Novel Real-Time System for Traffic Flow Classification and Prediction

YE Dezhong[1], LV Haibing[1], GAO Yun[2], BAO Qiuxia[2], and CHEN Mingzi[2]

(1. ZTE Corporation, Shenzhen, Guangdong 518057, China;
2. Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China)

Abstract: Traffic flow prediction has been applied into many wireless communication applications (e.g., smart city, Internet of Things). With the development of wireless communication technologies and artificial intelligence, how to design a system for real-time traffic flow prediction and receive high accuracy of prediction are urgent problems for both researchers and equipment suppliers. This paper presents a novel real-time system for traffic flow prediction. Different from the single algorithm for traffic flow prediction, our novel system firstly utilizes dynamic time wrapping to judge whether traffic flow data has regularity, realizing traffic flow data classification. After traffic flow data classification, we respectively make use of XGBoost and wavelet transform-echo state network to predict traffic flow data according to their regularity. Moreover, in order to realize real-time classification and prediction, we apply Spark/Hadoop computing platform to process large amounts of traffic data. Numerical results show that the proposed novel system has better performance and higher accuracy than other schemes.

Keywords: traffic flow prediction; dynamic time warping; XGBoost; echo state network; Spark/Hadoop computing platform

DOI: 10.12142/ZTECOM.201902003

http://kns.cnki.net/kcms/detail/34.1294.TN.20190611.1545.002.html, published online June 11, 2019

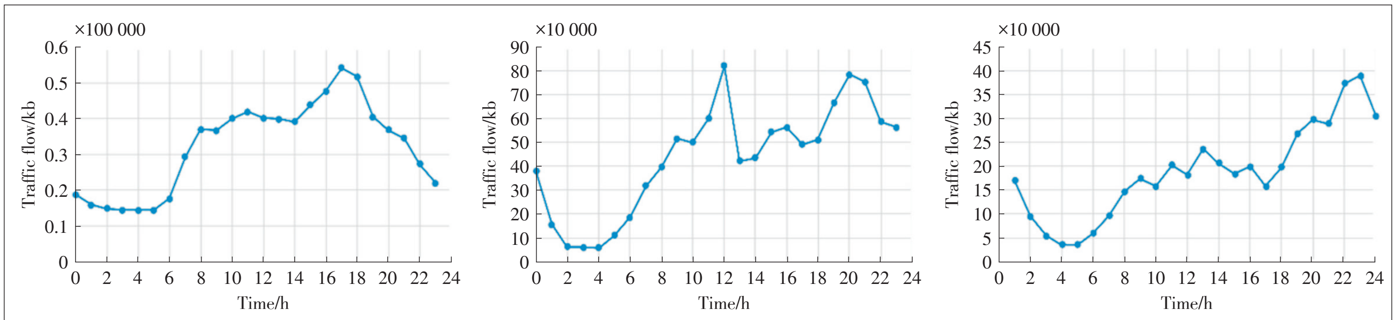
Manuscript received: 2018-09-28

1 Introduction

In recent years, as the advances of multimedia communications and technologies [1]–[3], traffic flow data has been widely studied and used in various situations. For one thing, mobile users pay more attention on their experience of surfing online; for another, some service providers like Zhongxing Telecommunication Equipment Corporation (ZTE) want to provide users with better service by traffic flow prediction and allocation. Therefore, traffic flow research has become one of the hottest research topics in our daily lives. As we all know, quality of experience (QoE) has been considered as one of the most evaluation criteria for user experience description, which is mixed with subjective and objective indicators [4]. To improve users' QoE, service providers will predict traffic flow data so that they could substitute their devices to provide users with higher satisfaction of service. Obviously, how to accurately predict traffic flow data and quickly estimate traffic flow trends are very important for both users and service providers.

Generally speaking, traffic flow prediction greatly depends on historical records from different base stations. With the increasing number of users, the amount of traffic flow data becomes larger than before, which may bring different types of traffic trends. For example, **Fig. 1** shows three different traffic trends among the dataset. Given these flow diagrams, we have the following observations. Figs. 1a and 1b show the strong regularity and characteristics in the usage of traffic flow data for different regions like commercial district and residential area. For example, Fig. 1a shows that the traffic flow data starts to

This work is partly supported by the National Natural Science Foundation of China (Grants No. 61571240, 61671474), the Jiangsu Science Fund for Excellent Young Scholars (No. BK20170089), the ZTE program "The Prediction of Wireline Network Malfunction and Traffic Based on Big Data," (No. 2016ZTE04-07), Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX18_0916), and the Priority Academic Program Development of Jiangsu Higher Education Institutions.



a) Commercial district b) Residential area c) Unknown area

▲**Figure 1. Different traffic trends among the dataset.**

rise at 8 a.m. and decrease at about 6 p.m. Fig. 1b is the opposite of Fig. 1a. Compared with the two flow diagrams mentioned before, Fig. 1c seems to have weak regularity. It may contain the combinations of many non-linear factors such as base station's positions and different time periods [5], [6]. Besides, there are some other types of traffic flow regularity in the dataset. Therefore, how to predict these different types of traffic flow data has become a urgent problem for researchers.

Moreover, based on the observations mentioned before, how to define the regularity for traffic flow data is also very important. As we all know, most of us only use subjective feelings to classify the regularity for traffic flow data. However, faced with increasing number of traffic flow data, it would be impossible for us to judge regularity manually. Essentially, we must find the objective method to classify these traffic flow data.

