Application Research of Short-term Traffic Flow Forecast Based on Bat Algorithm Support Vector Machine

Rongxia Wang*
School of Guangzhou Nanyang Polytechnic College, Guangzhou 510000, China.
*Corresponding author e-mail: ziming323@126.com

Abstract. Characteristics transition from certainty to randomness, and the prediction difficulty of traffic flow also increases. Short-term traffic flow prediction technology can help cities to induce intelligent traffic by a Urban road traffic is a dynamic and complex system. With the reduction of observation time range, traffic analyzing and predicting traffic flow. Through the analysis of traffic flow data and the identification and processing of erroneous and missing data, the influence of noise on the prediction process is reduced. Intelligent Transportation System (ITS) is getting more and more attention. At the same time, people put forward higher requirements for vehicle type recognition, license plate recognition, traffic flow prediction and other technologies. Support vector machine (SVM) can find a compromise between model complexity and learning ability according to limited sample data, in order to obtain the best generalization ability. Based on bat algorithm (BA) support vector machine, this paper studies the basic algorithm of pattern recognition and regression analysis and its application in short-term traffic prediction of intelligent transportation system.

1. Introduction
With the development of society, traffic accidents, traffic jams, environmental pollution and energy consumption are becoming more and more serious. Especially in some big cities, traffic congestion has become a bottleneck restricting the development of urban economy [1]. 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]. Intelligent transportation system is getting more and more attention. At the same time, people put forward higher requirements for vehicle type recognition, license plate recognition, traffic flow prediction and other technologies. Accurate and reliable traffic prediction information is the foundation and key of dynamic route guidance system, and is also an important component of intelligent traffic system [3]. Short-term traffic flow prediction has a wide application prospect in intelligent traffic control and guidance, which can alleviate or solve the problem of urban traffic congestion [4]. Short-term traffic flow prediction is an important link to realize intelligent traffic control and guidance system. Traffic flow is affected by many factors, which makes it nonlinear and uncertain [5]. Support vector machine (SVM) can seek a compromise between model complexity and learning ability based on limited sample data, in order to obtain the best promotion ability [6]. SVM transforms the optimization problem into a convex programming problem, and the resulting local optimal solution is the global optimal solution.

At present, urban traffic control and induction generally adopt preset schemes. A few cities adopt adaptive control mode based on traffic flow detection, and there are still few applications of intelligent traffic control and induction based on short-term traffic flow prediction. The incentives for traffic congestion are multi-faceted. Each city has its own special circumstances, and the decisive factors for
congestion are also different [7]. It can't just be seen as a simple technical issue, but a complex social issue. Bat Algorithm (BA) is a new group intelligent optimization algorithm. Compared with other algorithms, the BA algorithm has a significant improvement in validity and accuracy [8]. The discriminant function obtained by the bat algorithm SVM training is similar in form to the feedforward neural network, and its output is a linear combination of several intermediate layer nodes, and each intermediate layer node corresponds to the inner product operation of the input sample and a certain support vector [9]. Traffic flow short-term forecasting is the basis of traffic state prediction. Therefore, the key to realizing road traffic flow guidance system is the prediction of road traffic status. That is to say, the corresponding technology is used to effectively use the real-time traffic data information to predict the traffic state in the future period [10]. Based on the bat algorithm SVM, this paper makes a deep research on the basic algorithms of pattern recognition and regression analysis and its application in short-term traffic prediction in intelligent transportation systems.

2. Basic Characteristics of Traffic Flow and SVM Principle

2.1. Analysis of Traffic Flow Characteristics and Influencing Factors
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. People and traffic flow moving on the road constitute a complex road traffic, and both vehicles and pedestrians will show certain characteristics when moving on the road. People and traffic flow jointly constitute the traffic flow. Faced with more and more crowded traffic, limited resources and financial resources and environmental pressure, the construction of more infrastructure is restricted. Urban traffic roads crisscross and connect with each other, interweaving main roads and auxiliary roads, and each road has a certain traffic flow. Under the restriction of public traffic rules, the main behavior of traffic flow pursues safety, speed and smoothness criteria, so there is mutual cooperation and coordination. The change of traffic flow state has certain regularity under certain conditions. In fact, this characteristic tendency can also be expressed by the characteristic parameters of the traffic flow state. According to different unit time, we further divide traffic flow: time traffic flow, hourly traffic flow, daily traffic flow, weekly traffic flow, etc.

