Proceedings of the 2023 International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/10/20230183

Review on short-term traffic flow prediction methods under big data

Shiyao Tang

School of Traffic and Transportation, Beijing Jiatong University, Beijing, China.

21252018@bjtu.edu.cn

Abstract. The goal of traffic forecast is to predict the related traffic situation in the future according to the historical concept. The predicted angle can be divided into short-term prediction and long-term prediction. This method can be used to solve the increasingly serious urban traffic congestion problem, and researchers have proposed a deep learning model to help decision makers in the field of traffic control. It has made great contributions to improving future road capacity and optimizing intelligent transportation services. In this paper, short-term traffic forecast related documents under big data are helpful after sorting out, and the traffic flow data characteristics of intelligent transportation system and related correction methods are analysed. Secondly, it then classifies the applications involved by big data algorithm calculation according to various related principles, and summarizes the prediction accuracy, computational complexity, applicable interval and computation time of each type of algorithm in an overview. According to the relevant data, the combination forecasting model is effectively diversified, and combined with the related combination forecasting, I hope to improve the forecasting accuracy of the future development prospect.

Keywords: short-term traffic flow forecast, forecast models, intelligent transport system.

1. Introduction

From another point of view, the traffic time series has strong dynamic time performance. According to his frequent events, for example, when the traffic is in rush hour, it will produce different high-scoring time series, which will bring some trouble to traffic prediction. On the contrary, because of its complexity, the sensors on the traffic network have certain spatial correlation performance, which can be divided into three main stages: short-term, long-term and medium-term [1,2]. The first stage is the unbalanced demand forecasting stage. This stage is mainly to investigate and analyze the existing traffic system and land use status, and seek to explain the traffic demand more accurately under unbalanced conditions through the "evaluation, correction and verification" workflow.

The second stage is the forecasting stage within the balanced demand. In this stage, the traffic is considered to change again because of the psychological characteristics of its passengers, so there is a more accurate balance prediction method. The third stage focuses on the behavioral traffic demand forecasting stage, and there are two main classification directions in different community studies. First, it is through knowledge-centered research, and second, it can focus on big data. In terms of traffic and operational research, because of the knowledge method, it attempts to carry out the traffic calculation network mode generated by queuing theory to simulate the different behaviors of drivers in traffic. With

DOI: 10.54254/2755-2721/10/20230183

the increase of traffic data, data-centered machine learning traffic prediction has been put into an important research direction [3].

Through this short-term traffic flow prediction research, the theoretical concept of traffic flow is put forward, and in the second part, the traffic flow data correction method and related characteristics were introduced. In the third part, the big data traffic prediction models are divided into five categories, including the prediction of statistical analysis, the model of nonlinear theory and the simulation prediction model. This paper focuses on predicting the related content challenges faced by traffic, and discusses the modeling methods in practice or space. Finally, it predicts the important direction of future research.

2. Features of traffic flow big data and its correction methods

With the widespread deployment of sensors such as radio frequency identification, road induction coils, and electronic card cutters on urban roads, and the gradual improvement in the mastery of automatic traffic data collection technology, the in-depth study of urban traffic characteristics found through the use of automatically collected mass data has become a novel trend in the current development [4].

The data used in existing short-term forecasting studies include current historical and observed values [5,6]. Big traffic data is significantly different from traditional traffic environment data in many ways.

The first is that it is fast and data-rich, with a wide and very diverse range of data sources, but it also potentially suffers from data loss, errors, and redundancy. The second is that it is very rich in data value, with significant characteristics of spatial, temporal and historical dimensions. This makes it necessary to have a precise control on the quality of the data before analysing and using it to truly ensure the accuracy and reliability of the traffic data.

Second, short-term traffic forecasting needs to have better real-time performance. First, forecasting can provide more accurate data services for traffic routing and control, as well as other methods, but it needs to be done quickly to ensure timely and accurate prediction of traffic flow for the next cycle. Second, the traffic flow system provides researchers with an easier way to compare forecasting models and assess the accuracy of data by collecting and transmitting data in real time, facilitating continuous correction of forecasts and thus achieving an optimization effect on the forecast results. The process effectively integrates a series of forecasting methods with artificial intelligence techniques through a series of inference, judgment and learning to create a short-term traffic flow forecasting system capable of meeting various requirements of traffic flow forecasting.

