Research on Short-term Traffic Flow Forecast and Auxiliary Guidance Based on Artificial Intelligence Theory

Rongxia Wang*
School of Guangzhou Nanyang Polytechnic College, Guangzhou 510000, China;

*Corresponding author e-mail: ziming323@126.com

Abstract. With the development of social economy, intelligent transportation system (ITS) has been booming. It is mainly to realize the omni-directional, real-time, accurate and efficient guidance and control of transportation in a large range. Traffic flow prediction, especially short-term traffic flow prediction, is the basis of urban traffic control and guidance, so traffic flow prediction system is one of the important subsystems in its. Short-term traffic flow prediction technology belongs to the important research field of intelligent traffic control and vehicle guidance. It can help cities to conduct intelligent traffic guidance and enable users to choose the optimal path. Prediction of road traffic flow state and auxiliary guidance are very important basic theories. Based on the theory of artificial intelligence, this paper analyzes the short-term traffic flow prediction and auxiliary induction methods, points out the limitations of the existing prediction methods, and proposes a short-term traffic flow intelligent prediction method combined with artificial intelligence technology.

1. Introduction
For modern cities, transportation directly affects everyone's life. Whether the traffic problem can be solved will directly affect the improvement of people's living standards and the development of the national economy [1]. Short-term traffic flow prediction is the core content of modern ITS and the basis for achieving advanced traffic control and traffic guidance. Urban road traffic is a dynamic and complex system. With the reduction of observation time range, traffic characteristics transition from certainty to randomness, and the prediction difficulty of traffic flow also increases [2]. Traffic flow prediction, especially short-term traffic flow prediction, is the basis of urban traffic control and guidance, so traffic flow prediction system is one of the important subsystems in ITS [3]. The causes of traffic congestion are various, each city has its own special situation, and the determinants of congestion are also different. Short-term traffic flow prediction has broad application prospects in intelligent traffic control and induction, and can alleviate or solve urban traffic congestion problems [4]. The short-term prediction of traffic flow is an important part of the realization of intelligent traffic control and guidance systems. Because traffic flow is affected by many factors, it is non-linear and uncertain [5]. On the surface, the construction of new roads will increase road capacity and reduce travel time in a short period of time, but at the same time the traffic flow on existing roads will gradually shift, and the overall congestion of the city will return to its original level after a period of time.

The causes of traffic congestion are various, each city has its own special situation, and the determinants of congestion are also different. The realization of urban traffic flow control and guidance
system will effectively reduce traffic congestion and urban environmental pollution, improve road capacity and improve traffic safety [6]. At present, urban traffic control and induction generally adopt preset schemes. A few cities use adaptive control modes based on traffic flow detection. Intelligent traffic control and induction based on short-term traffic flow prediction are still few applications [7]. Among the many important components of ITS, the dynamic analysis and prediction of road traffic flow is a very important basic theory and is also one of the difficult problems currently being solved in the world traffic field. In the existing road system, the computer of the traffic control center is used to process and display various data, which come from traffic signal control machines, detectors on both sides of the road, etc. Short-term traffic flow prediction is the basis of traffic state prediction, therefore, the key to realize the road traffic flow guidance system is the prediction of road traffic state. In other words, corresponding technologies are adopted to effectively use real-time traffic data information to scroll and predict traffic conditions in the future [9]. Based on the theory of artificial intelligence, this paper points out the limitations of the existing prediction methods on the basis of analyzing the short-term traffic flow prediction and auxiliary guidance methods, and proposes a short-term traffic flow intelligent prediction method combined with artificial intelligence technology.

2. Traffic Flow Data Analysis

2.1. Quality Problems of Data Sources
With the development of transportation science and technology, more and more automatic traffic flow data acquisition devices have been installed on expressways or trunk roads in some cities in China. Obtaining massive traffic flow data has become a reality, providing data foundation guarantee for road traffic flow prediction. The traffic flow change process is a real-time, non-linear, multi-dimensional and stable random process. With the shortening of statistical period, the randomness and uncertainty of traffic flow change become stronger and stronger. In the real-time detection process of urban road traffic flow, noise interference and other reasons cause data loss or erroneous data. Correct identification and effective processing of erroneous data and missing data is a crucial step in the process of establishing a prediction model. Because during the data acquisition process, unexpected events such as hardware failure, noise interference, and communication failure of traffic flow detectors may occur, there is no guarantee that all data is perfect. Subject to the rules of public transportation, the main goal of traffic flow is to pursue safety, speed, and smoothness rules, so there is mutual cooperation and synergy. Before analyzing the data of traffic flow prediction, you should first control the quality of the data and use pre-processing techniques to pre-process the collected data. By identifying and repairing erroneous or missing data, and providing a reasonable data source for the next analysis of traffic flow data, the effect of the final traffic flow prediction can be guaranteed.