The main factors affecting the change of traffic flow are the influence of traffic control facilities and the capacity of roads. Traffic control facilities mainly refer to traffic signs, traffic lines and the timing of cross signals. According to the different ways of using the detection equipment and the location of use, the traffic flow information collection technology can be divided into two types: fixed type and mobile type. As the scope of research, development, and testing continues to expand, Intelligent Transportation Systems (ITS) have emerged. ITS is a broad term that includes a wide range of technologies and provides solutions to many traffic problems [11]. In the city, the delay caused by the vehicle mainly comes from the delay of the intersection. Because the traffic flow is periodically interfered by the intersection signal control, some vehicles in the flow direction stop periodically. Traffic congestion occurs when the amount of traffic on the road approaches the capacity of the road. At this time, all vehicles are lined up at the same speed. Once there is interference, it is easy to cause traffic congestion. In the real-time detection process of urban road traffic flow, noise interference and other reasons cause data loss or erroneous data. Intelligent transportation systems enable transportation infrastructure to be more efficient and enable society to use transportation facilities and energy efficiently, thereby achieving enormous economic and social benefits.

2.2. Basic Flow of Short-term Traffic Flow Forecast
With the change of modern traffic management mode from passive management to active guidance, prediction information becomes more and more important. The change process of traffic flow is a real-time, non-linear, high-regularity, non-stationary random process with uncertainty of change. The so-called short-term traffic flow prediction is to predict the changes of traffic flow in a certain road or corridor in the next few minutes. In SVM, the value of penalty factor represents the importance that the
classifier attaches to the points that are mistakenly divided. The higher the value of error penalty factor, the less attention the classifier attaches to the noise point, and vice versa [12]. Short-term traffic flow prediction can be used to guide the formulation and implementation of a traffic management plan and regulate traffic flow to reduce possible traffic congestion and potential dangers. Correct identification and effective processing of erroneous data and missing data is a crucial step in the process of establishing a prediction model. Once the traffic flow is abnormal, it should be able to dynamically feed back to the calculation model for adjustment according to the actual situation. According to the time series variation rule of traffic flow, traffic flow on the road section has an inevitable connection with traffic flow in the previous periods, thus traffic flow data sequence in the previous periods of the road section can be used to predict traffic flow in the future periods.

The problem to be solved in traffic flow prediction is how to systematically analyze the data according to traffic flow parameter data from various traffic flow information collection devices from the traffic flow changes with randomness and uncertainty, and establish corresponding prediction models and methods. The Fig. 1 shows the short-term traffic flow forecasting process.

![Fig. 1 Short-term traffic flow forecasting process](image)

The traffic flow on the road section is not only related to the traffic flow in the previous several periods of the road section, but also affected by the traffic conditions of the upstream and downstream sections. The collected samples are usually mixed with abnormal values, which often deviate from the actual values, which makes data analysis and modeling difficult. Statistical prediction methods can be divided into statistical methods and time series analysis methods. They try to regard traffic flow parameters as a time variable, so as to find the implicit statistical laws or relations in the time series. The detected real-time traffic data is a superposition of real traffic data and various interference noises. Direct observation will find that the data changes are more intense and the laws cannot be mastered [13]. Traditional statistics studies the gradual theory that the number of samples approaches infinity. When the number of samples is limited, it is difficult to obtain ideal results, so it is difficult to adapt to the current complex and changeable traffic conditions. Before data analysis and modeling, it is necessary to check and deal with abnormal values. The traffic flow in the first few periods of a certain road section and the traffic flow in the first few periods of the upstream and downstream road sections can be used to predict the traffic flow in the current period of the road section.

3. Short-term Traffic Flow Forecast Method Based on Bat Algorithm SVM

The integration of emerging information, control, communication and computer technologies with traffic engineering has formed various unique technologies in intelligent transportation systems. Such as intelligent traffic control technology for urban roads and expressways, traffic information collection and fusion technology, route navigation and traffic information service technology, etc.
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 $y_i$:

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$y_i = f(x_i)$$

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 $i$-th 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:

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

Bat algorithm has the advantages of fast convergence speed and simple and easy implementation of the algorithm. Therefore, when the bat algorithm is used to solve the problem quickly, it is often advantageous. This paper proposes to combine the bat algorithm with the support vector machine to optimize the model of the support vector machine through the process of search optimization. Support vector regression is a regression algorithm based on SVM and has been successfully applied in practical engineering. The basic idea of the algorithm is to increase the dimension of the data, and the construction of decision function is realized in high-dimensional space to achieve the purpose of linear regression. In this process, how to select regression parameters will determine whether its theoretical advantages can be realized. In the process of human evolution and development, learning ability plays a vital role, that is, through the process of acquiring, analyzing and summarizing the known information, we find some laws among them, and then predict the information that cannot be directly obtained through observation according to these laws. Since traffic flow prediction is essentially a prediction of basic parameters of traffic flow, its characteristics can be described from the basic parameters of road traffic flow.

As shown in Fig. 2 and Fig. 3, the absolute relative error of traffic flow prediction model based on BP neural network and support vector machine and the comparison between actual flow and predicted flow are respectively shown. Table 1 shows the evaluation index values of the two models.