Faced with issues mentioned before, we design a novel real-time system for traffic flow classification and prediction. It is an integrated system concerning both traffic flow classification and prediction. This system is divided into two main models: a traffic flow classification model and a traffic flow prediction model. In particular, for the classification model, we propose the dynamic time warping (DTW) algorithm for traffic classification, and it could classify traffic data into two categories including regularity and non-regularity. Subsequently, for the prediction model, we adopt XGBoost algorithm to predict traffic flow data with regularity. For traffic flow data with non-regularity, we first use wavelet transform (WT) to decompose them into different coefficients. Then we make full use of echo state network (ESN) to predict each coefficient. Specially, to meet the real-time requirement for the whole system, we also do data preprocessing and feature selection by Spark/Hadoop computing platform. In summary, the main contributions of this work could be summarized as follows:

(1) We propose the DTW algorithm to classify traffic flow data, which is beneficial to improve the accuracy of traffic flow prediction.

(2) We use the combination of WT and ESN to predict the traffic flow data without regularity. For the regular data, we adopt XGBoost algorithm to do prediction. In other words, for different categories of data, we select different adaptable algorithms to do the traffic flow prediction.

(3) We design the novel real-time system for traffic flow prediction, and it will favor us in handling traffic flow data automatically and quickly.

The rest of this paper is organized as follows. In Section 2, we discuss the related work in this area. In Section 3, system overview is described in detail. In Sections 4 and 5, we respectively introduce the traffic flow classification and prediction models. In Section 6 we give our experimental results. Finally, we give the conclusion in Section 7.

2 Related Work

There are some existing works that focus on the study of traffic flow data prediction and classification. Generally speaking, existing works usually do not consider different types of traffic trends or diagrams, and often use one single algorithm or its extensions to do the prediction. Actually, how to design a suitable system for different types of traffic data prediction is still an open problem for network service providers.

For one thing, the main solutions of traffic flow classification are divided into two categories, which are machine learning classification algorithms and time-series statistical similarity measurement. Williams et al. [7] introduce supervised algorithms which are Bayesian Network, C4.5 Decision Tree, and Naïve Bayes for traffic flow classification. Experimental results show that these supervised learning algorithms have better performances in the area of classification compared with other competing models. Moreover, Hu et al. [8] and Hochst et al. [9] both make use of the clustering algorithm which is the typical type of unsupervised learning algorithms for classification. In addition, some scholars concentrate on statistical learning to measure the similarity between different traffic diagrams. For example, Jeong et al. [10] propose weighted dynamic time warping (WDTW) for time series classification, which is a novel distance measurement for traffic data. Compared with machine learning algorithms for classification, statistical learning methods like DTW pay more attention on the similarity measurement in the shape of traffic diagrams.

For another, owing to the strong generality of boosting and decision tree, Chen et al. [11] propose the XGBoost algorithm to improve the accuracy of classification and regression. Moreover, with the huge development of neural networks, Zhu et al. [12] propose a novel BP neural network model for traffic prediction, which brings better learning ability than some common machine learning algorithms. However, faced with complex non-linear factors to traffic data, these algorithms would not perform better than other competing models for all of traffic diagrams. Therefore, Yang et al. [13] make full use of WT method to decompose the traffic flow data into different coefficients, and use radial basis function network for traffic prediction.

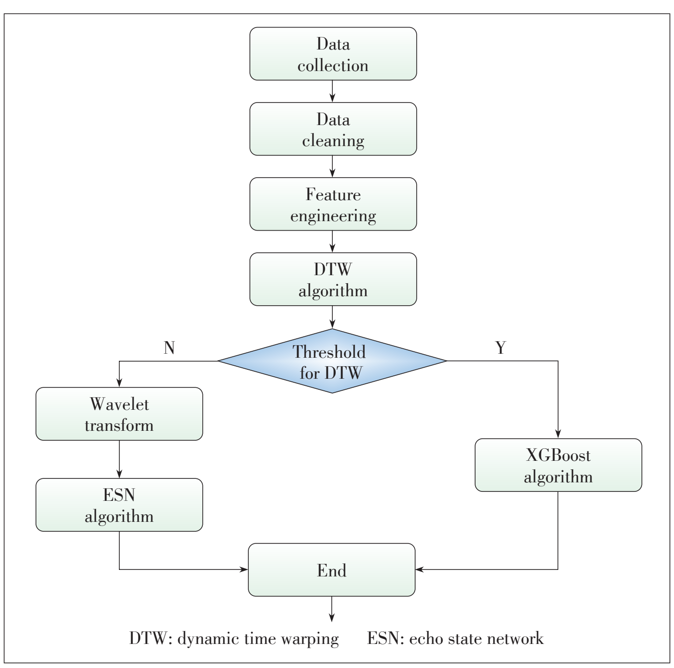
In summary, most of existing works do not concern about the classification of traffic data, but only use some techniques to do the optimization. Moreover, there is less study about the real-time system for traffic flow prediction and management. Based on these analysis, we design a novel real-time for both traffic classification and prediction.

3 System Overview

3.1 System Brief Description

For the purpose of the whole system for traffic flow data treatment and management, we fully consider the complete process for traffic flow data, including data preprocessing, traffic flow classification and traffic flow prediction. The flow diagram of framework for system is shown in **Fig. 2**.

Briefly speaking, we firstly do data preprocessing containing data cleaning and feature engineering by Spark/Hadoop computing platform, which will greatly improve the speed of data processing. Then, the processed data would be sent to traffic flow classification model. The realization of this model is based on DTW algorithm. After performing by DTW, if the ratings of traffic flow data are above the threshold value, they are taken as data with regularity. And they will be directly transferred to do the prediction by Xgboost algorithm. Otherwise, they are classified as data with non-regularity. They are decomposed into different coefficients by wavelet transform and then do the traffic prediction by ESN algorithm. Finally, we could get the promising results through this novel system even if the traffic flow data has different categories.