2.2. Correlation between Traffic Flow Data of Multiple Sections
Because the data comes from the same road network, there is a correlation between the traffic flow data of multiple sections in the road network. This correlation is affected by the distance between each other and the surrounding environment of the road. Faced with increasingly congested traffic, limited resources and financial resources, and environmental pressures, the construction of more infrastructure is restricted. When the internal and external factors affecting the traffic flow process remain relatively stable, the results obtained using these models are valid. Due to the short-term requirements of short-term traffic flow prediction, when the road network is relatively large, it is impossible to simultaneously predict all the cross-section traffic flow data in the road network [10]. The key of short-term traffic flow prediction is to be able to select the most suitable prediction model according to the current situation of traffic flow in real time, and to be able to judge according to the prediction effect, and constantly improve the prediction effect. In the city, the delay caused by vehicles mainly comes from the delay of intersections. Because the traffic flow is periodically interfered by the intersection signal control, the vehicles in some directions stop driving periodically. An accurate and efficient traffic flow prediction method should be able to correctly grasp the changes of the predicted process and its environment, and
adjust the structure of the model in time to make the prediction adaptive. Before the multi section short-term traffic flow prediction of road network, the research scope should be selected according to the correlation of section traffic flow data.

3. Short-term Prediction and Auxiliary Induction Method Based on Support Vector Machine

In machine learning, support vector machine (SVM) is a supervised learning model and related learning algorithm to analyze data in classification and regression analysis. SVM training algorithm creates a model that assigns a new instance to one of two categories, making it a non-probabilistic binary linear classifier. In SVM model, instances are represented as points in space, so the mapping makes the instances of individual categories separated by as wide and obvious intervals as possible. It is difficult to explain the traffic flow state with exact numbers, which has strong fuzziness and concealment. Due to different purposes, models and routes, the running state of various vehicles on the road varies with the road conditions, traffic environment and driver characteristics. The change process of traffic flow is a real-time, non-linear, high-law, non-stationary random process, and has uncertainties of change. The so-called short-term traffic flow prediction is to predict the changes in traffic conditions of a road or a traffic corridor in the next few minutes. Cross SVM can seek a compromise between model complexity and learning ability based on limited sample data in order to obtain the best promotion ability. SVM transforms the optimization problem into a convex programming problem, and the resulting local optimal solution is the global optimal solution. Once the flow is abnormal, it should be able to dynamically feed back into the calculation model for adjustment according to the actual situation. The Figure 1 shows the modeling process of short-term traffic flow based on SVM.

![Figure 1. SVM-based short-term traffic flow modeling process](image)

Since traffic flow prediction is related to many factors, assume \( x_i \in \mathbb{R}^n \) is the factor affecting traffic flow prediction and \( y_i \) is the traffic flow prediction value. The traffic flow prediction model based on support vector machine is to find the relationship between \( x_i \) and \( y_i \):

\[
\begin{align*}
  f : \mathbb{R}^n &\to \mathbb{R} \\
  y_i &= f(x_i)
\end{align*}
\]

Where: \( \mathbb{R}^n \) is the factor affecting traffic flow prediction. According to the support vector machine theory, the following expression is sought for the establishment of the traffic flow prediction model:

\[
  f(x) = \sum_{i=1}^{k} (a_i - a^*_i) K(x, x_i) + b
\]
Where: \( x \) is the factor affecting traffic flow, \( x_i \) is the \( i \) sample in \( k \) samples, \( K(x, x_i) \) is the kernel function, and the kernel function is radial basis function, as shown in the following formula:

\[
K(x, y) = \exp\left(\frac{-\|x - y\|^2}{2\sigma^2}\right)
\]

(4)

In order to quantitatively analyze the prediction effect of the model, two evaluation indexes are introduced, namely the mean square error (MSE) and the mean absolute percentage error (MAPE). The prediction error of the method in this paper is compared with the Kalman filtering method. MSE is defined as:

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} \left( Y_i - Y_i^* \right)^2
\]

(5)

MAPE is defined as:

\[
MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{Y_i - Y_i^*}{Y_i} \right| \times 100\%
\]

(6)

The prediction errors of the two methods for cross-section traffic flow are shown in the Table 1.

|Prediction method      | Section   | MAPE/\% | MSE      |
|-----------------------|-----------|---------|----------|
|Kalman filter method   |           |         |          |
|Section 1              | 6.17      | 417.55  |
|Section 2              | 9.48      | 213.64  |
|Section 3              | 5.05      | 315.28  |
|Methods in this paper  |           |         |          |
|Section 1              | 5.61      | 363.47  |
|Section 2              | 6.91      | 167.59  |
|Section 3              | 3.82      | 285.89  |

