International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.3 (2016), pp.229-238
http://dx.doi.org/10.14257/ijmue.2016.11.3.22

The Application Research on Busy-Time Traffic Prediction Based ESNs of Multivariate Principal Components Analysis

Liu Jun-Xia[1,2] and Jia Zhen-Hong[1]

1.School of Information Science and Engineering, Xinjiang University, Urumqi, xinjiang province ,China
2.Department of Electrical and Information, Xinjiang Institute of Engineering, Urumqi, xinjiang province, China
251414185@qq.com, jzhh@xju.edu.cn

Abstract

As single variable cannot provide comprehensive forecast information, Combining with the communication traffic data come from Xinjiang branch of china mobile communication corporation, the traffic and the four variables highly related with the traffic are done principal component analysis, the first principal component and the second principal component that they are obtained by principal components analysis (PCA) and dimension reduction are as input variables of ESNs. Therefore, the busy-time traffic forecasting model based ESNs (echo state networks) of multivariate principal component analysis was established. The simulation results show that the new prediction method has good generalization performance. Compared with single variable ESNs prediction model and LS-SVM forecasting model, the presented traffic forecasting method ensure the prediction efficiency while improving the prediction accuracy.

***Keywords**: Multivariate, Principal Components Analysis, Echo State Networks, busy-time traffic, Prediction*

1. Introduction

With the rapid development of communication services, especially mobile communications network is facing high traffic pressure on holiday, accurate traffic forecasting demand is prominently increased. Accurate and efficient traffic forecasting not only be able to provide early warning for mobile operators and process busy cells in advance, but also provide a theoretical basis for mobile communication network planning.

In recent years, many new methods are applied for communication traffic prediction, such as least squares support vector machine model, auto-regressive moving average model, neural network models and their improved algorithm. To some extent, the prediction accuracy of traffic forecasting model and generalization capacity are improved [1-3]. However, mobile communication traffic data is typical time series. As it is highly non-linear, the prediction accuracy is improved while the prediction efficiency is reduced by those algorithms. ESNs (echo state networks) is a novel recurrent neural network model [4], compared with the traditional recurrent neural networks, it has higher computational efficiency and stability. The core of the ESN is a large reservoir. As the reservoir contains a large number of randomly and sparsely connected neurons, it has a good short-term memory capability and better performance in terms of time series forecasting [5-6], while it has been successfully applied to prediction traffic [7-9].

On the basis of the existing study, aiming at the problem that traffic fluctuations affected by many factors, so single historical traffic does not provide more comprehensive forecasting information, however, the existing ESNs traffic forecasting

ISSN: 1975-0080 IJMUE

methods only consider the single historical traffic, not involving multiple input variables. This paper established a busy-time traffic forecasting model of multivariate echo state network, the model combined with predict demand and data from Xinjiang branch of China Mobile communication corporation . What's more, the prediction accuracy, efficiency and generalization performance are further improved.

2. ESNs

Classical ESNs contains three layers: an input layer, a hidden layer and a output layer. The hidden layer is between input layer and output layer, it consists of a large number of sparsely and randomly connected neurons. So the reservoir exhibits some nonlinear dynamics properties and is able to deal with nonlinear problems well. The hidden layer is also called dynamic reservoir (DR). The output layer is made of some neurons in parallel; the number of the putout neurons is equal to the output number. Known network input, the reservoir state and output are updated as follows :

$$X(k) = f^{DR}(W^{in}u(k) + WX(k-1) + W^{fd}y(k-1) + v_f(k)) \tag{1}$$

$$y(k) = f^{RD}(W^{out}X(k)) \tag{2}$$

Where f^{DR} and f^{RD} are the activation function of reservoir neurons and output neurons respectively, v_f is the system noise; $u(k)$, $X(k)$, $y(k)$ are the input variable , the state variables, output variables at the time k respectively; W^{in}, W, W^{fd} is the weight of the connections from input layer to DR, within DR, from the output to DR reversely, they are generated randomly; W^{out} is the connection weights from DR to output , W^{out} is calculated through training echo state network[8, 10-12].

