International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

Traffic Congestion Prediction using Soft computing Technique

Huma Parveen
Computer Network
BBD University, Lucknow, India
humapiu@gmail.com

Rishi Srivastava
Computer Network
BBD University, Lucknow, India
rishi.bbdu@gmail.com

Abstract--Network traffic forecasting has many important role to play in the domain of network traffic congestion control, its management and network traffic engineering. Characterizing the traffic and modeling are necessary for efficient functioning of the network. It is very vital for any model to depict self similarity, heavy tailed distribution and long range dependence (LRD). Thus modeling of time series is a challenging task. In the present work video stream prediction for application in services like video-on-demand, videoconferencing, video broadcasting, etc has been proposed. The main objective being is to forecast the variable bit rate (VBR) data stream for the allocation of efficient bandwidth of video signal. This plays an important role in traffic congestion control prediction.

The model is fitted on a real data, consisting of training and test sets, taken from the video stream files of Telecommunication Networks Group, Technical University of Berlin, Germany. Here, an artificial intelligence model, known as Adaptive Neuro Fuzzy Inference System (ANFIS) has been proposed. The actual traffic data and the predicted traffic data is compared for performance evaluation of the model. Based on the prediction error the performance metrics are evaluated. Results confirm the simplicity and the better performance of ANFIS model. The work shows that ANFIS is able to forecast traffic congestion control from the point of view of bandwidth allocation.

Key Words-- Network Traffic, ANFIS, LRD, VBR.

1. Introduction

Traffic modeling and analysis plays a vital role in determining network performance. A model which can accurately interpret the important characteristics of traffic is required for efficient critical study and simulation purposes. This leads to better network dynamics know how, being a vital help for network plan and control of bandwidth wastage. Traffic modeling originated from the study of telephone network with the Poisson assumption of the traffic arrival process. However, with the emergence of modern technology, network traffic has evolved tremendously becoming more complex and bursty than the earlier voice traffic. This has led to the introduction of many advanced stochastic models. However in order to accurately predict a network traffic, there is a need for traffic models which can copy the characteristics. Network engineering and management rely a lot on an appropriate model for traffic measurements. Traffic forecasting is foremost research interest for many network engineers.

Traffic modeling and forecasting also plays a significant role in achieving an optimum resource allocation by appropriate bandwidth provisioning and simultaneously maintaining the maximum network utilization. Thus it is worth having a good model to design future network capacity. Time series models, linear as well as non linear were used extensively to model network traffic by many researchers. Some of the popular time series models used are Auto Regressive Moving Average (ARMA), Auto Regressive Integrated Moving Average (ARIMA), Auto Regressive Conditional Heteroskedasticity (ARCH), etc.

2. Related Work

Network traffic prediction methods have been addressed by a number of researchers. Quite a few famous prediction models have challenged a few vital problems. **Ming Zhang and Yanhong Lu, (2015)** used an adaptive network traffic prediction algorithm based on BP neural network. It used an adaptive learning rate method to adjust the learning rate. **Manish, P. Ganvir, Dr. S.S.Salankar, (2015)** proposed a time-series prediction model for the packet loss rate (PLR). They showed that to forecast of PLR it is very much helpful in congestion control mechanisms. **C. Narendra Babu and B. Eswara Reddy, (2015)** studied about the suitability of different methods for better Internet traffic data. This suitability of hybrid ARIMA-ANN models is studied for both one-step ahead and multi-step ahead prediction cases. **Manish R. Joshi et. Al.(2012)** analysed many such network analysis and traffic forecast techniques. The individuality and rules of earlier studies were investigated. **Samira Chabaa, Abdelouhab Zeroual, Jilali Antari, (2010)** developed an artificial neural network (ANN) model based on the multi-layer perceptron (MLP) for internet traffic data analysis over IP networks. Used ANN to analyze a time series of calculated data for network response assessment. **Andreas Petlund, Pal Halvorsen, Pal Frogner Hansen, Torbjorn Lindgren, Rui Casais, Carsten Griwodz, (2012)** presented a dataset – a real-world, server-side packet trace – from Anarchy Online. Anarchy Online is a science fiction-themed massively multiplayer online role playing game (MMORPG). **Anukool Lakhina, et.al., (2005)** showed the usefulness of analysing anomalies as events that change traffic characteristic distributions. They showed that analysing anomalies in this way leads to considerable diagnostic authority, in discovering new anomalies, in

International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

comprehending the structure of anomalies, and in classifying anomalies. **Sun Guang, (2013)** proposed Network traffic forecasting based on the wavelet analysis and Hopfield neural network, which is the basic of the research on network traffic prediction model. **Han Song, Luying Gan (2015),** analysed that if one does do not impede with the network traffic, it will create network congestion and paralysis. **Tao Peng and Zhoujin Tang,(2015)** proposed a nonlinear forecasting algorithm based on a related local LSSVM regression model to forecast smaller network traffic.