Third, by using the three-parameter relationship of traffic flow and the theory of traffic flow balance, as well as the queering method to distinguish the erroneous data. For the identified erroneous data, deletion is usually adopted and the data is repaired and restored by data recovery methods. At this stage, most of them consider the change characteristics of historical data before sampling models and build models based on this to make inferences about the lost data. However, due to some special reasons, such as traffic accidents, weather, etc., the traffic flow distribution often changes as well, which has an impact on the recovery accuracy of traditional models.

A study proposed that the shortcomings of using only historical data in the traditional restoration model can be further improved effectively by considering the "time-space-period" data replenishment method [7]. Because the process of traffic flow is highly uncertain and complex, its state characteristics are different in different road environments and spatial locations. According to the relevant surveys, any traffic flow prediction method has a certain range of adaptation and application conditions, and none of them can show the absolute superiority of all methods [8]. Many forecasting models require the determination of many parameters before they start, and the model structure is determined by calibrating the model before the actual forecasting, by making a finalization of the parameters based on real-time or historical data [8]. The reliability of these models will be further effectively improved when the external and internal traffic flow factors are stable.

DOI: 10.54254/2755-2721/10/20230183

3. Short-term traffic flow forecasting model

Many short-term traffic flow forecasting models have been developed in the existing literature. There are five forecasting methods with certain research results and applications in short-term traffic flow forecasting, which are statistical analysis-based forecasting models, nonlinear theoretical models, simulation-based forecasting models, artificial intelligence models, and mixed models, shown on figure 1.

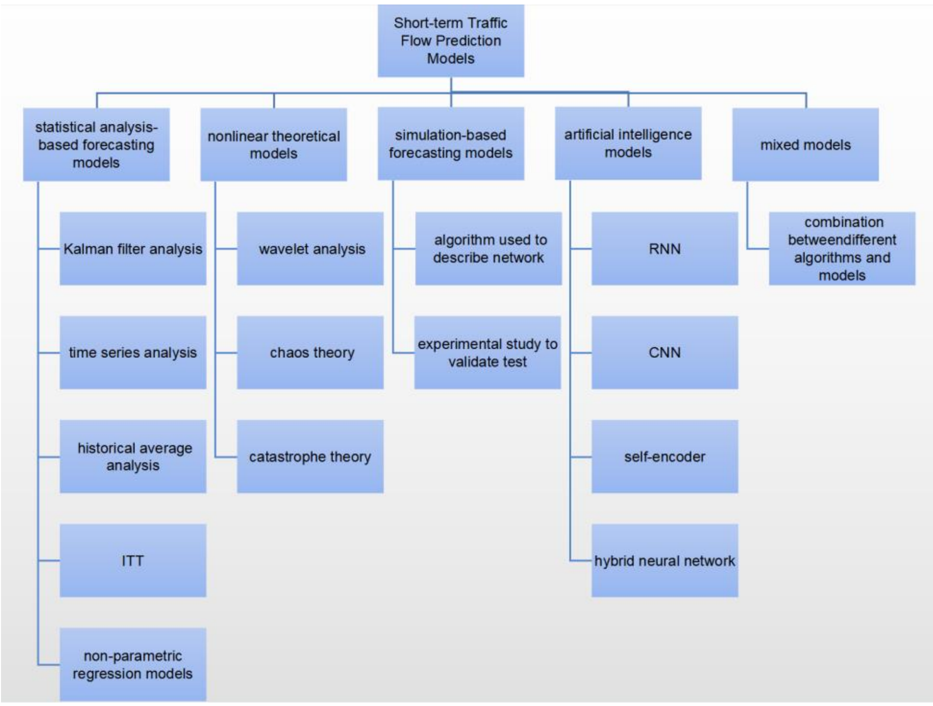

Figure 1. Classifications of short-term traffic flow prediction models.