The hidden layer contains a large number of neurons they are sparsely connected. The sparsity of W and the spectral radius is noted as XD and $\rho(W)$ separately. Typically XD is 1% to 5%. $\rho(W)$ is maximum absolute value of the eigenvalues of the matrix ,generally $\rho(W) < 1$ ensure the system is stabilized [8,12]. Due to DR has unique echo characteristic, ESNs is able to deal with nonlinear problem in the reservoir , therefore ESNs can approximate nonlinear system well and get good prediction effect in certain nonlinear prediction ; In terms of network training, W^{out} of ESNs is the global optimal value avoiding the local minimum problem of neural network ; The core problem is to calculate the W^{out} for Training ESNs, the network training process is simpler than the traditional neural network.

According to the nature of the echo state network and mobile telephone traffic data from xinjiang branch of china mobile communication corporation, following chapters will describe the method for establishing multivariate the echo state network prediction model.

3. Data Collection and Processing

3.1 Data Collection

Test data from xinjiang branch of China mobile is collected, the interval is one hour. Data attributes include: the cell name, recording time, traffic, Visitor Location Register (VLR) user number, system response times, the system connection ratio, the pager number *etc.* At this stage our research core contents are prediction traffic and traffic trends. The traffic is a typical time series, as show in Figure 1. It is traffic data of some cell xinjiang province from December 2, 2012 to December 8, 2012. From Figure 1, the change period of traffic is one day. The traffic peaks at 21 or 22 o'clock of Beijing time every day.

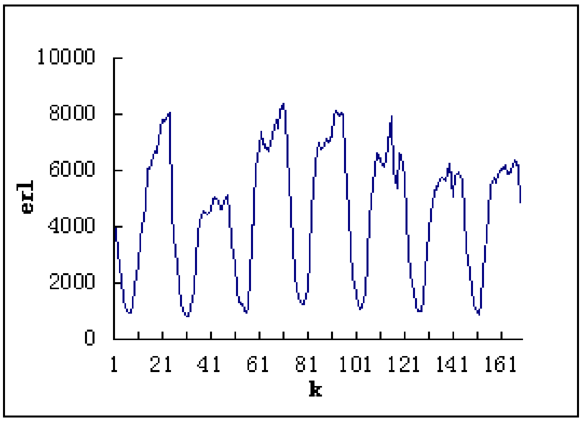

Figure 1. Traffic Graphs of One Week

Our experimental data taken from a region xinjiang on January 13, 2013 to on May 3, 2013. 111 days traffic data sum to 111 x 24 = 2644, the time of the maximum traffic daily is defined busy-time, busy-time traffic data is counted in 111 days, we can get 111 busy-time traffic data, as shown in Figure 2 we can see from Figure 2 that the busy-time traffic data show rising trend.

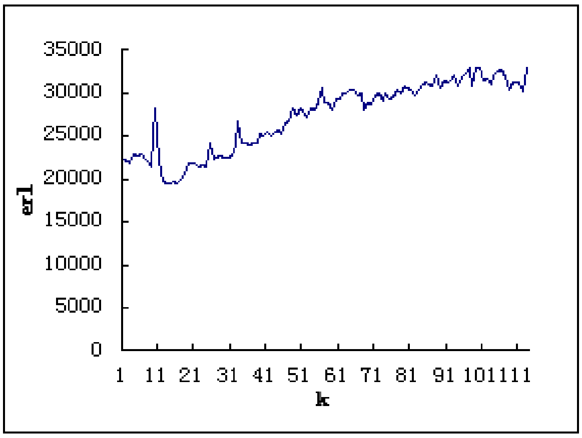

Figure 2. Busy-Time Traffic Graphs

3.2 Traffic Data Correlation Analysis

Various factors influence the traffic change, the independent variables using principal component analysis (PCA) is meaningless, so this paper takes spearman non-parametric correlation analysis method to lay the theoretical foundation for the following PCA method. Correlation coefficient represents a linear relationship between variables, the greater the absolute value of correlation coefficient is, the stronger the correlation between them is.

Collected data: Visitor Location Register (VLR) user number, system response times, the system connection ratio, the pager number and traffic. The peak traffic of everyday 24 hours is defined as busy-time traffic. The data when traffic peak is defined separately as busy-time Visitor Location Register (VLR)user number($h_1(n)$), busy-time system response times($h_2(n)$), busy-time system connection ratio ($h_3(n)$), busy-time pager number($h_4(n)$) and busy-time traffic($X^{max}(n)$).

