A Deep Prediction Architecture for Traffic Flow with Precipitation Information

Jingyuan Wang[1], Xiaofei Xu[1], Feishuang Wang[1], Chao Chen[2], and Ke Ren[1(✉)]

[1] School of Computer and Information Science, Southwest University, Chongqing, China
wjykim@email.swu.edu.cn, 513178837@qq.com

[2] Online and Continuing Education College, Southwest University, Chongqing, China

Abstract. Traffic flow prediction is an important building block to enabling intelligent transportation systems in a smart city. An accurate prediction model can help the governors make reliable traffic control strategies. In this paper, we propose a deep traffic flow prediction architecture P-DBL, which takes advantage of a deep bi-directional long short-term memory (DBL) model and precipitation information. The proposed model is able to capture the deep features of traffic flow and take full advantage of time-aware traffic flow data and additional precipitation data. We evaluate the prediction architecture on the dataset from Caltrans Performance Measurement System (PeMS) and the precipitation dataset from California Data Exchange Center (CDEC). The experiment results demonstrate that the proposed model for traffic flow prediction obtains high accuracy compared with other models.

Keywords: Traffic flow prediction · Bi-directional LSTM
Deep hierarchy · Precipitation information

1 Introduction

Traffic flow prediction has been long regarded as a critical problem for intelligent transportation systems (ITS) [1]. The aim of traffic flow prediction is to predict the number of vehicles within a given time interval on the basis of the historical traffic information. Traffic flow prediction has an important significance in real-time route guidance and reliable traffic control strategies [2]. Accurate real-time traffic flow prediction can offer information and guidance to road users to optimize their travel decisions and to reduce costs. And according to prediction, the authorities can use advanced traffic management strategies to mitigate congestion. Many traffic flow prediction models have been proposed in the past decades. They can be broadly divided into parametric and nonparametric models.

The traditional methods for traffic flow prediction are parametric approaches. In earlier studies, linear time series models have been widely applied. Time series

Y. Tan et al. (Eds.): ICSI 2018, LNCS 10942, pp. 329–338, 2018.
https://doi.org/10.1007/978-3-319-93818-9_31

methods, such as the autoregressive integrated moving average (ARIMA) [3], were employed to forecast short-term traffic flow. Moreover, some improved ARIMA models like space-time ARIMA [4] and seasonal ARIMA (SARIMA) [5] were also proposed to predict traffic flow. The parametric approach has simple and explicit architecture, which is easy for implementation. However, due to the stochasticity and nonlinearity of the traffic flow, parametric approaches cannot describe traffic flow precisely.

More and more researchers are dipping their toes in the waters of the nonparametric approach, for nonparametric approaches can capture the complicated nonlinearity of the traffic flow and take the uncertainty into consideration. Jin et al. [6] employed support vector regression (SVR) to predict traffic flow. Leshem et al. [7] developed a random forest regression (RF) method. And neural network (NN) models were reported in [8].

The rapid development of intelligent transportation system and data collecting technology makes it possible to get access to huge amount of traffic data as well as environmental data. However, both the typical parametric and nonparametric models tend to make assumptions to ignore additional influencing factors like precipitation. Recent advances in deep learning have enabled researchers to model the complex nonlinear relationships and have shown promising results in computer vision and natural language processing fields [9]. This success has inspired several attempts to use deep learning techniques on traffic prediction problems. Huang et al. [1] incorporated multitask learning (MTL) into deep belief networks (DBN) for traffic flow prediction. Lv et al. [10] proposed a stacked autoencoder (SAE) model. Tian et al. [11] used long short-term memory (LSTM) recurrent neural network to forecast traffic flow. Yao et al. [12] tried to model both spatial and temporal relations. The shallow structure of current deep learning models makes them unable to mine deep features of traffic flow. And ignoring additional influencing factors leads to the fact that they are hard to improve the prediction accuracy.