▲**Figure 2. Algorithm flow chart of the proposed novel real-time system.**

3.2 Data Collection and Preprocessing

For our study, the dataset we used comes from ZTE, which is one of the largest service providers in China. In particular, the traffic flow data in the dataset is from 5 762 different base stations. The specific meanings of the indicators are introduced in **Table 1**. Moreover, the collecting time of the data ranges from May 1, 2015 to June 20, 2017, and the order of magnitude for traffic flow data reaches GB.

Essentially, efficient and real-time system depends on not only the optimization of learning algorithms, but also the high speed of data processing. Owing to redundant and duplicate values in the dataset, we must do data preprocessing before traffic flow data classification and prediction. To satisfy the efficient requirement, we make full use of Spark/Hadoop computing platform to do the data preprocessing, including data cleaning and feature engineering. Specifically, Spark/Hadoop computing platform [14], [15], on behalf of the distributed computing, is a unified analytics engine for large-scale data processing. For so much missing data and outliers in the dataset, we clean these data by Spark/Hadoop. In addition, we process and perform the analysis of features which may influence the regularity of the traffic flow data.

As we all know, the important element of the traffic flow analysis is the inspection of the traffic flow fluctuations, related to the following factors:

- Time: time of day, day of week, month of year, etc.
- Base station position: school, high-speed railway station, business zone, etc.
- Traffic type: long term evolution (LTE), optical line terminal (OLT), etc.

Therefore, we observe that these following specific features are vital to the traffic flow prediction.

(1) The concreate date of the traffic data. For example, 2016-05

▼**Table 1. Data attributes**

Indicators	Meanings
Noid	ID of the base station
Name	Location information of the base station
Time	Collecting time of the data
kb	Traffic value of the base station by one day
Area	Administrative region of the base station

-10, etc.

(2) Whether the day is the working day plays an important role in prediction. For example, 2016-05-10 is the working day and we will mark this day 0 to represent working day. Otherwise, we could label 1.

(3) The day before a week is also one factor to do the prediction.

(4) The day before one year. For example, if we want to predict traffic flow data at 2016-05-10, the traffic data at 2015-05-10 will be this new feature.

(5) The sliding window of the current day derives some new features. From our analysis, we select current day before 1 day, 3 days and 5 days as new features. For instance, we will select 2016-05-09, 2016-05-07 and 2016-05-05 as new features to predict traffic flow data at 2016-05-10.

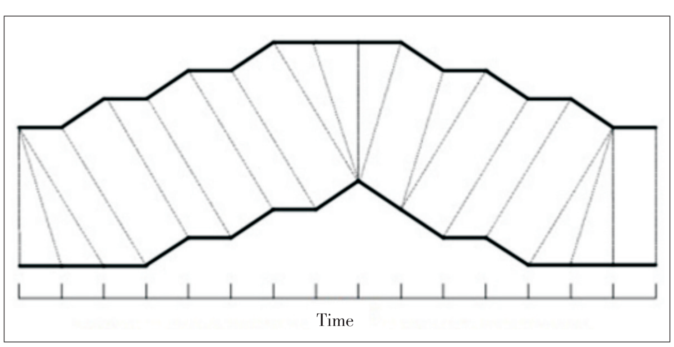
Theoretically, we should have selected base station's positions as a feature. To our disappointment, the base stations' positions are all different codes, and it is hard for us to translate the positions to real place in the map. Therefore, we do not deal with these values in our feature engineering and experiments.

4 Traffic Flow Classification Model

4.1 Introduction of DTW Algorithm

Dynamic time warping (DTW) is the algorithm to compare the similarity between two time series, which finds the minimum path by providing non-linear alignments. It has been applied into many fields, such as speech processing, time series classification, and clustering.

Briefly speaking, the methodology for DTW is as follows. For example, **Fig. 3** shows the optimal warping path between two different time series. We assume that one time series A has the length of p. We also assume another time series B with the length q, which is a different length from the time series A. Subsequently, we create the p-by-q path matrix where the element in the matrix represents the distance between two points. More importantly, there are also several constraints in building warping path [16]. The best match between two sequences is the one path with the lowest distance. Specifically, the optimal warping path could be found as the following equations:



▲**Figure 3. Optimal warping path between two different time series.**

$$DTW_p(A,B) = \sqrt[p]{\gamma(i,j)}\ , \tag{1}$$

where $\gamma(i,j)$ is the cumulative distance and the index p means the l_p norm in the equation. Generally speaking, most of us use Euclidean distance to calculate distance, where p equals two. Specifically, $\gamma(i,j)$ is described by

$$\gamma(i,j) = \left|a_i - b_j\right|^p + \min\{\gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1)\}, \tag{2}$$

where a_i and b_j are from two sequences with different lengths.

In summary, the DTW algorithm could accurately measure the similarity between time series even if two time series have different lengths.

4.2 Ratings of Different Traffic Data

In our system, we will select some typical traffic flow diagrams with obvious regularity as the templates, and these traffic flow data could be estimated and predicted easily. The DTW algorithm could measure the similarity between the other traffic flow data and the templets. Finally, we could get the ratings from the DTW algorithm. The distributions of the ratings are shown in **Table 2**.

We are surprised to find from the traffic flow diagrams that the lower rating the traffic flow diagram has, the more obvious regularity it has. On the basis of our analysis for the distributions of the ratings, we set the ratings to 5, which is the threshold for the regularity of traffic flow data. In other words, if the ratings of the traffic flow data beyond 5, we consider it regular in the dataset. **Fig. 4** shows the specific traffic flow diagram.

