New Method Research on Prediction Model for Busy Telephone Traffic

1Li Jiang-Bao, 1Jia Zhen-Hong, 1Qin xi-Zhong, 2Sheng Lei and 2Chen Li
1School of Information Science and Engineering, Xinjiang University, 830046, Urumqi, China
2Subsidiary Company of China Mobile In Xinjiang, 830063, Urumqi, China

Abstract: In order to improve the prediction accuracy of busy telephone traffic, this study proposes a busy telephone traffic prediction method that combines wavelet transformation and Least Square Support Vector Machine (LSSVM) model which is optimized by Particle Swarm Optimization (PSO) algorithm. Firstly, decompose the pretreatment of busy telephone traffic data with Mallat algorithm and get low frequency component and high frequency component. Secondly, reconfigure each component and use PSO_LSSVM model predict each reconfigured one. Then the busy telephone traffic can be achieved. The experimental results show that the prediction model has higher prediction accuracy.

Key words: Combined model, wavelet transformation, least square support vector machine, particle swarm optimization algorithm, prediction

INTRODUCTION

The number of mobile users has been continuing to increase in wireless communication, especially in holidays, mobile communication network are faced with the great influence of heavy telephone traffic. In order to prevent network congestion and improve the utilization ratio of network resources, so it has an important significance to predict the future telephone traffic accurately. Traditional linear prediction model of time series such as ARIMA (Chen et al., 2010) model and FARIMA (Chen and Wang, 2010) model can capture long and short data-related characteristics, but for non-linear and non-stationary traffic data, the methods do not forecast accurately. Neural network model (Deng et al., 2008; Zhang et al., 2011) reflect the inherent law of data through the training but the neural network is based on empirical risk mininization principle and it is easy to lead to less learning or more learning phenomenon in the training process. Literature (Mou et al., 2010) provides a new method for the large call center traffic prediction, but this method is not suitable for the holiday traffic. The method of Support Vector Machine (SVM) is based on structural risk minimization. It can better solve the data of the small sample, nonlinear, high dimension and other characteristics. But there are two problems about the traditional SVM (Elattar et al., 2010; Han and Jia, 2011; Bemolle and Rossi, 2009). One is that kernel function iteration error accumulation will lead to inaccurate results, the other is that Parameters are difficult to determine. In order to solve two problems, this study uses the optimized forecasting method of least squares support vector (LS_SVM) machine. Firstly, the method use equality constraints of least squares support vector machine (Shao and Ma, 2011; Li and Li, 2012) to replace the inequality constraints of SVM to solve the quadratic programming problem which is transformed into the problem of linear equations, so as to simplify the model structure and the training speed is improved. Secondly, use the improved Particle Swarm Optimization (PSO) algorithm to solve the problem of parameters of LS_SVM model. The wavelet transform has strong Multi-scale analysis capability. It can remove the correlation of traffic data. The literatures (Chen and Liu, 2011; Feng and Liu, 2011; Zhou et al., 2011) use the combined forecasting model with wavelet, the combination model can achieve better prediction. The single prediction model can't solve various characteristics of telephone traffic, so this study is combined with the wavelet prediction model.

For the non-stationary, self-similarity and multi-scale characteristic of telephone traffic, this study puts forward a busy telephone traffic prediction model that combines wavelet transformation and LS_SVM model which is optimized by PSO algorithm (PSO_LSSVM). Firstly, use the 3-layer wavelet of Mallat algorithm (Shensa, 1992) decomposition on the busy traffic data to get the low-frequency component and high-frequency components. Secondly, reconfigure each component and use PSO_LSSVM model to predict single reconstructed components. Finally the prediction results are superimposed. There must be some errors after PSO_LSSVM prediction model. These errors in the low frequency part and high frequency part of the prediction may be positive or negative. In the last step for telephone

Corresponding Author: Li Jiang-Bao, School of Information Science and Engineering, Xinjiang University, 830046, Urumqi, China
traffic synthesis, it can make positive and negative error offset each other, so the improvement of the wavelet transform can achieve better prediction results. PSO_LSSVM, LS_SVM and Elman prediction methods are chosen as compared model in this study and the simulation experiment shows that the proposed prediction model has higher prediction accuracy.