In this paper, we propose a deep architecture for traffic flow prediction considering precipitation impact. It is worthwhile to highlight the main contributions of our work: (1) We propose a deep bi-directional long short-term memory (DBL) by introducing long short-term memory (LSTM) recurrent neural network, residual connections, deeply hierarchical networks and bi-directional traffic flow. A regression layer is used above the DBL for supervised prediction. (2) We take precipitation factor into consideration when predicting traffic flow. The features of traffic flow under various precipitation conditions can be learned after training using traffic data and precipitation data. In this way, we promote the DBL model to the P-DBL model. (3) We adopt dropout training method to avoid overfitting problem.

The rest of this paper is organized as follows. Section 2 formalizes the problem of traffic flow prediction and introduces long short-term memory (LSTM) model. Section 3 proposes the P-DBL architecture for traffic flow prediction. Section 4 discusses the experiment design and performance of the proposed architecture, and comparison with several selected models. Finally, Sect. 5 is the conclusion.

2 Related Work

2.1 Traffic Flow Prediction

Traffic flow prediction is a typical temporal and spatial process. The traffic flow prediction problem can be stated as follows. The traffic flow of the i_{th} observation point (road, segment or station) at the t_{th} time interval is denoted as $f_{i,t}$. At time t', the prediction task is to forecast the traffic flow $f_{i,t'+1}$ at time $t'+1$, which is based on the traffic flow sequence $F = \{f_{i,t} | i \in O, t = 1, 2, ..., t'\}$ in the past. O is the full set of observation points. The prediction time interval is the interval between time t and $t+1$, which is denoted as $\varDelta t$. According to the length of the prediction time interval, the traffic flow prediction can be divided into three types: long-term, mid-term and short-term traffic flow prediction. Traffic flow prediction has a significant meaning in real-time route guidance and reliable traffic control strategies.

2.2 Long Short-Term Memory Network

Long Short-Term Memory (LSTM) [13] is an effective approach to handle sequential data, which takes advantage of the three multiplicative units in the memory block to determine the optimal time lags dynamically.

$$h_t = H(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{1}$$

$$y_t = W_{hy}h_t + b_y \tag{2}$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \tag{3}$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \tag{4}$$

$$c_t = f_tc_{t-1} + i_tg(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{5}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \tag{6}$$

$$h_t = o_th(c_t) \tag{7}$$

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{8}$$

$$g(x) = \frac{4}{1+e^{-x}} - 2 \tag{9}$$

$$h(x) = \frac{2}{1+e^{-x}} - 1 \tag{10}$$

2.3 Precipitation Impact

It is widely accepted that weather plays a significant role in the performance of the surface transportation system [14]. Certainly, extreme weather events, such as thick fog, can bring traffic to a halt. However, beyond these extreme conditions, more common weather events, such as precipitation, have also been shown to impact traffic conditions. Many researchers have focused on studying the precipitation impact on traffic characteristics such as road capacity and vehicle speed. Both road capacity and vehicle speed will greatly impact the traffic flow. Therefore, it is quite necessary to consider precipitation impact when predicting traffic flow.

Due to the lack of data, many previous scholars have to make assumptions to ignore weather factors like precipitation. However, with the rapid development of intelligent transportation system and data collecting technology, tons of weather data can be collected for traffic flow research. The speed-flow-occupancy relationship under adverse weather conditions was examined by using dummy variables within a multiple regression model, and the results of capacity loss are marginal in light rain but about 15% in heavy rain are indicated [15]. Thus, using precipitation data can make the model be more in line with the reality and gain better performance.