![Fig. 2 Absolute relative error](image-url)
Fig. 3 Actual traffic and predicted traffic

Table 1 Evaluation index values of the two models

| Model                        | Evaluation index |
|------------------------------|------------------|
|                              | marerr | mxarer | mrerr | Learnt | genability | contime | opt       |
| Based on BP neural network   | 2.83%   | 60.58% | 78.59% | 45.23s | Weak       | Slow    | Locally optimal |
| Based on SVM                 | 7.59%   | 19.62% | 11.34% | 0.87s  | Strong     | Fast    | Global optimization |

The working process in the contention-based error correction coding access SVM is still controlled by the master node. Unlike the polling method, the master node does not know which slave node to query in advance. Only when the master node receives the response signal from the slave node does it begin to query the slave node, thus the access of the slave node presents a certain disorder. The application of SVM in short-term traffic flow prediction has good effect. If we can deepen our understanding of the inherent characteristics of traffic flow and effectively combine other technical methods, such as correlation analysis and outlier detection, the quality of prediction will be further improved.

4. Conclusions
The urban transportation system itself is a very complicated and changeable system. With the rapid development of social economy, the acceleration of urbanization, the increase of vehicle ownership and the sharp increase of transportation demand have brought great pressure to the urban transportation. The application of SVR in short-term traffic flow prediction has good effect. If we can deepen our understanding of the inherent characteristics of traffic flow and effectively combine other technical methods, such as correlation analysis and outlier detection, the quality of prediction will be further improved. The classification accuracy of SVM is inseparable from the selection of its parameters. Using the advantages of RBF kernel function and selecting the corresponding parameters can greatly improve the learning ability of SVM. Through the empirical analysis and comparison of the short-term traffic flow measured at a non-detector intersection, it is proved that the proposed prediction model is qualified, and the fitting accuracy is high, the prediction effect is good, and it can be used for short-term traffic flow prediction. SVR technology is not only feasible but also effective when dealing with specific problems in intelligent transportation systems. With the deepening of theoretical research and the continuous expansion of applied research, SVR will have a broader development prospect.

Acknowledgements
This is the phased research result of the “Research on short-term traffic flow prediction based on bat algorithm support vector machine” from Guangzhou Nanyang Polytechnic College. (Project No: NY-2019KYYB-22).
References

[1] Huang S, Chang J, Huang Q, et al. Monthly streamflow prediction using modified EMD-based support vector machine[J]. Journal of Hydrology, 2014, 511:764-775.

[2] Agafonov A A, Myasnikov V V. An algorithm for traffic flow parameters estimation and prediction using composition of machine learning methods and time series models[J]. Computer Optics, 2014, 38(3):539-549.

[3] Tang M, Li Z, Tian G. A Data-Driven-Based Wavelet Support Vector Approach for Passenger Flow Forecasting of the Metropolitan Hub[J]. IEEE Access, 2019, 7:7176-7183.

[4] Abdi J, Moshiri B. Application of temporal difference learning rules in short-term traffic flow prediction[J]. Expert Systems, 2015, 32(1):49-64.

[5] Bagui S, Fang X, Kalaimannan E, et al. Comparison of machine-learning algorithms for classification of VPN network traffic flow using time-related features[J]. Journal of Cyber Security Technology, 2017, 1(2):108-126.

[6] Cheng M Y, Wu Y W, Syu R F. Seismic Assessment of Bridge Diagnostic in Taiwan Using the Evolutionary Support Vector Machine Inference Model (ESIM)[J]. Applied Artificial Intelligence, 2014, 28(5):449-469.

[7] Hammoud R I, Sahin C S, Blasch E P, et al. Automatic Association of Chats and Video Tracks for Activity Learning and Recognition in Aerial Video Surveillance[J]. Sensors, 2014, 14(10):19843-19860.

[8] Haworth J, Shawe-Taylor J, Cheng T, et al. Local online kernel ridge regression for forecasting of urban travel times[J]. Transportation Research Part C: Emerging Technologies, 2014, 46:151-178.

[9] Huang M L. Intersection traffic flow forecasting based on ν-GSVR with a new hybrid evolutionary algorithm[J]. Neurocomputing, 2015, 147(1):343-349.

[10] Kim, E. Y. MRF model based real-time traffic flow prediction with support vector regression[J]. Electronics Letters, 2017, 53(4):243-245.

[11] Wongsathan R, Seedadan I. A Hybrid ARIMA and Neural Networks Model for PM-10 Pollution Estimation: The Case of Chiang Mai City Moat Area[J]. Procedia Computer Science, 2016, 86:273-276.

[12] Jun K, Zong-Tao D, Lei T, et al. A Short Term Traffic Flow Prediction Method Based on Gaussian Processes Regression[J]. Journal of Transportation Systems Engineering and Information Technology, 2015, 15(4):51-56.

[13] Guo J, Liu Z, Huang W, et al. Short-term traffic flow prediction using fuzzy information granulation approach under different time intervals[J]. IET Intelligent Transport Systems, 2018, 12(2):143-150.