5 Integrated Traffic Flow Prediction Model

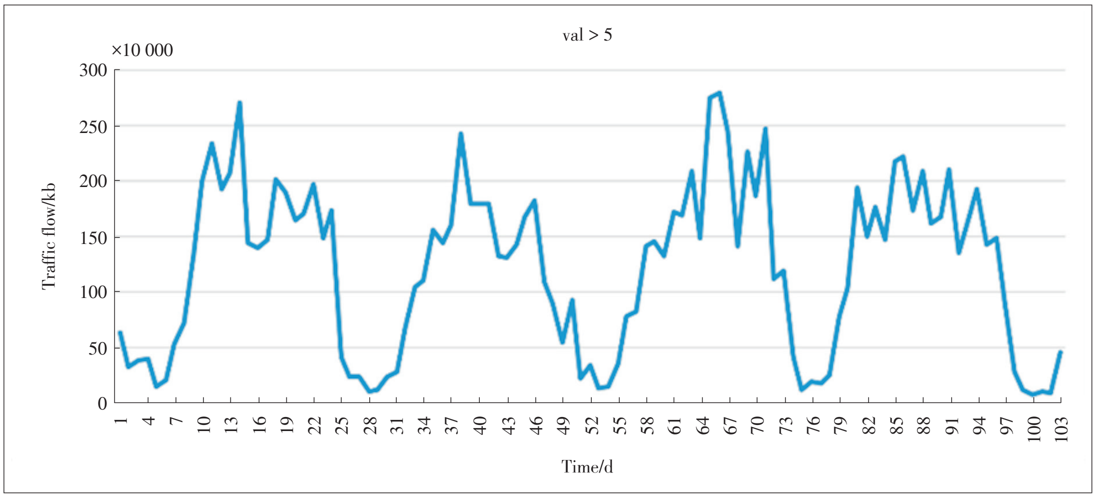
In this section, we mainly introduce the integrated traffic flow prediction model. For the classified traffic flow data, we adopt different algorithms for prediction, which are the XGBoost and WT-ESN. Therefore, this section is divided into two parts. One is for the traffic data with obvious regularity, and the other is for the other traffic data.

5.1 Prediction for Traffic Data with Strong Regularity

Figs. 1a and 1b display strong regularity of traffic flow data. For the business zone, the traffic flow reaches at the peak val-

▼**Table 2. Distributions of ratings for dynamic time warping (DTW)**

Ratings	Percentage
0–5	57%
5–10	15%
10–15	12%
more than 15	11%



◀Figure 4. Specific traffic flow for dynamic time warping (DTW).

ue at around 9 a.m−11 a.m. Accordingly, it decreases to the valley value at 1 a.m−4 a.m. Theoretically, machine learning algorithms could learn the knowledge and regularity from different datasets like the human, which is the main reason for its popularity. XGBoost algorithm [11], as the emerging algorithm in the supervised learning, has better learning ability compared with others supervised learning algorithms. In fact, XGBoost combines the ensemble method with the optimization of gradient descent. The following equations mainly introduce the core of this algorithm:

- Objective

$$Obj = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{K} \Omega\left(f_k\right). \tag{3}$$

- Objective, with constants removed

$$\sum_{i=1}^{n}\left[g_i f_t\left(x_i\right) + \frac{1}{2} h_i f_t^2\left(x_i\right)\right] + \Omega\left(f_t\right), \tag{4}$$

where $g_i = \partial_{\hat{y}^{(t-1)}} l\left(y_i, \hat{y}^{(t-1)}\right), h_i = \partial^2_{\hat{y}^{(t-1)}} l\left(y_i, \hat{y}^{(t-1)}\right)$. Moreover, in Eqs. (3) and (4), y_i means the true value and $\hat{y}_i$ is the predicted value. l represents the loss function and Ω is the penalty term. f is the regression tree. Compared with traditional optimization methods like Decision Tree, Gradient Boosting Decision Tree (GBDT) [17], XGBoost uses the second order Taylor expansion to approximate the objective method, which could optimize it quickly in the general setting.

5.2 Prediction for Traffic Data with Non-Regularity

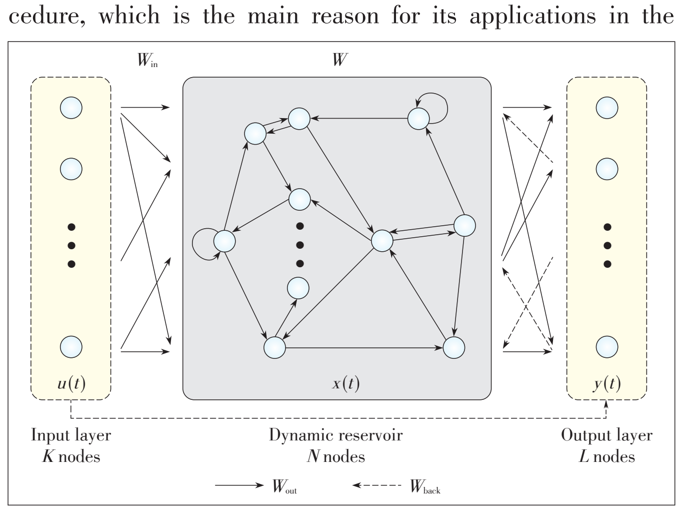
As mentioned before, we consider that the traffic flow data has regularity when the rating of traffic flow data beyond 5. Faced with these data, we believe that the combinations of nonlinear factors result in this phenomenon. Therefore, we adopt WT to decompose the traffic flow data into different coefficients, which could predict it separately. Moreover, with the better generality and fast computing in neural networks, we choose a kind of recurrent neural network called ESN to do each prediction.