3 Traffic Flow Prediction Architecture: P-DBL

3.1 Deep Bi-Directional Long Short-Term Memory (DBL) Model

The structure of deep bi-directional long short-term memory model is shown in Fig. 1. The input n-length historical traffic flow sequence is denoted as $x^0 = \{x_0^0, x_1^0, x_2^0, ..., x_{n-1}^0\}$, where $x_i^0 (i = 0, 1, 2, ..., n-1)$ is the traffic flow at the t_{th} time interval. $x^i (i = 1, 2, ..., m)$ is the output of the i_{th} layer. BiLSTM is the bi-directional long short-term memory network, where $\overrightarrow{LSTM}$ encodes the input sequence from the start to the end and $\overleftarrow{LSTM}$ encodes the input sequence from the end to the start. Due to the strong ability to handle the sequential data, biLSTM has been successfully applied in natural language processing and image processing. By using biLSTM, the traffic flow information of both directions can be taken into consideration.

In this paper, we use 7 biLSTM layers. The deep hierarchy structure always results in the gradient vanishing problem. To achieve the idea of modeling differences between an intermediate layers output and the targets, we introduce residual connections among the DBL layers in a stack (the red line shown in Fig. 1. Residual connections performed well in the past [16,17]. With residual connections in the DBL model, the equations are as follows.

$$c_t^i, h_t^i = biLSTM^i(c_{t-1}^i, h_{t-1}^i, x_t^{i-1}; \theta^i) \tag{11}$$

$$x_t^i = x_t^{i-1} + h_t^i \tag{12}$$

$$c_t^{i+1}, h_t^{i+1} = biLSTM^{i+1}(c_{t-1}^{i+1}, h_{t-1}^{i+1}, x_t^i; \theta^{i+1}) \tag{13}$$

where c_t^i and h_t^i are the memory states and hidden states of $biLSTM^i$ at the t_{th} time interval for the i_{th} layer, respectively; x_t^i is the input at the t_{th} time interval for the i_{th} layer; θ^i is the set of parameters of $biLSTM^i$ for the i_{th} layer.

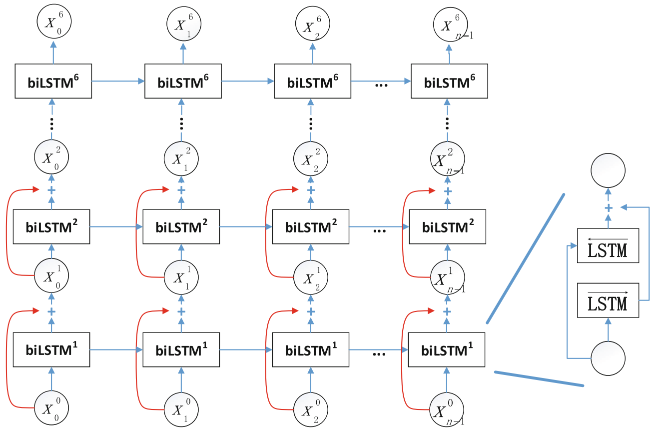

Fig. 1. The structure of deep bi-directional long short-term memory model (Color figure online)

3.2 The Prediction Architecture: P-DBL

As shown in Fig. 2, the prediction architecture mainly consists of four parts: the embedding layer, the DBL, the mean pooling layer and the softmax regression layer. $\{x(0), x(1), ..., x(n-1)\}$ is the input which represents corresponded traffic flow data and precipitation data, and each $x(i)(i = 0, 1, ..., n-1)$ is a piece of corresponded traffic flow data and precipitation data at a time interval encoded by one-hot representation. The corresponded traffic flow data and precipitation data is mapped into a space of same dimension, which is a 128-dimensional vector space. After the DBL encodes the time-aware traffic flow information and precipitation data, a sequence $\{h(0), h(1), ..., h(n-1)\}$ is produced. Then, the mean pooling layer extracts mean values of the sequence over time intervals. Besides, the mean pooling layer makes the features encoded into a vector h. The vector h is fed into the softmax regression layer at the top of the prediction architecture.