3.1. Forecasting model based on statistical analysis

Based on the prediction model used in statistical analysis, statistical methods and mathematics are used as effective methods to deal with past traffic data. The basic premise of this method is to assume that historical data and future prediction have the same characteristics. Five main models can be classified, namely Kalman filter analysis and forecasting, time series analysis and forecasting, historical average analysis and forecasting, instantaneous travel time (ITT) and non-parametric regression models.

Firstly, the instantaneous travel time method is used as a prediction for the next measurement by using the actual measurement. It is characterized by being very fast and does not use calculations at all. Therefore, in most cases, the prediction accuracy of this approach is poor, especially when the prediction range is very long. The historical averaging method is the original method for short-term traffic flow forecasting. This method was originally used to calculate the average value of data for the whole period to obtain the forecast results. Although the historical averaging method is relatively convenient and simple, the prediction accuracy is not optimistic and performs poorly for dealing with unexpected situations and complex traffic conditions; whereas a static forecasting system would be more suitable for this method in comparison [9]. The time series method is used to forecast the next traffic state by applying real-time and past traffic data and modeling the traffic flow based on a scalar data model (time series), which mainly includes moving average, autoregressive, autoregressive moving average and autoregressive integrated moving average (ARIMA). Based on ARMA and ARIMA, further generalizations of Arfima (moving average of long-lived autoregressive scores and integrals) and VAR (vector autoregressive model vector autoregressive) have been promoted [10]. The Kalman filter model [11], the state space of linear stochastic systems is a way to describe filtering. It obtains the best estimation result of state variables and filtering noise results through algorithms. Zhang Chunhui has developed a short-term passenger flow forecasting model based on Kalman filter model, after which the resulting prediction error is smaller compared with that of artificial neural networks.

However, the Kalman filter model does not effectively handle contingencies, and the adjustment weights need to be reset each time the equations are calculated.

3.2. Nonlinear theoretical models

He Guoguang finally discussed the feasibility of this method in the field of short-term traffic flow forecasting, and showed the corresponding solutions to the public, and conducted simulation experiments [12]. Oikawa used wavelet analysis to predict short-term traffic flow [13]. After introducing the hidden Markov model, the performance and prediction accuracy of the model have been greatly improved. Forbes applied catastrophe theory in his research, and assumed that the traffic system was disastrous by using the traffic flow model of expressway [14].

The accuracy of models based on statistical analysis is lower than the prediction accuracy of nonlinear theoretical models, especially because chaos theory still has a large potential for development, while most nonlinear theoretical models are computationally intensive and complex to operate. Therefore, this model is more suitable for complex and possible traffic systems.

3.3. Prediction models based on simulation

There are many similar applications of traffic forecasting in transportation, which are closely related to it. At present, the most popular nonlinear correlation modeling method can transmit neural networks. However, due to the long-term over-reliance on modeling and the low training efficiency of RNN, it is necessary to capture a long-term and time-consuming micro-inversion technology for this method, and make a certain fusion of traffic flow model and policy software model to predict and pay attention to their dynamic flow. which is less suitable for large-scale real-time prediction of complex traffic systems. Experimental research, analysis, validation and prediction observation of small intersections are more suitable for this method. The simulation model is more suitable for the prediction of urban rail traffic flow, such as metro and light rail.

3.4. Artificial intelligence models

Recurrent neural network modeling methods RNN [16] are widely used in text [16,17] classification, speech recognition, traffic flow prediction [18] LSTM and end-to-end. Among them, LSTM updates and continuously transmits the temporal trend features of the input sequence in the form of hidden cells, and the storage cells retain the historical features, which effectively overcome the problem of gradient disappearance; however, LSTM is more difficult to capture the periodicity of time series data, and thus is more adaptable to short-term traffic prediction scenarios. A traffic flow prediction model based on the integration of standard convolution and attention mechanisms was designed by fusing adjacent time series information and applying the dependence of traffic flow features in different time series to update future time series features.

The prediction of RNN's agent trajectory has changed from single to multi-modal trajectory, the prediction scene is no longer single, the data set contains more and more information, but there are some deficiencies in modeling spatial relations and image data. With the development of technology, the trajectory prediction method based on LSTM will be more mature.