TELEPHONE TRAFFIC PREDICTION ALGORITHM BASED ON IMPROVED WAVELET TRANSFORM AND PSO_LSSVM MODEL

Algorithm steps:

1. Preprocess the original telephone traffic data, then get \( x'(n) \)
2. Decompose \( x'(n) \) with three layers wavelet of Mallat algorithm and get low frequency component (a3) and high frequency components (d1, d2, d3)
3. Use Mallat algorithm to reconfigure low frequency component (a3) and high frequency components (d1, d2, d3), get A3 and D1, D2, D3
4. Use improved particle swarm algorithm to optimize the parameters of LS_SVM model, then get optimal prediction model (PSO_LSSVM)
5. Use PSO_LSSVM model to predict A3, D1, D2, D3, get A3' and D1', D2', D3'
6. Get the final forecast results: \( x'(n) = A3' + D1' + D2' + D3' \)

Data pretreatment: In mobile communication, there are some changes about the telephone traffic every day and the amount of telephone traffic is not the same at different times. Especially on holidays, the busy telephone traffic (peak value) has the largest effect on the mobile communication network. Collect 24 sample telephone traffic data every hour per day. Then compare the 24 data and treat the maximum as busy telephone traffic data. The statistics of telephone traffic data from September 1, 2011 to October 1, 2011 and from September 1, 2012 to October 1, 2012. It is a total of 62 days' traffic. Then get 62×24 data and 62 busy data. Decompose 62 busy data and reconfigure each component, then get A3 and D1, D2, D3 four components. Mobile traffic trends not only has close relationship with historical traffic data but also affected by other factors' change, so it is necessary to analyze the correlation of telephone traffic and put the factors as the training data, so that we can achieve the purpose of accurate prediction. In consideration of the busy VLR users, VLR boot users, pager number and system connection rate impacts on telephone traffic, treat the four factors as impact factor of the telephone traffic, then use the four factors and single reconstruction data to compose four new matrix 62×5. Conduct the front of 61 data as the training sample data to predict the final busy telephone traffic data (The National Day). If the collection of traffic data is \( x(n), n \in [1,61] \) the days, \( t \in [0,23] \) the sampling time every day) and assuming the largest traffic is \( x_{\text{max}}(n) \) per day, then use the busy data of \( X_{\text{max}}(1), X_{\text{max}}(2), \ldots, X_{\text{max}}(61) \) to predict The National Day of busy telephone traffic data.

Wavelet decomposition and reconstruction: In 1987, Mallat and some others put forward Mallat algorithm for rapid decomposition and reconstruction. The algorithm can solve the problem that the greater the alpha value, the more sample value. Literatures (Chen and Liu, 2011; Feng and Liu, 2011; Zhou et al., 2011) use the combination forecast model based on the Mallat algorithm of modified wavelet transform and obtain a better prediction result.

The decomposition and reconstruction of Mallat algorithm equation:

\[
\begin{align*}
\alpha_{j,i} &= H_{\alpha} \\
\beta_{j,i} &= G_{\beta}
\end{align*}
\]  \quad (1)

\[
\begin{align*}
a_{j,i} &= a_{j,\alpha} + d_{j,\beta}, \quad j = 0,1,2,3,4...
\end{align*}
\]  \quad (2)

The H is a low-pass filter and G is a high-pass filter. The original signal is decomposed into low frequency part and high frequency part by Mallat algorithm. The low frequency component reflects the outline characteristics and changing tendency of the busy telephone traffic data and high frequency component reflects the impact of dynamic factors such as random disturbance. Low frequency part can be further decomposed, so we can obtain the new low frequency component and high frequency component. Single reconstruction is not to reconstruct the low frequency part and high frequency part at the same time but separately reconstructed. That's to say when a certain part is reconstructed that the other part should be set to zero.

The difficulty of wavelet decomposition and reconstruction lies in the selection of wavelet basis. This study uses different wavelet basis to superpose the reconstruction of the low frequency and high frequency data. Then compare with the original data and find that when using biol.3 as wavelet basis, the error is \( 10^{-11} \) (Fig. 1). While using wavelet basis such as dbN and sym, the error is \( 10^{-8} \). So this study chooses biol.3 as wavelet basis.
Fig. 1: Data error of using bio1.3 as wavelet basis

**Model of LS-SVM:** The main idea of nonlinear regression of Least Squares Support Vector Machine (LSSVM) is that the input data is mapped to high dimensional feature space \( \psi \) through the nonlinear mapping \( \psi (\bullet) \), so low dimensional nonlinear regression problem is transformed into linear regression problems in high dimensional feature space. The principle is that we assume that there are \( m \) sample data, such as \((x_i, y_i)\), \((x_2, y_2)\) \(\ldots\) \((x_m, y_m)\), and so on. \( x_i \in \mathbb{R} \) is sample input and \( y_i \in \mathbb{R} \) is the sample output. Construct the optimal regression function:

\[
f(x) = \mathbf{w}^T \bullet \psi(x) + b
\]  
\[  \tag{3} \]

The \( \mathbf{w}^T \in \mathbb{R}, b \in \mathbb{R}, b \) is bias, \( \psi (\bullet) \) : a run-H nuclear space mapping function.

