Short-Term Traffic Flow Prediction of Highway Network Based on Network-Constrained Lasso and NN

Cong Wu[1], Zhaozheng Chen[2] and Xiaofei Li[1]

[1]*Engineering Research Center of Wideband Wireless Communication Technology, Ministry of Education, Nanjing University of Posts and Telecommunications, 210003 Nanjing, China*
[2]*Jiangxi Expressway Networking Management Center, 330036 Nanchang, China*

Abstract. Accurate and timely traffic flow prediction is important for the successful deployment of intelligent transportation systems. Most of existing methods have not made good use of information from adjacent sections to analyse the trends of the object section. A new method for traffic flow prediction of highway network, namely network-constrained Lasso (Least absolute shrinkage and selection operator) and Neural Networks, was proposed. Unlike existing methods, our approach incorporated all the spatial and temporal information available in a highway network to carry our short-term traffic flow prediction for the objective section. To capture the spatial correlation of traffic network, the Laplacian matrix was introduced to describe the highway network structure. Subsequently, a network-constrained Lasso method was applied for sparse variable selection. With the extracted historic and real-time data, the back propagation neural networks were implemented to predict traffic flow at different time intervals in future. The experimental results verified that the proposed method could achieve above 90% average accuracy in the 30-minutes speed predictions for 78 road sections.

1 Introduction

Real-time and accurate traffic flow prediction is the successful deployment of advanced traffic control and dynamic route guidance which can reduce road congestion and improve mobility of transportation. Due to the dynamic and stochastic characteristics of traffic flow, short-term prediction has become one of the difficulties of Intelligence Transportation System.

Traffic researchers have conducted substantial studies, e.g., ARIMA[1], time series[2], Kalman[3], Neural Networks (NN)[4, 5] and SVR[6]. Nevertheless, most of them have not achieved satisfactory performance in traffic network, because they didn't make good use of spatio-temporal correlation between multiple sections.

In order to reveal the potential influence of traffic flow changes between sections of highway network, some researchers have made attempts to increase the prediction accuracy. S Sun [7] and E Castillo [8] performed a short-time traffic flow prediction based on Bayesian network, using space structure information of adjacent sections in highway. B Ghosh [9] reported an time-series analysis, which was propitious to analyze the relationship of traffic flow network. In addition, Y Kamarianakis [10]and C Hu [11] have taken use of topology and distance information between multiple section, and performed short-time traffic flow prediction by ARIMA and Hybrid Process Neural Network. However, the prediction algorithm mentioned above were

mainly based on prior knowledge and mass calculation. Consequently, those algorithms haven't obtained an ideal prediction result in real-time systems, because of their complexity in describing substantive characteristics of the high dimensional traffic data.

In this paper, the Laplacian matrix was introduced to depict the highway network structure. Subsequently, a network-constrained least absolute shrinkage and selection operator (Lasso) method was implemented as the penalty function for sparse variable selection. As neural network has been proved to do well in complex and nonlinear system, the back propagation (BP) neural networks with the extracted data were used to predict nonlinear traffic flow at different time intervals in future.

2 Variable selection

2.1 Laplacian matrix representation of spatial correlation

There were many influencing factor of traffic flow between sections of highway network. From the perspective of qualitative analysis, the relationship was more significant while a section was adjacent and close to another. With the increase of distance and converged traffic flow from other sections, the relationship was gradually diminished.

In this work, we used a weighted graph $G = (V, E, W)$ to describe the highway network structure, where V is a set of vertices corresponding to the sensors (e.g. loop detectors) of independent road sections, $e(u, v)$ is the set of edges indicating that the section u and v are linked, and $w(u, v)$ is the set of weights of the edges, which represent the spatial correlation governed by the network topology. Note that the value of $w(u, v)$ was calculated according to the number and location of upstream and downstream sections in our work.