ESN is a kind of dynamic neural network, which mainly includes input layer, middle layer and output layer. The middle layer is also called dynamic reservoir and it contains a large amount of sparse connected neurons, which plays the important role in short-term memory achievement and prediction. The structure of network is shown in the **Fig. 5**. Specifically, the memory and prediction function depend on two equations called the state update equation and the output state equation.

$$x(t+1) = f\left(W_{in} \cdot u(t+1) + W \cdot x(t)\right), \tag{5}$$

$$\hat{y}(t+1) = f^{out} \cdot \left(W_{out} \cdot \left(u(t+1), x(t+1), y(t)\right)\right), \tag{6}$$

where $x(t)$ is the state of reservoir at time t and $u(t)$ is the state of input. f is the neuron activation function. W_{in} and W are the input and middle layer weight matrix. $y(t)$ is the output at time t and $\hat{y}(t+1)$ is the predicted output at time $t+1$. Essentially, ESN could avoid the shortcomings of typical gradient-descent-based RNN algorithm and have the simple learning procedure, which is the main reason for its applications in the



▲Figure 5. Network structure of echo state network (ESN).

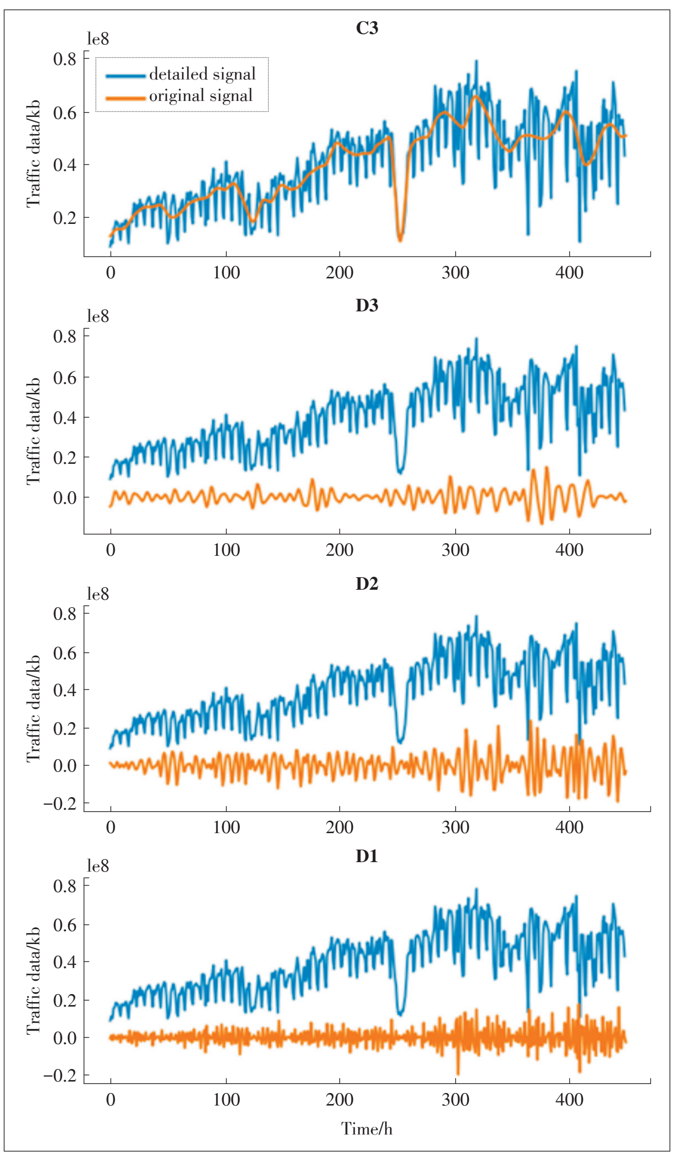
field of time series prediction. However, faced with traffic data without regularity, ESN could not accurately predict each type of traffic data well for its different regularities.

To solve these issues, we make efforts to adopt wavelet transform (WT) to decompose the initial traffic flow data into different coefficients, which could reduce the complexity of prediction. Briefly speaking, WT is a method of transform analysis and its main characteristic is to analyze the localization of frequency. For example, **Fig. 6** shows the details of WT, where the blue one represents original signal and yellow ones represent different degrees of details. Specifically, C3 is the lowest frequency factor among all coefficients, and it has the familiar trend with original signal. D1, D2 and D3 are all the high-frequency parts of the original signal and we use ESN algorithm to predict them. For non-regularity traffic flow data, we utilize Haar function as the core of the WT. The concrete steps of non-regularity traffic flow data prediction are described as follows:

Step 1: WT decompose the initial non-regularity data y_i into four coefficients called C3, D1, D2 and D3, which are the inputs of the ESN.

Step 2: Through single reconstruction, four coefficients including C3, D1, D2 and D3 are respectively sent to ESN for network parameters W adjustment. And it receive the predicted value for each single line.

Step 3: After training ESN, four predicted values are aggregated as the final output, which is also the prediction of the non-regularity traffic flow data $\hat{y}_i$.



▲Figure 6. Three-layer wavelet decomposition and single reconstruction.

6 Results and Analysis

In this section, we will conduct our experiments on the dataset collected from ZTE. We will first introduce two evaluation indexes in detail and then show and analyze the experimental results.

6.1 Evaluation Indexes Introduction

Here we briefly introduce two evaluation indexes called Normalized Root Mean Square Error (NRMSE) and R-Squared (R^2), which are the most common evaluation indexes in traffic flow prediction. Specifically, the definitions of NRMSE and R^2 are shown as follows:

$$NRMSE=\sqrt{\frac{\sum_{t=1}^{n}\left(\hat{y}_t-y_t\right)^2}{n}}/\left(y_{\max}-y_{\min}\right), \tag{7}$$

$$R^2=1-\sum\left(\hat{y}_i-\bar{y}\right)^2/\sum\left(y_i-\bar{y}\right)^2, \tag{8}$$

where $\hat{y}$ means the predicted value and y_t is the true value. On one hand, lower values of NRMSE represent better accuracy of the model. On the other hand, R^2 indicates the degree of fit for traffic flow data, and thus higher values of R^2 represent better degree of fit. Compared with NRMSE, R^2 focuses on the similarity of traffic flow data. In our experiments, we put more emphasis on NRMSE, because accuracy is our first choice.