According to the principle of solving objective and the structural risk minimization, use Lagrange multiplier method to solve optimization problem. The optimization problem of a Lagrange function for least squares support vector machine is:

\[
L(w, b, \xi, \alpha) = \frac{1}{2} ||\mathbf{w}||^2 + C \sum_{i=1}^{m} \xi_i^2 - \sum_{i=1}^{m} \alpha_i (\mathbf{w} \cdot \psi(x_i) + b + \xi_i - y_i)
\]  
\[  \tag{4} \]

\( C \) is regularization parameters, \( \xi_i \) is the error function, \( \alpha_i \) is the Lagrange multipliers, \( i = 1, 2, 3 \ldots M. \)

According to the KKT optimization conditions, get the partial derivative from the type (4) and make the result 0. Use radial basis kernel function (RBF) to solve the problem of high dimensional calculation. The optimization problem can be further transformed into the solution of linear equations:

\[
\begin{bmatrix}
0 & e^T \\
e & \Omega + \frac{1}{2} C \cdot I
\end{bmatrix}
\begin{bmatrix}
b \\
\alpha
\end{bmatrix}
= \begin{bmatrix}
0 \\
Y
\end{bmatrix}
\]  
\[  \tag{5} \]

where, \( e = [1, 1 \ldots 1], \alpha = [\alpha_1, \alpha_2 \ldots \alpha_m], Y = [y_1, y_2 \ldots Y_m], \Omega = k (x, x_i) \).

The final nonlinear regression model is:

\[
f(x) = \sum_{i=1}^{m} \alpha_i K(x, x_i) + b
\]  
\[  \tag{6} \]

And:

\[
K(x, x_i) = \exp\left[-(x_i - x)^2 / 2\sigma^2\right]
\]

**Improved particle swarm algorithm:** The shortcoming of LS-SVM model is that regular parameter and Kernel function are difficult to determine. The two parameters are mapped a group of particles in this study. Then get optimal particle with the improved particle swarm algorithm, so the problem of parameter is solved.

In order to improve the convergent performance of the basic PSO algorithm, this study adopts the evolution equation where contraction factor \( \chi \) is introduced to the speed:

\[
v_i(t+1) = \chi (v_i(t) + c_1 r_1(t) [p_i(t) - x_i(t)]) \\
+ c_2 r_2(t) [p_g(t) - x_i(t)]
\]  
\[  \tag{7} \]

\( \chi \) is contraction factor. \( c_1, c_2 \) is the acceleration. \( c_1 = 2 \) and \( c_2 = 2 \). \( r_1 \) and \( r_2 \) are two independent random variables and they are obeying uniform distribution between 0 and 1. \( t \) is the times of generation about particles. \( x_i \) is the current location of the particles i. \( p_i \) is the best location where particles i has been experienced. In the Eq. 7:

\[
\chi = \frac{2}{2 - \lambda \sqrt{\lambda^2 - 4\lambda}}
\]  
\[  \tag{8} \]

And \( \lambda = c_1 + c_2, \lambda > 4. \)

This experiment adopts the shrinkage factor to control the speed of particles. Define \( c_1 - c_2 = 2.05 \), then get \( \lambda = 4.1 \). Finally through Eq. 7, get \( \chi = 0.73 \). Presetting the size of the search space in the global search of particles, so that it can prevent the shrinkage factor of PSO can not meet the global extreme value point within a given number of iterations. This study iteration is 10. When, for example, forecasting the low frequency components, the online 10 times iterations convergence diagram of the optimal prediction model can be shown in Fig. 2.
ANALYSIS OF SIMULATION EXPERIMENT

Source of telephone ADTA: The experimental data is provided by a subsidiary company of China mobile in Xinjiang. It collected the traffic data since September 1, 2011 to October 1, 2011 and September 1, 2012 to October 1, 2012 from the Yili Prefecture of Xinjiang. It has 62×24 data totally, then extract 62 busy telephone traffic data. Generally speaking, the more data is selected, the more correctly the learning and training results reflect the relationship between the input and output and the higher precision of prediction is. However, in practical applications, it is impossible to increase the sample data without restrictions, in which case it should try to select a representative sample. So the telephone traffic data was selected from the area of Yili, Xinjiang province (Fig. 3). The changed trend of telephone traffic data in September, 2012 was similar to the data in September, 2011. As a whole, the data was rising in shock type. Especially, considering the correlation between four factors and telephone traffic, the study made some of data pretreatment in this paper to guarantee that it would improve the efficiency of the algorithm without losing the original data characteristics. During the Natural Day holiday, the traffic is larger and on October 1 the traffic is the largest. This experiment is mainly to predict the busy traffic of October 1, 2012.