The busy-time traffic ($X^{max}(n)$) and the non-parametric Spearman correlation coefficient (θ) of impact factor h1(n), h2 (n), h3 (n), h4 (n) , n∈ [1, 111] is calculated as follows:

$$\theta = \frac{\sum (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum (R_i - \bar{R})^2 (S_i - \bar{S})^2}} \tag{3}$$

where R_i is the rank of the i-th X^{max}, S_i is the rank of i-th $h_j(n)$,j∈ [1,4]). $\bar{R}$ and $\bar{S}$ is the average of R_i and S_i respectively.

Do hypothesis test, assuming *p* is the probability that the final result is true, namely the probability of $\theta = 0$ is *P*, *P* is expressed as:

$$P = \frac{\sqrt{n-2} \times \theta}{\sqrt{1-\theta^2}} \tag{4}$$

Where n-2 is degrees of freedom, the value of *n* is 111.The results is shown in Table 1.

Table 1. Correlation Coefficients and Probability of Hypothesis Test

	$h_1(n)$	$h_2(n)$	$h_3(n)$	$h_4(n)$
θ	0.959	0.898	0.735	0.812
P	0.000	0.000	0.000	0.000

Table 1 shows that: the correlation coefficients (θ) of busy-time traffic (Xmax (n)) and h1 (n), h2 (n), h3 (n), h4 (n) are greater than 0.5, indicating that the four impact factors and Xmax (n) are significantly correlated; *P* in the second row of Table 1 are all 0.000, manifesting that the correlation of these four impact factors and busy-time traffic is true in the limited sampling set.

Form what has been discussed, the five variables h1(n), h2 (n), h3 (n), h4 (n) and Xmax(n) together are employed as input variables of traffic forecasting model in this paper.

3.3 Principal Component Analysis (PCA) Dimension Reduction

Although multiple input variables can more fully describe the traffic information, but it increases the input dimension and the complexity of the forecasting model. In this paper, we use principal component analysis (PCA) to reduce the dimension of input variable. It not only can keep effectively multivariate sequence information, but also can simplify the training process of the traffic prediction model, further improve the calculation efficiency and accuracy. PCA is used to reduce dimension of input variable, the steps is as follows:

Step1: The input variables are standardized to obtain the normalized input variables X =[X1, X2, ... Xn] ^{T}, $X \in R^{n \times m}$.

Step2: the introduction of PCA, applying singular value decomposition to the X by the formula (5) , theoretical basis see paper [13] - [15].

$$X^{\wedge} = U \sum V^T \tag{5}$$

Where U, V^T are orthogonal matrices, $\sum=[Q_1,Q_2,...Q_R,0,...O]$, $Q_1 \geq Q_2 \geq ... \geq Q_R > 0$ are R singular values.

Step 3: the variance contribution ratio of the singular value is calculated according to the formula (6). η_0 is a constant in the formula (6), usually $0.8 \leq \eta_0 \leq 1$, take the first *p* principal components $V^{\sim}$,$V^{\sim}=[V_1,V_2,...V_P]$ at the same time.

$$\begin{cases} \sum_{j=1}^{p} \eta_j > \eta_0 \\ \eta_j = \dfrac{Q_j^2}{\sum_{k=1}^{R} Q_k^2} \quad \text{j=1,2,...,R} \end{cases} \tag{6}$$

Step4: get the primary matrix for Z=X^V~.

In accordance with the above steps of PCA, the collected five busy-time variables data: h1(n), h2(n), h3(n), h4 (n) and Xmax(n) are applied PCA to get the singular value variance contribution ratio(VCR) and cumulative variance contribution ratio (CVCR) of the principal component as shown in Table 2.