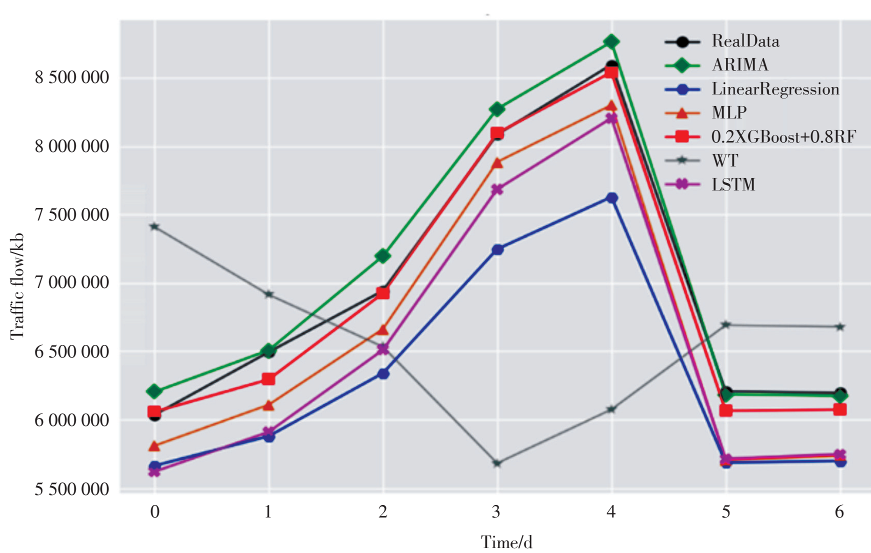
6.2 Analysis of Results

As mentioned before, we select the dataset that ranges from May 1, 2015 to June 20, 2017 as a training set. In order to show the accuracy and degree of fit in detail, we choose one week's data from June 21, 2017 to June 27, 2017 as testing set. To prove the better performances of XGBoost for regular traffic flow data, we compare the NRMSE and the degree of fit with some supervised learning algorithms and other commonly used algorithms. From the results shown in **Table 3** and **Fig. 7**, we could conclude that XGboost has the best performances among all algorithms. Specifically, the prediction of XGBoost fits the real data very well. From the perspective of evaluation

▼Table 3. Results of NRMSE for six algorithms

Algorithms	NRMSE
XGBoost	0.012
Linear regression	0.089
ARIMA	0.017
MLP	0.050
Wavelet transform	0.157
LSTM	0.062

ARIMA: autoregressive integrated moving average
LSTM: long short-term memory
NRMSE: normalized root mean square error
MLP: multi-layer perception



ARIMA: autoregressive integrated moving average
LSTM: long short-term memory
MLP: multi-layer perception
WT: wavelet transform

▲Figure 7. Degrees of fit for different algorithms.

index, XGboost has the lowest value of normalized root mean square error (NRMSE) among all algorithms, which means high accuracy of prediction.

Moreover, we compare performances of our integrated model with traditional algorithms for traffic flow data prediction, including autoregressive moving average model (ARMA) time series analysis, LSTM and WT algorithm. **Tables 4** and **5** show

▼Table 4. Results of Normalized Root Mean Square Error (NRMSE) in the Experiments

MAE	Integrated model	ARMA	WT	LSTM
≤15%	60.12%	22.9%	51.3%	56.6%
≤25%	81.14%	35.4%	70.5%	75.1%
≤30%	86.06%	52.6%	77.4%	82.5%
≤50%	93.23%	74.6%	86.5%	90.0%
≤1	96.98%	95.0%	95.3%	96.0%

ARMA: autoregressive moving average model
LSTM: long short-term memory
MAE: mean absolute error
WT: wavelet transform

▼Table 5. Results of R^2 in the Experiments

R^2	Integrated model	ARMA	WT	LSTM
≥0.8	8.5%	6.11%	7.50%	7.46%
≥0.5	19.46%	15.2%	18.4%	17.39%
≥0	59.32%	41.0%	59.0%	59.1%

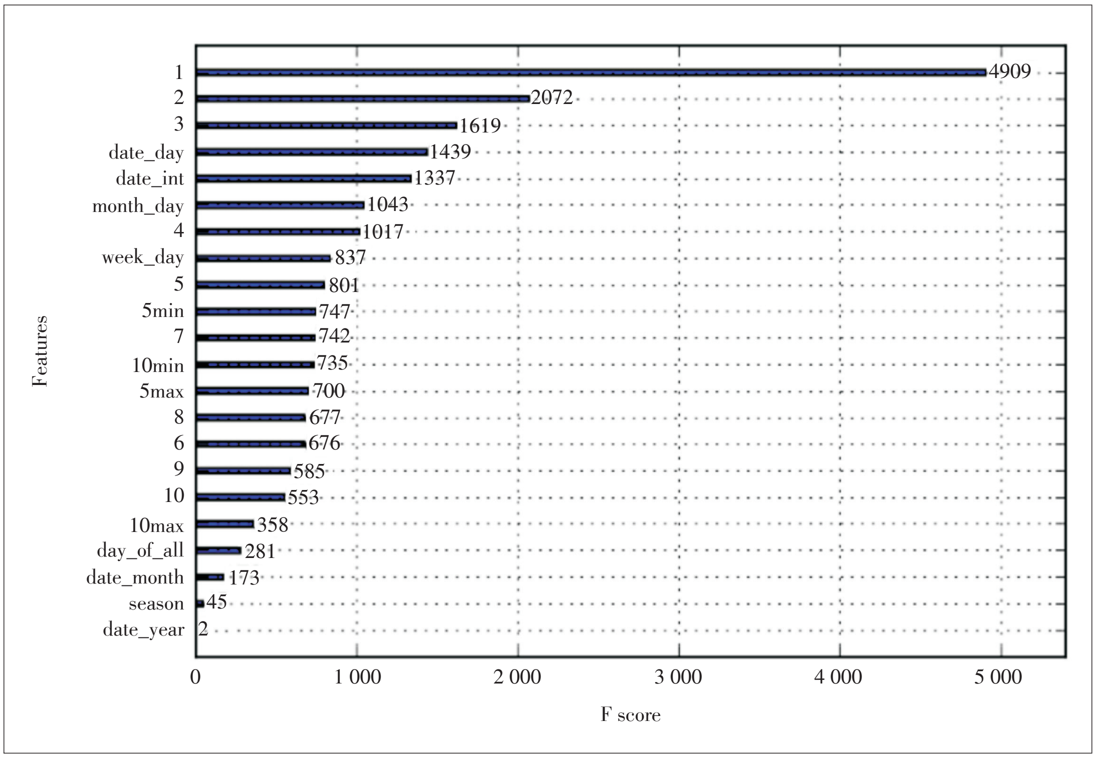
ARMA: autoregressive moving average model
LSTM: long short-term memory
WT: wavelet transform

