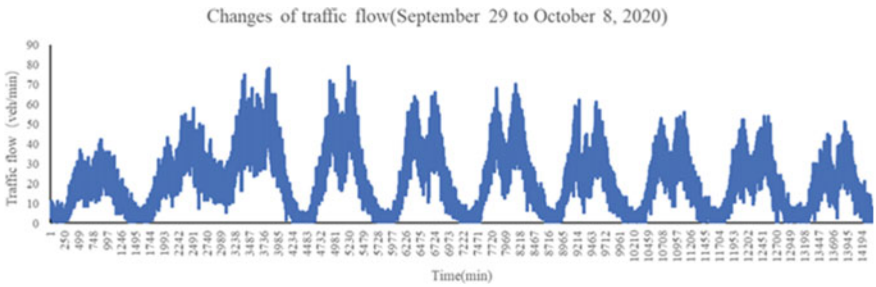


Fig. 3 Gantry G000237011000510040 traffic changes per minute during holidays

Data characteristics

The changes of traffic flow of gantry G000237011000510040 within the range affected by holidays in 10 days is shown in the Fig. 3. The average daily traffic flow value increases from September 29, and reaches its maximum value on October 2. It becomes stable on October 3 and 4, and gradually decreases on October 5, which conforms to the traffic flow rule of National Day holiday.

4.2 Prediction Results

The training process of all models was realized in MATLAB R2020b. The calculation formulas of the prediction result evaluation index are as follows:

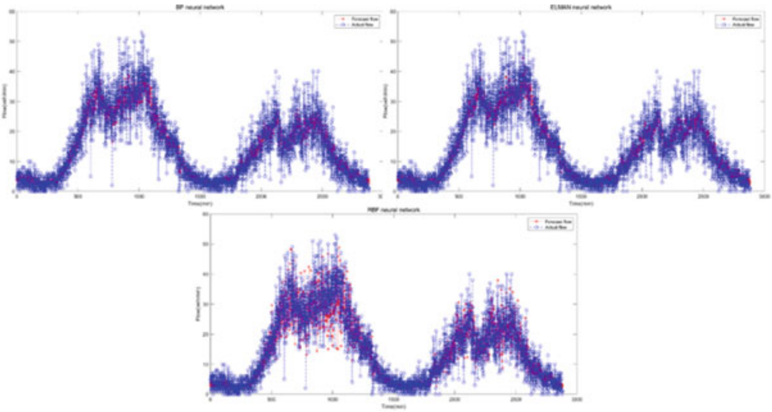
$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{v}_i - v_i\right| \tag{10}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\hat{v}_i - v_i\right|}{v_i} * 100\% \tag{11}$$

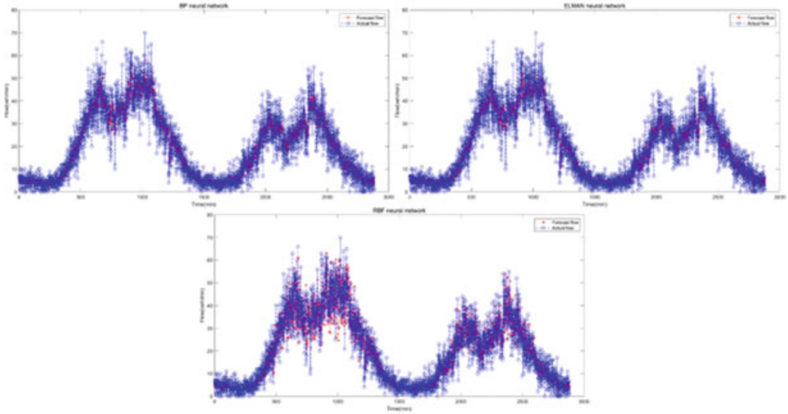
$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{\left(\hat{v}_i - v_i\right)^2}{n}} \tag{12}$$

As shown in this formula, $\hat{v}_i$ indicates the predicted value of traffic flow at time i, and v_i indicates the observed value of traffic flow at time i.