Wavelet decomposition and single reconstruction: The busy traffic variation tendency of former 30 days of October 1, 2012 is similar to 2011 and there is a strong correlation between them. So this study use the Mallat algorithm to decompose the 62 busy data for three layers, thus the purpose of decorrelation can be achieved and the traffic of low frequency components and high frequency components (Fig. 4) also can be got. Then reconfigure the low frequency component and the high frequency components respectively to the original level (Fig. 5). From the results of the decomposition and reconstruction, low frequency component conforms to the trend of traffic variation and have a rising trend on the whole and high frequency component is of the shock type change and can reflect the effects of random disturbance factors. With the increase of decomposition layers, low frequency and high frequency components become more and more smooth.

Prediction of each component by PSO_LSSVM model:
Set the busy VLR users, VLR boot users, pager number and system connection rate, respectively as $\eta_1(n), \eta_2(n), \eta_3(n)$ and $\eta_4(n)$. When predicting the low frequency component information ($A_0(n)$), namely put the $[\eta_1(n), \eta_2(n), \eta_3(n), A_0(n)]$ as input variables ($n$ is between 0 and 61). Finally the component predictive value of the last day (October 1, 2012) can be obtained by PSO_LSSVM model. The method to predict high frequency components is similar to the low frequency. Low frequency component ($A_0$) and the high frequency component ($D_1$, $D_2$, $D_3$) are predicted respectively as shown in Fig. 6. And then the predicted value of the respective components will be superimposed to obtain the prediction busy telephone traffic of the October 1, 2012.

Results analysis: This study chooses PSO_LSSVM, LS_SVM and Elman neural network prediction model as contrast model. The result of prediction is shown as Fig. 7. The method of analyzing relative error is used in this study. The error formula is $\xi = (x^* - x)/x$ ($x^*$ is prediction and $x$ is practical value). If the busy telephone traffic on October 1, 2012 is ‘$S$’ and low frequency
component \((A_i)\) and the high frequency component \((D_i, D_j, D_k)\) are predicted, respectively as \(A_i', D_j', D_k'\) and \(D_i'.\) Then the relative error of WT, LSSVM prediction model is 
\[
\xi = \frac{|(A_i + D_j + D_k + D_i') - S|}{S}
\]
The results of the prediction are shown as Fig. 7. Experiments were conducted 5 times in this study. The error of the prediction model for the statistics is shown as in Table 1. Compared with PSO_LSSVM model, the method of this study makes the prediction accuracy improve 3.09%, mainly because predicted deviation can be cancelled out when add each fraction together after wavelet transform and the error of predict results of each fraction either positive or negative, thus the prediction accuracy can be improved. Compared PSO_LSSVM model with LS_SVM model, prediction accuracy improve 2.08 percent. It's shown that PSO_LSSVM model can properly solve the problem of assurance of regular parameters and kernel.

Fig. 4(a-f): Three layer wavelet decomposition of busy telephone traffic data with Mallat algorithm

Fig. 5(a-d): Single reconstruction of the low frequency component and high frequency component with Mallat algorithm
Fig. 6(a-d): Prediction of the low-frequency(A3) and high-frequency(D1,D2,D3) with PSO_LSSVM model on October 1, 2012

Fig 7: Result of this study and compared model on October 1, 2012

function parameters. Compared with Elman model, prediction accuracy of the formal three methods improved 9.66, 6.57 and 4.49%, respectively. It’s shown that LS_SVM is more suitable than neural network to predict the busy traffic.

CONCLUSION

This study puts forward a kind of improved telephone traffic measurement algorithm based on wavelet transforms and PSO_LSSVM model. At first, decompose preprocessed telephone traffic data with Mallat algorithm and get low frequency component and high frequency component. Then reconfigure each component and use the PSO_LSSVM model to predict single reconstructed components. In the last step for telephone traffic synthesis, it can make positive and negative error offset each other, so the improvement of the wavelet transforms model can achieve better prediction results and more stable. The next step is to study the online prediction model based on multiple regression analysis to improve the timeliness.
ACKNOWLEDGMENT

First of all, I’d like to express my thankfulness to the judges for selecting my work among all the works of the contestants. The experimental data is provided by subsidiary company of China mobile in Xinjiang. The research project stems from China Mobile Group Xinjiang Limited Company Development Found Projects (Projects No.XJM2012-01). Finally, thanks for the guidance of my teacher.

REFERENCES