In graph theory, Laplacian Matrix L, also known as Admittance Matrix or Kirchohoff Matrix, is a matrix form for graph G. As in [12], the element $L(u, v)$ in the normalized Laplacian matrix L was defined in equation (1), where the vertices degree was defined $d_v = \Sigma_u w(u, v)$.

$$L(u,v)= \begin{cases} 1 - w(u,v)/d_u & \text{if } u = v \text{ and } d_u \neq 0 \\ -w(u,v)/\sqrt{d_u d_v} & \text{if } u = v \text{ are adjacent} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

2.2 Laplacian matrix representation of spatial correlation

Compared with classical variable selection methods, such as subset selection, the Lasso is computationally feasible for high-dimensional data. The selection procedure is continuous and hence more stable than the subset selection which is discrete[13]. The traditional penalty regression can be formulated as follows in Equation (2), as a combination of regularized least squares and a penalty term[14].

$$\hat{\beta} = \arg\min_\beta \left\{ \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i \beta_i)^2 + \sum_{j=1}^{p} \lambda |\beta_j| \right\} \quad (2)$$

In this work, a network-constrained penalty was constructed by the Lasso penalty and a penalty induced by the Laplace matrix of the graph[15]. A network-constrained penalty function was defined in Equation (3), where $|\beta|_1 = \sum_{j=1}^{p} |\beta_j|$ was the L1-norm.

$$J(\lambda_1, \lambda_2, \beta) = (y - \mathbf{X}\beta)^T (y - \mathbf{X}\beta) + \lambda_1 |\beta|_1 + \lambda_2 \beta^T L\beta \quad (3)$$

In Equation (3), the tuning parameters λ_1, λ_2 control the amount of regularization for sparsity and smoothness. When $\lambda_2 = 0$, the estimate reduces to the traditional Lasso. It is easy to verify[16] that this matrix L is nonnegative definite and can be decomposed as $L=SS^T$. Equation (3) can be rewritten as Equation (4).

$$J(\lambda_1, \lambda_2, \beta) = (y - \mathbf{X}\beta)^T (y - \mathbf{X}\beta) + \lambda_1 |\beta|_1 + \lambda_2 \sum_{u \sim v} \left(\frac{\beta_u}{\sqrt{d_u}} - \frac{\beta_v}{\sqrt{d_v}} \right)^2 w(u,v) \quad (4)$$

The network-constrained estimation $\hat{\beta}$ was defined as the minimizer of Equation (4).

$$\hat{\beta} = \arg\min_\beta \left(J(\lambda_1, \lambda_2, \beta) \right) \quad (5)$$

As in [17], the network-constrained regularization problem was solved by a similar efficient computation procedure.

3 Prediction model with network-constrained Lasso and NN

As shown in Figure 1, we adopted Laplacian matrix to depict the traffic network topology and spatial correlation, and then applied network-constrained Lasso for variable selection with historical data. After finding the most helpful sections which have great relativity to the objective section, we constructed many BPNNs to predict traffic flow in a region of highway network. In Figure 1, n and m were the number of T, time interval of 5 minutes.
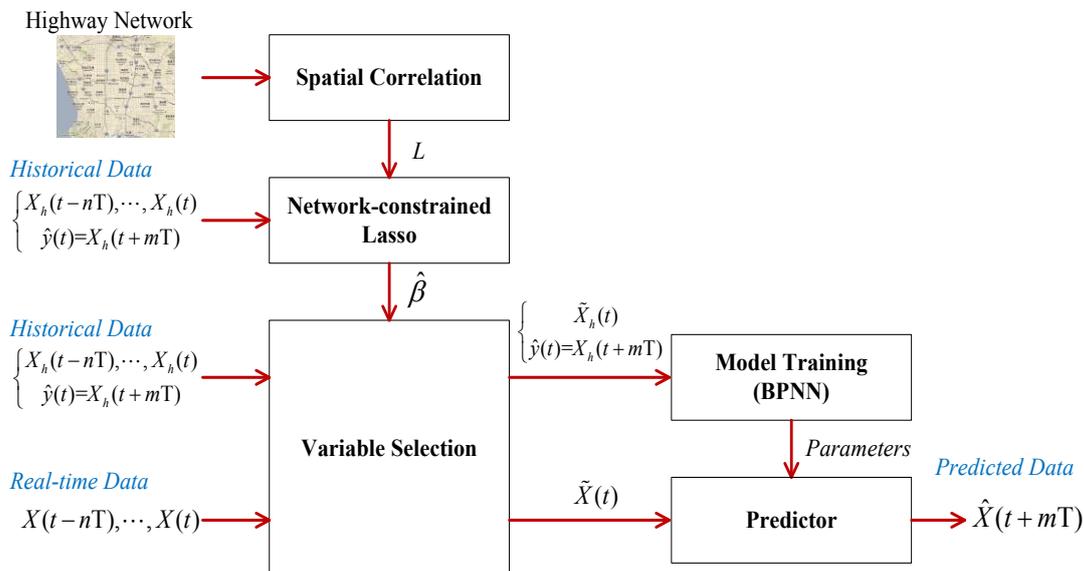
Figure 1. Prediction flowchart with network-constrained Lasso and NN.