CNN is a deep neural network, which consists of input layer, line layer, pool layer, full connection layer and output layer. Generally speaking, the greater the standard deviation, the more difficult it is to predict the traffic situation in this place. It is controversial that the location and time with large errors, such as busy intersections during rush hours, are more important to predict [19]. Recently, the spatial correlation in traffic network is captured from two-dimensional spatiotemporal traffic data by using CNN. CNN learns the spatial features of different regions and extracts the spatial correlation to predict traffic flow. There is literature that CNN was used to predict short-term vehicle trajectories, and the surrounding environment of a single participant was encoded as a raster image input, but the output trajectory was unitary [20]. The CNN trajectory prediction method is mainly used to process image features, agent trajectories, simulate agent interactions and driving scenes, and output them as occupation maps, this is necessary for an agent's trajectory to change over time.

Auto-encoder as a learning algorithm is unsupervised in nature. The concept of auto-encoder was introduced by Rumhar et al [21], who considered auto-encoder as a data compression algorithm by using an encoder to compress the data and decompress it by a decoder. In the encoding period, high-dimensional data is mapped into low-dimensional data, and in this way the amount of data is reduced, and in the decoding period, the reproduction of the input data is achieved. A generic accessible workflow has been proposed to obtain large-scale traffic congestion data and to build an image analysis-based traffic congestion dataset [22]. Further research has been done to design and propose layer recursive autoencoder (LRA) for predicting a wider range of traffic, in this paper, a three-layer stacked automatic encoder (SAE) architecture is used to obtain the time-flow correlation of three different time scales. For each output of different time scales, a special neural network is used to predict the time-flow correlation.

Hybrid neural network is a kind of synthetic model. According to different traffic conditions, different forecasting methods are combined and the advantages of each model are synthetically used to forecast. It is precisely because the previous prediction models have their own conditions and advantages and disadvantages, some scholars have proposed an autoregressive integrated moving average model and a Kalman filter combination real-time road state prediction model [23]. Through the training process of historical road traffic data, the optimal parameters of the model are determined, and the difficulties in predicting the complex traffic flow data with a single model are solved, and the prediction accuracy is improved. An improved LSTM network and autoregressive integrated moving average model (SDLSTM-ARIMA) have been proposed and applied to an embedded system, realized the real-time distributed traffic flow prediction and calculation, and achieved good results.

3.5. Mixed models

Because the single prediction model has its own advantages, disadvantages and applicable conditions, and the traffic flow system is complex and changeable, no one model is considered to be able to predict the real-time and efficient traffic flow. This dilemma urges researchers to consider the way of combining models and choose two or more models according to different situations to achieve the effect of integrating the advantages of all kinds of models. After the training of the mixed model is completed, the short-term time period data can be predicted based on the historical data in the historical-time period. The specific prediction steps of the combined model are as steps shown on Figure 2. J. N. Bates and C. W. J. Granger [24] first put forward combined forecasting in 1969, and put it into practice. He chose different models to combine, and finally got better forecasting effect. Because the combination model needs to optimize the weight distribution algorithm, it increases the model complexity, which is also a defect of the combination model, is to improve accuracy at the expense of some performance. There are many researches on combination forecasting models, including constrained least square method, minimum variance method, recursive method and Bayesian network structure-based combination forecasting method.

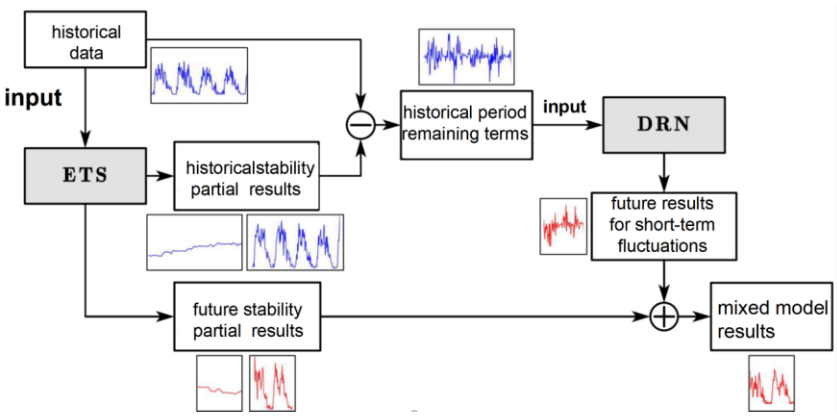

Figure 2. Mixed model forecast process.