Table 2. Variance Contribution Ratio and Cumulative Variance Contribution Ratio

principal component	Z1	Z2	Z3	Z4	Z5
VCR %	70.221	22.017	5.689	1.721	0.352
CVCR %	70.221	92.238	97.927	99.648	100.000

From Table 2 shows that the cumulative variance contribution rate of the first 2 principal component is 92.238%, it has been able to represent the most information in the sample data. The data of the first principal component Z1 and the second principal component Z1 are obtained by the formula (7) and (8) :

$$Z1=0.931X_1+0.938X_2+0.852X_3-0.402X_4+0.937X_5 \tag{7}$$

$$Z2=0.235X_1+0.186X_2-0.371X_3+0.886X_4+0.299X_5 \tag{8}$$

4. ESNs Busy-Time Traffic Prediction Model

4.1 ESNs the Structure of Busy-Time Traffic Forecast Model

Five input variables are respectively h1 (n) , h2 (n), h3 (n), h4 (n) and Xmax (n) and one output variable y (n + L)(L as the prediction step length) in ESNs busy-time traffic forecast model, y (n + L) is busy-time traffic prediction value. The principal component analysis is used to reduce dimensions after Standardized input variables. The first principal component and the second principal component are obtained after applying PCA. The two principal component are used as input variables of the echo state network. The two principal components after principal component dimension reduction were nonlinear and independent each other. As unique features of DR, ESNs can better deal with nonlinear problem of traffic. So the busy-time traffic forecast model structure diagram is established as shown in Figure 3.

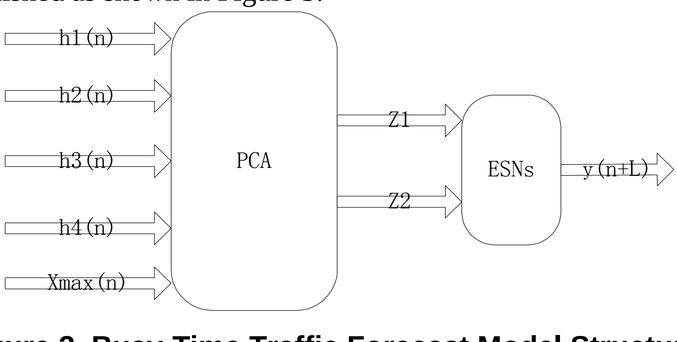

Figure 3. Busy-Time Traffic Forecast Model Structure

4.2 Training ESNs Traffic Prediction Model

According to traffic data, setting parameters refer to papers [16-17]: the DR size N = 200, sparse XD = 3%, $\rho(W)$ = 0.85, W^{in}, W, W^{fd} randomly generate. f^{DR} is hyperbolic tangent Sigmoid function, and f^{RD} is the identity function, system noise v_f = 0, the model state variables x(0) =0 .

Taking first 105 busy-time data as training data, Z(n) instead of u (n) is putted into formula (1) and get value of the state variable x (n). Every x(n) constitute a state variable matrix B. The output state variable y (n) is calculated by the formula (9). $\hat{y}(n)$ is noted as expected output state variable.

$$y(n)=W^{out}x(n) \tag{9}$$

Every y(n) forms output state matrix Y, every $\hat{y}(n)$ constitutes expected output state matrix $\hat{Y}$, the training goal is the minimum deviation between output and the expected output, as shown in the formula (10):

$$\min\left\|\hat{Y}-Y\right\|^2 \tag{10}$$

Violations method is used to solve the values of output weight W^{out}, the expression is showed as the equation (11):

$$W^{out}=Y*B^{\dagger} \tag{11}$$

where $B^{\dagger}$ is violations matrix of B . Now the training ESNs is completed.

5. The Simulation Results

The five-variable sequences $h_1(n)$,$h_2(n)$, $h_3(n)$, $h_4(n)$and $X^{max}(n)$ are put into trained ESNs(N, XD, $\rho(W)$, W^{in}, W, W^{fd}, W^{out}) traffic prediction model .Then running the network model can get predict putout y(n), the closer the value between y(n) and $\hat{y}(n)$ are, the higher prediction accuracy is . The error evaluation index is the standard mean square root error(E_N) in this paper, expressed as formula (12) [18]:

$$E_N=\sqrt{\frac{\sum_{n=a+1}^{b}(\hat{y}(n)-y(n))^2}{(b-a)*\sigma^2}} \tag{12}$$

Where *a* is the beginning time of the test sample , *b* is the end point of the test sample, σ^2 is the variance of busy-time traffic sequence, the smaller the value of E_N is, the higher precision of prediction error is. Selected the data come from 50 cells A region in xinjiang province on January 13, 2013 to May 4, 2013.The 7 days busy-time traffic from April 28, 2013 to May 4, 2013 are predicted. Training error and the prediction error are shown in Figure 4 and Figure 5 respectively.