4 Application and experimental results

4.1 Data description

The Freeway Performance Measurement System (PeMS) of California collects and processes the data of loop detectors every 30s such as flow, occupancy and speed. This paper focused on one region of Los Angeles and selected the average traffic flow data in 5 minutes' interval from 2013/3/4 to 2013/3/31. A total number of 78 loop detectors in this freeway network were listed in Figure 2, some of which locate near the crossways of traffic network. There were 8064 records in each detector with heavy, medium, and low traffic loads.
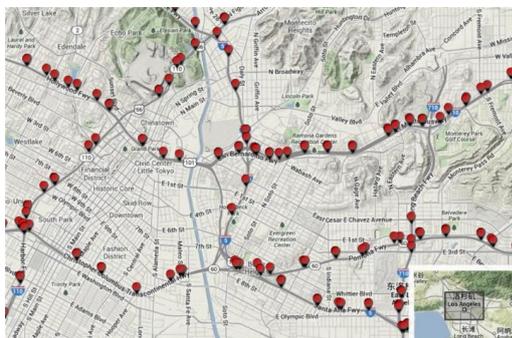

Figure 2. The map of traffic network and loop detectors.

4.2 Comparison of 15-min traffic flow prediction

In this part, we used the average speed data collected in 5 minutes as the input $X(t)$, and predicted average speed $X(t+3T)$ in next 15 minutes with NN and the proposed method. Results were shown in Table 1 and Table 2, corresponding to 15 sections far away from crossways and 15 sections near crossways in the network region. In this paper, Mean Absolute Percent Error (MAPE) was calculated to evaluate the effectiveness of NN and the proposed model.

In Table 1 and Table 2, we could also find that the proposed method proved to be more accurate than NN model for the short-term prediction of average speed. With the proposed method for 15-min prediction, the average accuracy of sections far away from crossways, and sections near crossways was 90.9%, and 93.6%, respectively. Meanwhile, the variable number of the proposed method had a large drop from 78 to 23.0, and 16.3, respectively.

4.3 Comparison of traffic flow prediction from 5-min to 60-min

With the same input X(t), NN model and the proposed method was used to predict 5-min, 10-min, 15-min, 20min, 30-min and 60-min traffic flow. As shown in Table 3, for 78 loop detectors, the average accuracy of the proposed method was 96.7%, 95.3%, 94.3%, 93.2%, 91.2%, and 87.6%, respectively. The maximum average

prediction accuracy improvement of the proposed method was up to 4.6% compared with NN.

Table 1. Performance comparison of sections near crossways.

Section	NN		The proposed method	
	Variable Number	MAPE	Variable Number	MAPE
718173	78	16.9%	14	6.3%
773281	78	14.9%	7	7.8%
717264	78	15.0%	2	7.6%
716942	78	15.0%	21	7.7%
763237	78	14.3%	27	10.4%
717047	78	14.4%	53	12.9%
718045	78	17.0%	21	8.6%
760650	78	12.2%	29	7.4%
773024	78	9.9%	15	5.1%
716946	78	12.7%	28	5.9%
716951	78	18.3%	46	10.8%
772501	78	12.5%	43	8.5%
716940	78	14.2%	5	13.5%
716939	78	15.4%	5	13.7%
718166	78	15.9%	29	9.8%
Average	**78**	**14.6%**	**23.0**	**9.1%**

Table 2. Performance comparison of sections far away from crossways.