In order to facilitate comparison, the prediction models are classified according to the location of the prediction points, the size of saturation, the length of the prediction period, the prediction period, etc. To obtain the accuracy of each model for different prediction periods in different positions and different flow sizes and different time periods. In SVM, the value of penalty factor represents the importance that the classifier attaches to the incorrectly divided points. The higher the value of error penalty factor, the more attention the classifier attaches to the noise point, and vice versa. For the model with few parameters, the process of determining parameters is relatively simple and the required time is relatively short. When deciding to use the model, the current and recent historical data can be used to determine the pending parameters in a timely manner. In different cases, the results of traffic flow prediction using the model are stored in the corresponding database of the system. When the predicted time period is completed, the system can obtain the corresponding actual flow value by setting detectors at different positions of the road section. Some models, such as neural network model, need a lot of parameters to predict, and the method to determine the parameters is complex, which is difficult to complete online and real-time. For these models, the state space of model parameters can be determined according to different forecast locations, flows, time periods and forecast periods. If the test results show that the prediction errors of various models are greater than the given threshold, it means that the traffic process and environment have changed dramatically, and the current model or parameters can not correctly describe its characteristics.
4. Conclusions
More and more serious traffic problems make people's demand for intelligent traffic management more and more urgent. Intelligent transportation system is an effective way to solve many urban traffic problems. The urban traffic system itself is a very complex and changeable system. With the rapid development of social economy, the acceleration of urbanization, the increase of vehicle ownership and the rapid increase of traffic demand bring great pressure to urban traffic. This paper studies the problem of short-term traffic flow prediction and auxiliary guidance of road network based on real-time data. Through the analysis of traffic flow data, the theory and method of short-term traffic flow prediction of multiple sections in road network considering time and space changes are established. The classification accuracy of SVM is closely related to the selection of its parameters. Using the advantages of RBF kernel function and selecting corresponding parameters can greatly improve the learning ability of SVM. For the model with few parameters, the process of determining parameters is relatively simple and the required time is relatively short. When deciding to use the model, the current and recent historical data can be used to determine the pending parameters in a timely manner. In the short-term traffic flow prediction system, it is necessary to adopt an efficient and adaptable intelligent prediction method according to the characteristics of randomness, uncertainty and obvious nonlinear characteristics of the traffic flow process.

Acknowledgements
This is the phased research result of the “Research on short-term traffic flow prediction based on bat algorithm support vector machine” (Project No: NY-2019KYYB-22) and "XAPI-based wisdom education cloud classroom higher vocational teaching activity model research"(Project No: NY-2019CQJGYB-05) from Guangzhou Nanyang Polytechnic College.

References
[1] Habtemichael F G, Cetin M. Short-term traffic flow rate forecasting based on identifying similar traffic patterns [J]. Transportation Research Part C Emerging Technologies, 2015, 66:61-78.
[2] Hou Y, Edara P, Sun C. Traffic Flow Forecasting for Urban Work Zones [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(4):1761-1770.
[3] Tan H, Wu Y, Shen B, et al. Short-Term Traffic Prediction Based on Dynamic Tensor Completion [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(8):1-11.
[4] Guo J, Huang W, Williams B M. Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification [J]. Transportation Research Part C, 2014, 43:50-64.
[5] Agafonov A A, Myasnikov V V. Method for the reliable shortest path search in time-dependent stochastic networks and its application to GIS-based traffic control [J]. Computer Optics, 2016, 40(2):275-283.
[6] Lee D N. Solving large telecommunications network loading problems [J]. At & T Technical Journal, 2014, 68(3):48-56.
[7] Romauch M. Product mix optimization for a semiconductor fab: Modeling approaches and decomposition techniques [J]. Computers & Operations Research, 2015, 53:338-352.
[8] Chen D, Hu M, Zhang H, et al. Short-term traffic flow prediction of airspace sectors based on bayesian estimation theory [J]. Xinan Jiaotong Daxue Xuebao/Journal of Southwest Jiaotong University, 2016, 51(4):807-814.
[9] Yuan P C, Lin X X. How long will the traffic flow time series keep efficacious to forecast the future [J]. Physica A: Statistical Mechanics and its Applications, 2016, 467:419-431.
[10] Dan C, Minghua H U, Honghai Z, et al. Forecast Method for Medium-Long Term Air Traffic Flow Considering Periodic Fluctuation Factors [J]. Journal of Southwest Jiaotong University, 2015, 50(3):562-567.