the results of two evaluation indexes mentioned before for our integrated model and four extra algorithms. From these tables, we could conclude that our integrated model has the lowest values of NRMSE that correspondingly represents highest accuracy among all algorithms. In addition, another evaluation index R^2 for traffic flow prediction pays more attention on the degree of fit for them. As shown in the tables, our integrated model has better degree of fit than others, which means the shape of traffic flow prediction is more similar with the true data. Generally speaking, XGBoost and WT-ESN have their own advantages but it is impossible for them to accurately predict all types of traffic flow data. In this case, our integrated model makes full use of advantages of these two algorithms, obtaining the best prediction performance.

On the theoretical aspect, we analyze the feature importance and contributions to the predicted value among all features, which are also the interesting aspects for service providers. Among all features referred in Section 3.2, we extract top-5 features influencing the predicted value, as shown in **Fig. 8**. In particular, F-score in this figure measures the importance of the features, and higher value represents more effective feature. The concrete meanings of these features are presented in **Table 6**. This could help us know what features have deep impact on the predicted value. Finally, **Fig. 9** shows the overview of our system. There are mainly four parts in our system, including raw data overview, data statistical analysis, data classification, and data prediction. When the raw data is processed in the background, relevant information will be displayed in this system.

7 Conclusions

Traffic flow prediction has been a hot research topic in recent years. However, how to systematically realize real-time traffic prediction in the context of big data era is urgent to be solved. In this paper, we propose a system for traffic flow prediction and classification. Compared with traditional algo-



◀**Figure 8.**
Feature importance rankings in the experiment.

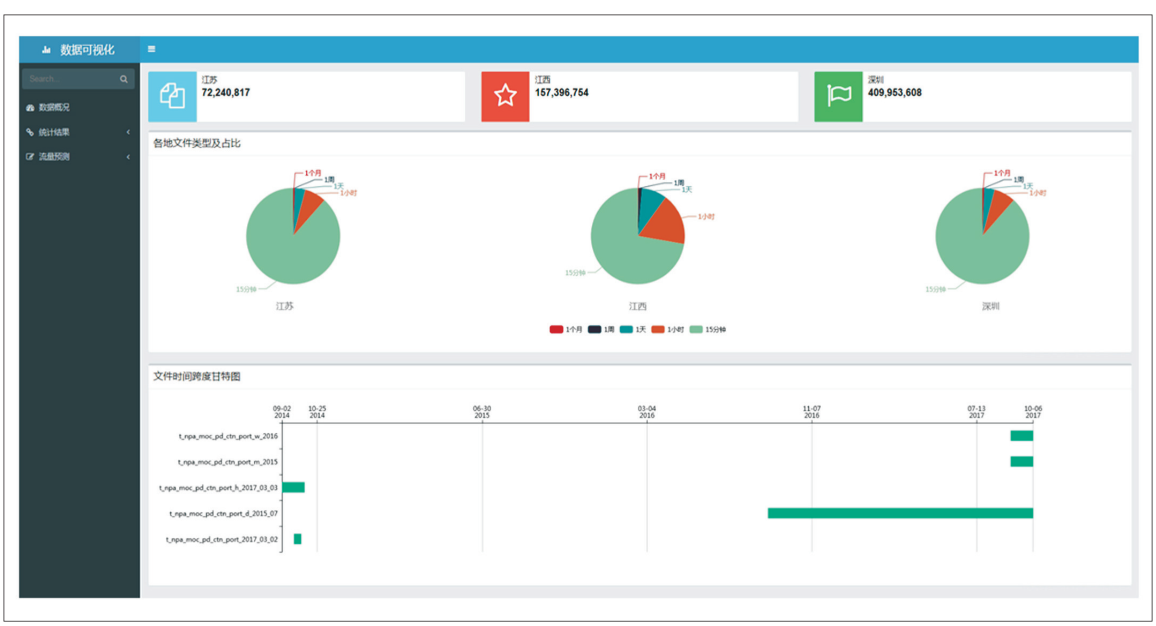
▼**Table 6. Feature selection in the experiment**

Features	Meanings
1	Average values one week before
2	Average values two weeks before
3	Average values three weeks before
Week_day	Today is a weekend or working day
Date_int	The concrete date of today

rithms or models for traffic flow prediction, our system firstly utilizes DTW algorithm for classification, which is a method to solve different trends of traffic flow prediction. Then, we mainly make full use of XGBoost and ESN to predict traffic flow data. Experiment results show that our system has better performances than other traditional algorithms.