4. Summary

In the past, old-fashioned trajectory prediction methods could only work in simple scenarios and short-term predictions, but the methods based on deep learning are able to perform accurate predictions in a longer time frame and are superior to traditional trajectory prediction models.

Through the analysis of different short-term traffic flow prediction models, it is found that the immediate single prediction model is difficult to get rid of its limitations. For example, most of the prediction models based on statistical analysis have the advantage of being simple and fast, but they are difficult to deal with unexpected traffic conditions in a timely manner. Moreover, the accuracy achieved is quite common, and it can only play a role in the interval with stable traffic conditions and low accuracy. Most nonlinear theoretical models are complex and computationally intensive. Simulation prediction models have a more adequate theoretical basis and incorporate some of the more complex influencing factors of the traffic system in the study, but they are computationally intensive and not suitable for large-scale real-time prediction of complex traffic systems because of their low real-time performance.

Moreover, there is not yet an optimal prediction model that can be applied to all traffic scenarios, so traffic flow prediction with combined prediction models is still a future development trend in this field, and with the update of technology and the improvement of algorithm performance, intelligent trajectory prediction will definitely usher in a new period. With the improvement of neural network models and the increase of computer computing power, intelligent trajectory prediction will become the focus of future research in this area. Meanwhile, the incorporation of vehicle data and pedestrian movement trajectory data into traffic flow prediction, forming a multi-data fusion prediction method, will also be a new development direction in the future.

References

[1] Alexopoulos C, Keramidis P, Pereira G V, Charalabidis Y 2022 Towards Smart Cities 4.0: Digital Participation in Smart Cities Solutions and the Use of Disruptive Technologies *M. Themistocleous, M. Papadaki. Information Systems (emcis 2021). Cham: Springer International Publishing Ag*, 437: 258–273.

[2] Fu R, Zhang Z and Li L 2016 Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction. *31st Youth Academic Annual Conference of Chinese Association of Automation New York: Ieee*, 324–328.

[3] Nagy A M, Simon V 2018 Survey on traffic prediction in smart cities. *Pervasive and Mobile Computing*, 50: 148–163.

[4] Kong F, Li J, Jiang B, Tianyuan Z and Houbing S 2019 Big data-driven machine learning-enabled Traffic flow prediction. *Transactions on Emerging Telecommunications Technologies*, Hoboken: Wiley, 30(9): E3482.

[5] Jing Y, Hu H, Siye G, Xuan W, and Fangqiu C 2021 Short-term Prediction of Urban Rail Transit Passenger Flow in External Passenger Transport Hub Based on LSTM-LGB-DRS. *IEEE Transactions on Intelligent Transportation Systems, Piscataway: Ieee-Inst Electrical Electronics Engineers Inc*, 22(7): 4611-4621.

[6] Meng K, Shi C and Meng Y 2018 Vehicle Action Prediction Using Artificial Intelligence. M. A. Wani, M. Kantardzic, M. Sayedmouchaweh, wait. *17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. New York: IEEE, 2018:1231-1236.

[7] Jiahe Y, Honghui L, Yanhui B and Yingli L 2021 Spatial-temporal Traffic flow Restoration Data and Prediction Method Based on the Tensor Decomposition. *Applied Sciences, Multidisciplinary Digital Publishing Institute*, 2021,11(19): 9220.

[8] Ma, C, Dai, G, and Zhou J 2021 Short-term Traffic Flow Prediction for Urban Road Sections Based on Time Series Analysis and LSTM_BILSTM Method. *IEEE Transactions on Intelligent Transportation Systems*, 23(6), 5615-5624.