To avoid overfitting problem [18] and improve the generalization capability of the model, we adopt the dropout method [19,20] in the embedding layer. The key idea of the dropout method is to randomly drop units (along with their connections) from the neural network during training, which can prevent units from co-adapting too much. During training, dropout samples from numerous different thinned networks. When testing, it becomes easy to approximate the effect of averaging the predictions of all these thinned networks and it can be achieved by a single unthinned network with smaller weights.

4 Experiments and Results

4.1 Experimental Settings

There are mainly two types of traffic flow data in the real world. The first type is the loop detector data, which is collected by sensors on each road, such

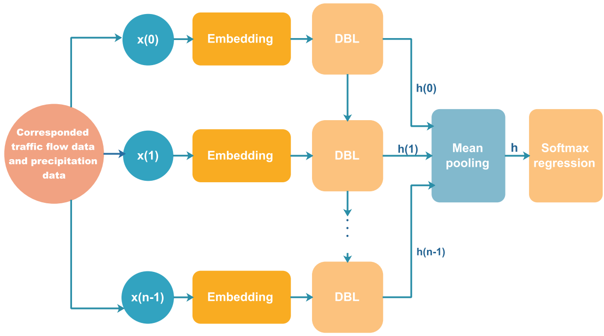

Fig. 2. The structure of the prediction architecture: P-DBL

as inductive loops. The second type is the entrance-exit station data, which is collected at the entrance and exit of a road segment. The prediction task for the first type of data is to forecast the traffic flow on each road or segment, while the prediction task for another type is to forecast the traffic flow in each station, particularly the exit station.

In this paper, we evaluate our model and other comparison models on the dataset from Caltrans Performance Measurement System (PeMS)[1]. The precipitation data is from California Data Exchange Center (CDEC)[2]. PeMS is the most widely used dataset in traffic flow prediction. PeMS constantly collects loop detector data in real time for more than 8100 freeway locations throughout the State of California. Thus, the PeMS dataset is a typical dataset of the loop detector data and the prediction task for PeMS is to forecast the traffic flow on each road or segment. The California Data Exchange Center installs, maintains, and operates an extensive hydrologic data collection network including automatic snow reporting gages for the Cooperative Snow Surveys Program and precipitation and river stage sensors for flood forecasting. We obtain corresponding hourly precipitation data from CDEC. We use the data of five months (from January to May) in 2017 as the training set and the later one month (June) as the testing set.

We used one-hot encoding to transform discrete features (e.g., holidays and weather conditions) and used Max-Min normalization to scale the continuous features (e.g., the average of demand value in last three time intervals).

The models used in comparison experiments are explained as follows.

- **ARIMA:** autoregressive integrated moving average;
- **SVR:** support vector regression [6];
- **DBN:** deep belief network [1];
- **SAE:** stacked auto-encoder [10];
- **LSTM:** long short-term memory recurrent neural network [11];

[1] Caltrans Performance Measurement System (PeMS), http://pems.dot.ca.gov.
[2] California Data Exchange Center (CDEC), http://cdec.water.ca.gov.

- **DBL:** we remove the precipitation data from the proposed architecture;
- **P-DBL:** the proposed prediction architecture.

4.2 Experimental Results

To evaluate the effectiveness of the traffic flow prediction models, we use two performance indexes, which are the Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE). According to them, we can evaluate the relative error and the absolute error. They are defined as follows.

$$MAPE(f,\hat{f}) = \frac{1}{n}\sum_{i=1}^{n}\frac{|f_i - \hat{f}_i|}{f_i} \tag{14}$$

$$RMSE(f,\hat{f}) = \left[\frac{1}{n}\sum_{i=1}^{n}(|f_i - \hat{f}_i|)^2\right]^{\frac{1}{2}} \tag{15}$$

where f is the observation (real) value of traffic flow, and $\hat{f}$ is the prediction value of traffic flow.