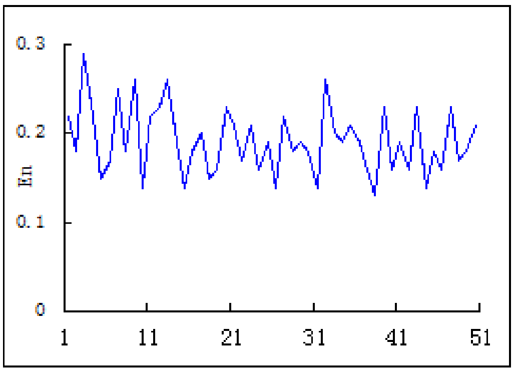

Figure 4. Trained Error

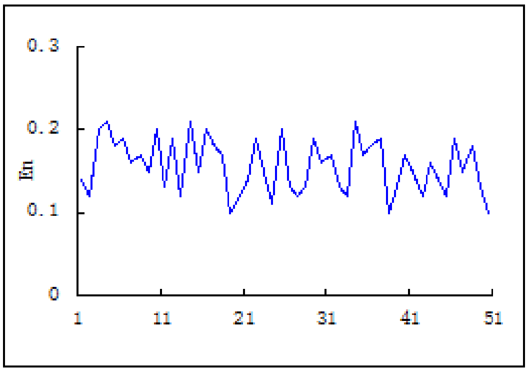

Figure 5. Prediction Error

Average training error is 0.19, the average prediction error is 0.15.The different traffic cells are selected for the experiment, the average prediction error is able to meet the actual demand. So it is proved that generalization of the proposed forecasting model is good.

To compare the forecasting performance of prediction methods, as shown in Table 3 prediction methods include: proposed method (ESNs of Multivariate Principal Components Analysis)in this paper, least-square support vector machines (LS-SVM) forecasting model based single variable (method and parameter Settings see the paper [1]) and the ESNs prediction model based single variable (refer to the paper [8]). The computation time is the time of completing once prediction after parameters are set.

Table 3. Comparison of Prediction Performance

Prediction model	this paper	LS-SVM	ESNs
E_N	0.18	0.51	0.32
Computation time(s)	0.96	2.59	1.17

Table 3 shows: prediction accuracy of this paper is superior to LS-SVM and ESNs forecast model based single variable; on the computation time, prediction efficiency of this paper proposed prediction method is 0.96s. It is significantly higher than 2.59 seconds of LS-SVM forecasting model; As processing multivariate data occupied more time, compared to the single variate ESNs predict model, this paper computation time is less 0.21 seconds, there is no obvious superiority. In summary, the new prediction method used in this paper both on the prediction accuracy and prediction efficiency, is the superior than LS-SVM and ESNs based single variable.

6. Conclusion

Four variables of significantly relevant to busy-time traffic are obtained by means of correlation analysis in this paper, these variables and busy-time traffic constitute the multivariate input sequences. Combining with PCA set up a busy-time traffic prediction model based ESNs. The simulation shows that the proposed new model not only has excellent generalization, but also highly prediction accuracy and prediction efficiency. The proposed new prediction method can lay theory and practice foundation for fine planning of mobile communication network.

The next step work is to study the stability of ESNs traffic prediction model from the perspective of the theoretical analysis.

Acknowledgement

This study is supported by natural science fund project young teachers in colleges and universities department of education in Xinjiang Province, the serial number of the project: XJEDU2014S074.