[9] El Faouzi N-E, Leung H, Kurian A. 2011 Data fusion in intelligent transportation systems: Progress and challenges - A survey. *Information Fusion*, Amsterdam: Elsevier, 12(1): 4-10.

[10] Ding, F, Zhang, Z, Zhou, Y, Chen, X, and Ran B. 2019 Large-scale full-coverage traffic speed

DOI: 10.54254/2755-2721/10/20230183

estimation under extreme traffic conditions using a big data and deep learning approach: Case study in China. *Journal of Transportation Engineering, Part A: Systems*, 145(5), 05019001.

[11] Okutani I, Stephanedes Y. 1984 Dynamic Prediction of Traffic Volume Through Kalman Filtering Theory. *Transportation Research Part B-Methodological*, Oxford: Pergamon-Elsevier Science Ltd, 18(1): 1-11.

[12] Guoguang H, Shoufeng M and Yu L 2002 Time Series Prediction method based on wavelet decomposition and reconstruction. *Acta Automatica*, 2002(6) : 1012-1014.

[13] Zhai, H, Cui, L, Nie Y, Xu X, and Zhang W. 2018 A comprehensive comparative analysis of the basic theory of the short term bus passenger flow prediction. *Symmetry*, 10(9), 369.

[14] Forbes G, Hall F. 1990 The Applicability of Catastrophe-Theory in Modeling Freeway Traffic Operations. *Transportation Research Part a-Policy and Practice*, Oxford: Pergamon-Elsevier Science Ltd, 24(5): 335-344.

[15] Bengio Y, Simard P, Frasconi P. 1994 Learning Long-Term Dependencies with Gradient Descent Is Difficult. *IEEE Transactions on Neural Networks*, Piscataway: Ieee-Inst Electrical Electronics Engineers Inc, 5(2): 157-166.

[16] Yao L, Mao C, Luo Y. 2019 Graph Convolutional Networks for Text Classification. T*hirty-Third Aaai Conference on Artificial Intelligence / Thirty-First Innovative Applications of Artificial Intelligence Conference / Ninth Aaai Symposium on Educational Advances in Artificial Intelligence. Palo Alto: Assoc Advancement Artificial Intelligence*, 2019: 7370-7377.

[17] Zhu M, Yang X. 2019 Chinese Texts Classification System. *IEEE 2nd International Conference on Information and Computer Technologies (icict)*. New York: IEEE, 2019: 149-152.

[18] Yu F, Wei d, Zhang S, et. al. 2019 3D CNN-based Accurate Prediction for Large-scale Traffic flow. *International Conference on Intelligent Transportation Engineering (icite 2019)*. New York: IEEE, 2019:99-103.

[19] Huang y, Du j, Yang Z, et. al. 2022 A Survey on Trajectory-Prediction Methods for Autonomous Driving. *IEEE Transactions on Intelligent Vehicles*, Piscataway: Ieee-Inst Electrical Electronics Engineers Inc, 7(3): 652-674.

[20] Djuric N, RADOSAVLJEVIC v, Cui H, et. al. 2020 Uncertainty-aware Short-term Motion Prediction of Traffic Actors for Autonomous Driving. *IEEE Winter Conference on Applications of Computer Vision (WACV)*. Los Alamitos: IEEE Computer Soc, 2020:2084-2093.

[21] Rumelhart D, Hinton G, Williams R. 1986 Learning Representations by Back-Propagating Errors. *Nature*, London: Macmillan Magazines Ltd, 323(6088): 533-536.

[22] Zhang S, Yao y, Hu J, et. al. 2019 Deep Autoencoder Neural Networks for Short-Term Traffic Congestion Prediction of Transportation Networks. *Sensors*, Basel: MDPI, 19(10): 2229.

[23] Xu D, Wang y, Jia l, et. al. 2017 Real-time road traffic state prediction based on Arima and Kalman filter. *Frontiers of Information Technology and Electronic Engineering*, Hangzhou: Zhejiang Univ, 18(2): 287-302.

[24] Bates J, Granger C 1969 Combination of forecasts. *Operational Research Quarterly, Oxford: Pergamon-Elsevier Science Ltd, 1969, 20(4): 451-455.*