The Prediction Accuracy Comparison. To evaluate the prediction accuracy of the P-DBL model, we use P-DBL to predict 30-min interval traffic flow of a whole day (June 1, 2017) using the data collected from No.312139 observation road on D03-5 freeway in California. As shown in Fig. 3, the red line represents the real traffic flow, and the blue line shows the prediction of traffic flow. From Fig. 3, we can notice that the performance of the P-DBL model is quite good during most of the day. Besides, there are mainly three fluctuating periods, which are around 7:00, 12:00 and 19:00 respectively. Those fluctuating periods are all peak traffic periods during which the performance of the P-DBL model is not as stable as other periods. The MAPE of P-DBL is 3.23% and the RMSE of P-DBL is 110.43, which manifests that P-DBL obtains a high prediction accuracy.

Table 1. The results of traffic flow prediction with 30-min time intervalThe results of traffic flow prediction with 30-min time interval

Models	MAPE(%)	RMSE
ARIMA	8.63	171.26
SVR	5.97	143.42
DBN	7.44	165.43
SAE	6.81	151.09
LSTM	4.26	135.32
DBL	**3.87**	**124.31**
P-DBL	**3.23**	**110.43**

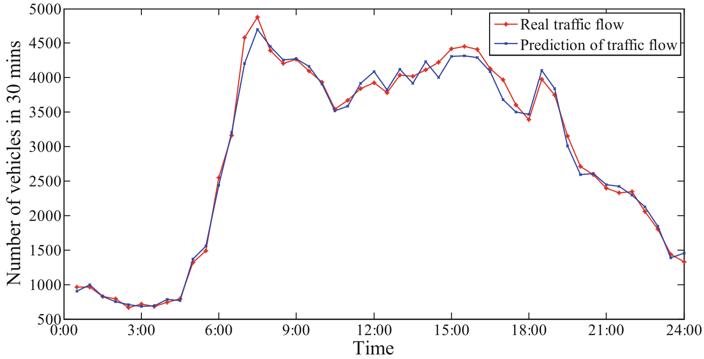

Fig. 3. The comparison between real traffic flow and prediction of traffic flow (Color figure online)

The performance of seven models is tested and the results of them are listed in Table 1. As shown in Table 1, both MAPE and RMSE of P-DBL are lowest among the prediction models. We can notice that deep bi-directional architecture improves the performance of LSTM. And the precipitation information enhances the prediction ability of DBL.

The Effect of Precipitation Information. To verify the effectiveness of the precipitation information, we use DBL and P-DBL to predict 30-min interval traffic flow of a whole day. We evaluate the proposed method with the baseline model on 876 observation roads and achieve the similar results. Next, we will illustrate the details for one randomly selected road and omit the details of others owing to the space limit. The comparison between DBL and P-DBL is shown in Fig. 4. It is immediately visible from the graph that P-DBL outperforms DBL significantly. In our experiment, we notice that a precipitation event occurs from 10:30 to 14:30, and P-DBL predicts more precisely than DBL does in this period.

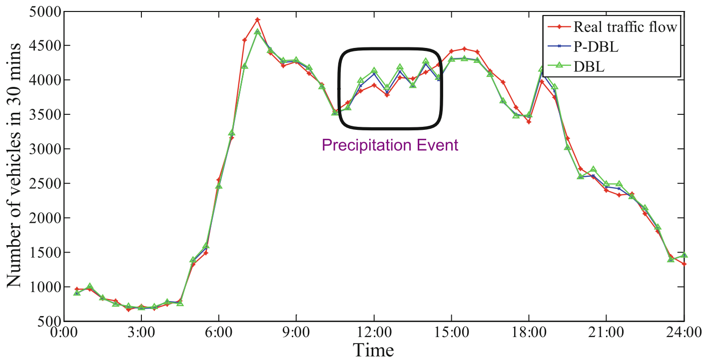

Fig. 4. The comparison between P-DBL and DBL

We find that the MAPE and the RMSE of P-DBL are both lower than those of DBL, which means that the precipitation information is an effective tool to improve the prediction accuracy.