References

[1] ZHOU L, WU D, DONG Z J, et al. When Collaboration Hugs Intelligence: Con-



◀**Figure 9.**
Overview of the system.

tent Delivery over Ultra-Dense Networks [J]. IEEE Communications Magazine, 2017, 55(12): 91–95. DOI: 10.1109/mcom.2017.1700481
[2] ZHOU L, WU D, CHEN J X, et al. Greening the Smart Cities: Energy-Efficient Massive Content Delivery via D2D Communications [J]. IEEE Transactions on Industrial Informatics, 2018, 14(4): 1626–1634. DOI: 10.1109/tii.2017.2784100
[3] WU D, WANG J L, HU R Q, et al. Energy-Efficient Resource Sharing for Mobile Device-To-Device Multimedia Communications [J]. IEEE Transactions on Vehicular Technology, 2014, 63(5): 2093–2103. DOI: 10.1109/tvt.2014.2311580
[4] ZHAO T S, LIU Q, CHEN C W. QoE in Video Transmission: A User Experience-Driven Strategy [J]. IEEE Communications Surveys & Tutorials, 2017, 19(1): 285–302. DOI: 10.1109/comst.2016.2619982
[5] PAMULA T. Classification and Prediction of Traffic Flow Based on Real Data Using Neural Networks [J]. Archives of Transport, 2012, 24(4): 519–529. DOI: 10.2478/v10174-012-0032-2
[6] GUO F C, POLAK J W, KRISHNAN R. Comparison of Modelling Approaches for Short Term Traffic Prediction under Normal and Abnormal Conditions [C]//13th International IEEE Conference on Intelligent Transportation Systems. Funchal, Portugal, 2010: 1209–1214. DOI: 10.1109/ITSC.2010.5625291
[7] WILLIAMS N, ZANDER S, ARMITAGE G. A Preliminary Performance Comparison of Five Machine Learning Algorithms for Practical IP Traffic Flow Classification [J]. ACM SIGCOMM Computer Communication Review, 2006, 36(5): 5. DOI: 10.1145/1163593.1163596
[8] HU C C, YAN X H. Mining Traffic Flow Data Based on Fuzzy Clustering Method [C]//Fourth International Workshop on Advanced Computational Intelligence. Wuhan, China, 2011: 245–248. DOI: 10.1109/IWACI.2011.6160011
[9] HÖCHST J, BAUMGÄRTNER L, HOLLICK M, et al. Unsupervised Traffic Flow Classification Using a Neural Autoencoder [C]//IEEE 42nd Conference on Local Computer Networks (LCN). Singapore, Singapore, 2017: 523–526. DOI: 10.1109/LCN.2017.57
[10] JEONG Y S, JEONG M K, OMITAOMU O A. Weighted Dynamic Time Warping for Time Series Classification [J]. Pattern Recognition, 2011, 44(9): 2231–2240. DOI: 10.1016/j.patcog.2010.09.022
[11] CHEN T, GUESTRIN C, XGBoost: A Scalable Tree Boosting System [C]//ACM International Conference on Knowledge Dicovery and Data Mining. San Francisco, USA, 2016: 785–794
[12] LI Z, LEI Q, XUE K Y, et al. A Novel BP Neural Network Model for Traffic Prediction of Next Generation Network [C]//Fifth International Conference on Natural Computation, Tianjin, China, 2009: 32–38. DOI: 10.1109/ICNC.2009.673
[13] YANG W, YANG D Y, ZHAO Y L, et al. Traffic Flow Prediction Based on Wavelet Transform and Radial Basis Function Network [C]//International Conference on Logistics Systems and Intelligent Management (ICLSIM). Harbin, China, 2010: 969–972. DOI: 10.1109/ICLSIM.2010.5461098
[14] ZAHARIA M, CHOWDHURY M, FRANKLIN M, et al. Spark: Cluster Computing with Working Sets [C]//2nd Usenix Conference on Hot Topicsin Cloud Computing. Boston, USA, 2010: 10–10
[15] DEAN J, GHEMAWAT S. MapReduce: Simplified Data Processing on Large Clusters [C]//6th Symposium on Operating Systems Design & Implementaion. San Francisco, USA, 2004: 137–149
[16] SAKOE H, CHIBA S. Dynamic Programming Algorithm Optimization for Spoken Word Recognition [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1978, 26(1): 43–49. DOI: 10.1109/tassp.1978.1163055
[17] MA X L, DING C, LUAN S, et al. Prioritizing Influential Factors for Freeway Incident Clearance Time Prediction Using the Gradient Boosting Decision Trees Method [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(9): 2303–2310. DOI:10.1109/tits.2016.2635719

Biographies

YE Dezhong received his B.S. degree from Jilin University, China in 1996. He is currently a chief big data engineer at ZTE Corporation. His research interests include big data in wireline communication and fixed network product.

LV Haibing received his B.S. degree from China University of Mining and Technology, China in 2002. He is currently a wireline product architecture director at ZTE Corporation. His research interests include fixed network products and IPTV systems.

GAO Yun received his B.E. degree from Nanjing University of Posts and Telecommunications (NUPT), China in 2016. He is currently pursing Ph.D. degree in NUPT. His research interests include machine learning, deep learning, and QoE in multimedia communication.

BAO Qiuxia received her B.E. degree from Nanjing University of Posts and Telecommunications (NUPT), China in 2017. She is currently pursing M.S. degree in NUPT. Her research interests include machine learning in big data, etc.

Chen Mingzi (1018010415@njupt.edu.cn) received her B.S. degrees in computer science from both New York Institute of Technology, USA and Nanjing University of Posts and Telecommunications (NUPT), China in 2018. She is currently pursing Ph.D. degree in NUPT. Her research interests include machine learning on big data, QoE in multimedia communication, and 5G tactile internet toward artificial intelligence.