Section	NN		The proposed method	
	Variable Number	MAPE	Variable Number	MAPE
718421	78	11.6%	24	5.6%
717262	78	6.5%	10	3.1%
717265	78	8.1%	10	6.9%
717268	78	8.7%	14	5.6%
716069	78	6.9%	6	4.6%
716328	78	9.0%	8	4.9%
773062	78	9.6%	10	5.8%
737320	78	7.8%	21	6.6%
717052	78	6.9%	35	5.8%
716943	78	8.3%	8	7.9%
716944	78	17.7%	27	12.0%
771667	78	9.4%	20	4.3%
718371	78	10.0%	12	8.0%
717445	78	13.0%	36	8.6%

717263	78	10.0%	4	6.2%
Average	**78**	**9.6%**	**16.3**	**6.4%**

Table 3. Performance comparison of predicting different time intervals in future.

Prediction	NN		The proposed method	
	Variable Number	MAPE	Variable Number	MAPE
5-min traffic flow $X(t+T)$	78	4.7%	3.0	3.3%
10-min traffic flow $X(t+2T)$	78	6.8%	9.4	4.7%
15-min traffic flow $X(t+3T)$	78	8.8%	20.9	5.7%
20-min traffic flow $X(t+4T)$	78	10.6%	24.3	6.8%
30-min traffic flow $X(t+6T)$	78	12.8%	34.3	8.8%
60-min traffic flow $X(t+12T)$	78	17.0%	44.5	12.4%

4.4 Comparison of traffic flow prediction with historic data

Usually historic data played an important role in traffic flow prediction. In this part, we predicted 15-min traffic flow with real-time data and historic data at previous time intervals, i.e., X(t-T), X(t-2T). From the results in Table 4, it is apparent that high accuracy result was not due to the variable number but owing to spatio-temporal correlation of traffic network.

Table 4. Performance comparison of predicting different time intervals in future.

Prediction	NN		The proposed method	
	Variable Number	MAPE	Variable Number	MAPE
$X(t)$	78	8.8%	20.9	5.7%
$X(t), X(t-T)$	156	9.5%	33.5	5.9%
$X(t), X(t-T), X(t-2T)$	234	10.3%	34.3	5.8%

5. Conclusion

According to large amount of actual data testing and comparison mentioned above, we proposed a traffic flow prediction method for highway networks. The advantage of the proposed method is mainly due to its simplicity and effectiveness. The simplicity enables its wide applications in real-time traffic management systems.

For future work, we will investigate deep learning algorithms with Lasso for traffic flow prediction and apply these algorithms on different public open traffic data sets to verify their effectiveness.

Acknowledgement

References

1. BM Williams, LA Hoel, J. Transp. Eng. **129**, 6 (2003)
2. T Thomas, W Weijermars, EV Berkum, IEEE T. Intell. Transp. **11**, 1 (2010)
3. J Guo, B Williams, Transport. Res. Rec. **2175**, 1 (2010)
4. X Jiang, H Adeli, J. Transp. Eng. **131**, 10 (2005)
5. KY Chan, TS Dillon, J Singh, E Chang, IEEE T. Intell. Transp. **13**, 2 (2012)
6. M Castro-Neto, YS Jeong, MK Jeong, LD Han, Expert Syst. Appl. **36**, 3 (2009)
7. S Sun, C Zhang, G Yu, IEEE T. Intell. Transp. **7**, 1 (2006)
8. E Castillo, M Nogal, JM Menéndez, SS Cambronero, P Jimenez, IEEE T. Intell. Transp. **13**, 2 (2012)
9. B Ghosh, B Basu, M O'Mahony, IEEE T. Intell. Transp. **10**, 2 (2009)
10. Y Kamarianakis, P Prastacos, Transport. Res. Rec. **1857**, 1 (2003)
11. C Hu, K Xie, G Song, T Wu, *11th International IEEE Conference on Intelligent Transportation Systems* (Beijing, China, 2008)
12. FRK Chung, *Spectral graph theory* (American Mathematical Soc., 1997)
13. Y Kamarianakis, W Shen, L Wynter, Appl. Stoch. Model. Bus., **28**, 4 (2012)
14. S Sun, R Huang, Y Gao, J. Transp. Eng. **138**, 11 (2012)
15. C Li, H Li, Bioinformatics, **24**, 9 (2008)
16. C Li, H Li, *The annals of applied statistics* **4**, 3 (2010)
17. H Zou, T Hastie, R Tibshirani, J. Comput. Graph. Stat., 15, 2 (2006)