References

[1] W. Shaojun, L. Qi, P. Xiyuan, L. Datong and C. Qiang, "Study on LS-SVM method for multi-step forecasting of mobile communication traffic", Chinese Journal of Scientific Instrument, vol. 32, no. 6, **(2011)**, pp. 1258-1264.
[2] Y. Peng, M. Lei and J. Guo, "Clustered complex echo state networks for traffic forecasting with prior knowledge", Proceedings of 2011 IEEE International Instrumentation and Measurement Technology Conference, Hangzhou, China, **(2011)**.
[3] X. Wang, C. Zhang and S. Zhang, "Modified Elman neural network and its application to network traffic reduction", Proceedings of2012 IEEE 2nd International Conference, Hangzhou, China, **(2012)**.
[4] H. Jaeger, "The echo state approach to analysis and training recurrent neural networks", Tech. Rep., German national research center for information technology, Berlin German, **(2001)**.
[5] H. Jaeger and H. Haas, "Harnessing non-linearity: predicting chaotic systems and saving energy in wireless communication", Science, vol. 304, no. 5667, **(2004)**, pp. 78-80.
[6] P. A. Kountouriotis, D. Obradovic, S. L. Goh and D. P. Mandic, "Multi-step forecasting using echo state networks", Proceedings of IEEE the International Conference on Computer as a Tool, Belgrade, Serbia, **(2005)**.
[7] L. Miao, P. Yu, G. Jia and P.Xiyuan, "Traffic forecasting for prior knowledge based clustered complex echo state networks", Chinese Journal of Scientific Instrument, vol. 32, no. 10, **(2011)**, pp. 2190-2197.
[8] P. Yu, W. J. Min and P. X. Yuan, "Researches on Time Series Prediction with Echo State Networks", Acta Electronica Sinica, vol. 38, no. 2A, **(2010)**, pp. 148-154.
[9] Q. Ge and C. Wei. "Multi-resolution based Echo State Network and its Application to the long-term Prediction of Network Traffic", IEEE proceeding of International Symposium on, Wuhan, China, **(2008)**.
[10] H. Min and W. X. Ying, "An Effective Online Sparse Learning Algorithm for Echo State Networks", Acta Automatica Sinica, vol. 37, no. 2, **(2012)**, pp.1536-1540.
[11] P. Yu, W. J. Min and P. X. Yuan, "Survey on Reservoir Computing", Acta Electronica Sinica, vol. 39, no. 10, **(2011)**, pp. 2397-2396.
[12] H. Cui, X. Liu and L. Li, "The architecture of dynamic reservoir in the echo state network", Chaos: An Interdisciplinary Journal of Nonlinear Science 22, vol. 033127, **(2012)**.
[13] W. Minghu, C. Rui, L. Ran and Z. Shangli, "Dynamic Global-Principal Component Analysis Sparse Representation for Distributed Compressive Video Sampling", China Communications, vol. 10, no. 5, **(2013)**, pp. 20-29.
[14] X. Ding, L. He and L. Carin, "Bayesian Robust Principal Component Analysis", IEEE Transactions on Image Processing, vol. 20, no. 12, **(2011)**, pp. 3419-3430.
[15] M. Jolliffee, "Principal Component Analysis", Springer-Verlag, New York, USA, **(1986)**.
[16] D. Li, M. Han and J. Wang, "Chaotic Time Series Prediction Based on a Novel Robust Echo State Network", IEEE Transactions on Neural Networks and Learning Systems, vol. 23, no. 5, **(2012)**, pp. 787-799.
[17] A. Rodan and P. Tiňo, "Minimum Complexity Echo State Network", IEEE transaction on neural networks, vol. 22, no. 1, **(2011)**, pp. 131-144.
[18] M. C. Neto, Y. S. Jeong, M. K. Jeong and L. D. Han, "Online-SVR for Short-term Traffic Flow Prediction under Typical and Atypical Traffic Conditions", Expert systems with Applications, vol. 36, no. 3, **(2009)**, pp. 6164-6173.

Authors

Liu Jun Xia, She was born in 1980, is a lecturer of Xinjiang Institute of Engineering of china. She received B.S. degree from Xinjiang Normal University of China. M.S. degree was obtained from Xinjiang University of China. She is a Ph. D. candidate of Xinjiang University now. Her research interests are mobile communication network planning and modeling.

Jia Zhen-Hong, He was born in 1964, is doctoral supervisor and professor of Xinjiang University of China. He received B.S. degree from Beijing Normal University of China. M.S. degree and doctor's degree were obtained from Shanghai Jiao Tong University of China. At present, he is a member of the Optical Society of China holography and optical information processing Committee, Committee members of National optoelectronic technology and systems, doctoral supervisor of Xi'an Jiao Tong University and Xinjiang University of China. His research interests are optical communication technology, signal and information processing *etc*.