5 Conclusion

To address the traffic flow prediction problem, we propose a deep bi-directional long short-term memory (DBL) model which is able to capture the deep features of traffic flow and take full advantage of time-aware traffic flow data. Moreover, we introduce the DBL model, precipitation information, regression layer and dropout training method into a traffic flow prediction architecture P-DBL. To verify the performance of the P-DBL model, the dataset from PeMS is used in the proposed model and other six comparison models (ARIMA, SVM, DBN, SAE, LSTM, DBL) and the dataset from CDEC is used for precipitation information. In the exprimental results, P-DBL obtains high accuracy.

In the follow-up work, we will take both spatial features and temporal features into consideration. We attempt to use convolutional neural network (CNN) to capture the spatial features and use long short-term memory model (LSTM) to capture the temporal features.

Acknowledgement. This work is supported by "Fundamental Research Funds for the Central Universities" (XDJK2017C027) and "CERNET Innovation Project" (NGII20170516).

References

1. Huang, W., Song, G., Hong, H., Xie, K.: Deep architecture for traffic flow prediction: deep belief networks with multitask learning. IEEE Trans. Intell. Transp. Syst. **15**(5), 2191–2201 (2014)
2. Abadi, A., Rajabioun, T., Ioannou, P.A.: Traffic flow prediction for road transportation networks with limited traffic data. IEEE Trans. Intell. Transp. Syst. **16**(2), 653–662 (2015)
3. Ahmed, M.S., Cook, A.R.: Analysis of freeway traffic time-series data by using Box-Jenkins techniques (1979)
4. Kamarianakis, Y., Vouton, V.: Forecasting traffic flow conditions in an urban network: comparison of multivariate and univariate approaches. Transp. Res. Rec. **1857**(1), 74–84 (2003)
5. Williams, B.M., Hoel, L.A.: Modeling and forecasting vehicular traffic flow as a seasonal arima process: theoretical basis and empirical results. J. Transp. Eng. **129**(6), 664–672 (2003)
6. Jin, X., Zhang, Y., Yao, D.: Simultaneously prediction of network traffic flow based on PCA-SVR. In: Liu, D., Fei, S., Hou, Z., Zhang, H., Sun, C. (eds.) ISNN 2007. LNCS, vol. 4492, pp. 1022–1031. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72393-6_121
7. Leshem, G.: Traffic flow prediction using adaboost algorithm with random forests as a weak learner. Enformatika **193** (2011)

8. Chan, K.Y., Dillon, T.S., Singh, J., Chang, E.: Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg-Marquardt algorithm. IEEE Trans. Intell. Transp. Syst. **13**(2), 644–654 (2012)
9. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436 (2015)
10. Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.Y.: Traffic flow prediction with big data: a deep learning approach. IEEE Trans. Intell. Transp. Syst. **16**(2), 865–873 (2015)
11. Tian, Y., Pan, L.: Predicting short-term traffic flow by long short-term memory recurrent neural network. In: IEEE International Conference on Smart City/SocialCom/SustainCom, pp. 153–158 (2015)
12. Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J.: Deep multi-view spatial-temporal network for taxi demand prediction. arXiv preprint arXiv:1802.08714 (2018)
13. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735 (1997)
14. Wang, Y.Q., Jing, L.: Study of rainfall impacts on freeway traffic flow characteristics. Transp. Res. Procedia **25**, 1533–1543 (2017)
15. Ibrahim, A.T., Hall, F.L.: Effect of adverse weather conditions on speed-flow-occupancy relationships (1994)
16. Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K.: Google's neural machine translation system: bridging the gap between human and machine translation (2016)
17. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Computer Vision and Pattern Recognition, pp. 770–778 (2016)
18. Hawkins, D.M.: The problem of overfitting. Cheminform **35**(19), 1 (2004)
19. Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing co-adaptation of feature detectors. Comput. Sci. **3**(4), 212–223 (2012)
20. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**(1), 1929–1958 (2014)