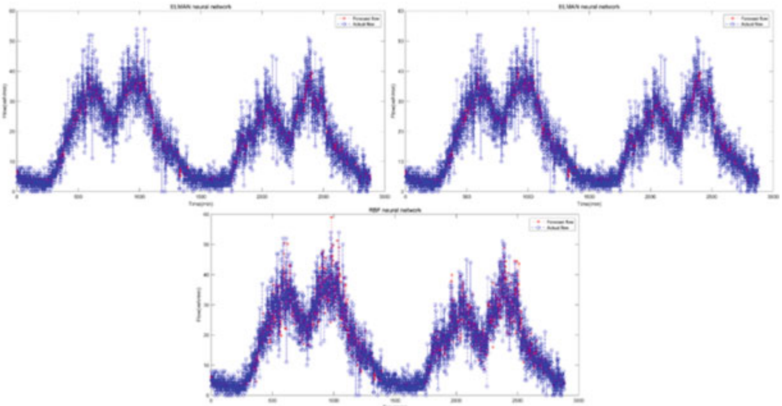
Figure 4 shows a comparison of the traffic flow of three gantries on holidays by using BP, ELMAN and RBF algorithms. It can be seen from the figure that the prediction results of the three gantries obtained by RBF neural network algorithm are better than the other two methods, and the prediction errors are different due to the complexity of the sections where the three gantries belongs. The prediction



(1) G000237011000320070



(2) G000237011000420080



(3) G000237011000510040

Fig. 4 Comparison of holiday flow forecast results of BP, ELMAN and RBF

Table 1 Holiday traffic prediction error

The number of gantries	Algorithm	MAE/ (veh/min)	RMSE/ (veh/min)	MAPE/ (%)
G000237011000320070	BP	84.86	5.19	3.78
	ELMAN	83.06	5.12	3.74
	RBF	74.99	5.07	3.70
G000237011000420080	BP	59.87	5.98	4.41
	ELMAN	58.99	5.96	4.39
	RBF	54.14	5.91	4.36
G000237011000510040	BP	34.51	5.55	4.05
	ELMAN	33.72	5.53	4.04
	RBF	31.78	5.49	4.01

results of G000237011000510040 gantry are significantly better than the other two gantries.

Table 1 is a summary of the average error of traffic prediction obtained by using the three algorithms. It can be seen from the table that the RBF neural network algorithm has the best prediction result among the three algorithms, followed by BP and ELMAN neural network algorithm. The MAE of G000237011000320070 holiday traffic obtained by the algorithm proposed in this study is within 75 veh/min, the MAE of G000237011000420080 is within 55 veh/min, the MAE of G000237011000510040 is within 32 veh/min. The RMSE of the three gantries were all within 6 veh/min, and the MAPE were all less than 4.5%.

5 Conclusion

Considering that ETC gantry data contains a lot of information, this study proposes an RBF neural network algorithm, which uses historical traffic flow data trend to predict the trend of gantry section flow.

(1) RBF neural network algorithm is adopted to predict the changing trend of flow of gantry section on holidays by using historical trend, which increases the prediction robustness of the model.
(2) RBF neural network algorithm is used to predict the flow of gantry section during holidays: three gantries of Beijing-Shanghai Expressway in Shandong province were selected and BP, ELMAN and RBF neural network algorithms were used to predict, which proves the superiority of the proposed algorithm in this study. The MAE were less than 75veh/min, and the RMSE were less than 6veh/min, and the MAPE were less than 4.5%.

References

1. Liu, J., Guan, W.: A summary of traffic flow forecasting methods. J. Highw. Traffic Sci. Technol **03**, 82–85 (2004). (in Chinese)
2. Li, Y., Yu, R., Shahabi, C., et al.: Diffusion convolutional recurrent neural network: data-driven traffic forecasting (2017)
3. Zhang, J., Zheng, Y., Sun, J., et al.: Flow prediction in spatio-temporal networks based on multitask deep learning. IEEE Trans. Knowl. Data Eng. **1**(1) (2019)
4. Mou, L., Zhao, P., Xie, H., et al.: T-LSTM: a long short-term memory neural network enhanced by temporal information for traffic flow prediction. IEEE Access **7**, 98053–98060 (2019)
5. Zhao, L., Song, Y., Zhang, C., et al.: T-GCN: a temporal graph convolutional network for traffic prediction. IEEE Trans. Intell. Transp. Syst. **99**, 1–11 (2019)
6. Guo, S., Lin, Y., Feng, N., et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In: Proceedings of the AAAI conference on artificial intelligence, vol. 33, pp. 922-929 (2019)
7. Song, C., Lin, Y., Guo, S., et al.: Spatial-temporal synchronous graph convolutional networks: a new framework for spatial-temporal network data forecasting. In: Proceedings of the AAAI conference on artificial intelligence, vol. 34(1), pp. 914–921, (2020)