3. Dataset Used

The data used for model development has been taken from the video stream files of Telecommunication Networks Group, Technical University of Berlin, Germany [6]. We used trace file from MPEG-4 Jurassic Park I movie in high quality. Both the training and the test set consist of 3000 patterns (first 2000 representing the training, next 1000 the test set). The descriptions of the features are taken from http://trace.eas.asu.edu/TRACE/trace.html [6].

4. ANFIS Model Develoment

Model Selection

In the present work ANFIS Network Structure model consisting of one input layer with five input variables and an output layer consisting of variable bit rate of video traffic as the output variable. Here predictions were based on taking *N* previous patterns to predict one following pattern. This is shown in Fig. 1 below.

In the present work various ahead prediction models have been tested using ANFIS. This includes 1- step ahead, 3-step ahead, 5-step ahead and 10-step ahead prediction models.

All the models include the time series predictions analysis, with the following modeling variables,

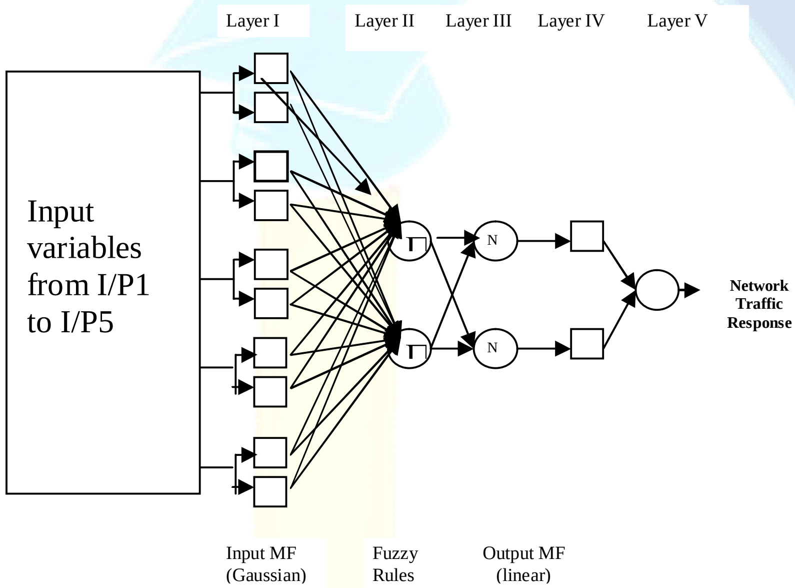

Fig. 1. ANFIS architecture

	Input Variables (Time series VBR traces)					Output Variable
Model-1	VBR(t-4)	VBR(t-3)	VBR(t-2)	VBR(t-1)	VBR(t)	VBR(t+1)
Model-3	VBR(t-4)	VBR(t-3)	VBR(t-2)	VBR(t-1)	VBR(t)	VBR(t+3)
Model-5	VBR(t-4)	VBR(t-3)	VBR(t-2)	VBR(t-1)	VBR(t)	VBR(t+5)
Model-10	VBR(t-4)	VBR(t-3)	VBR(t-2)	VBR(t-1)	VBR(t)	VBR(t+10)

Parameter Selection

As discussed earlier, ANFIS is a well judged combination of FIS and ANN, having learning, thinking and reasoning capability [4] and it uses the benefits of these two techniques into a single capsule [3].

Identification of the rule base is the key of a FIS. The problems are (1) there are no standard methods for transformimg human knowledge or experience into rule base; and (2) it is required to further tune the MFs to minimise the output error and to maximise the performances. Hence for generation of FIS using ANFIS, it is vital to choose proper parameters, which includes the number of membership functions (MFs) for each individual antecedent variables. It is also significant to choose proper parameters for learning and refining process, including the initial step size (ss). Here frequently used rule extraction method i.e. subtractive clustering has been used for FIS recognition and refinement. The ANFIS is generated using the MATLAB version R2013a Fuzzy Logic Toolbox [9].

In ANFIS, the primary parameters of the ANFIS are recognized using the subtractive clustering method. However, subtractive clustering algorithm parameters still need to be specified. In subtractive clustering algorithm clustering radius palys a very vital role and is found out

International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

through a trial and error procedure. By varying the clustering radius r_a between 0.1 and 1 with a step size of 0.01, the optimal parameters are sought by minimizing the root mean squared error obtained on a representative validation set. Clustering radius r_b is selected as 1.5 r_a. Default values are used for other parameters in the subtractive clustering algorithm [7].

Gaussian membership functions are used for each fuzzy set in the fuzzy system. Subtractive clustering algorithm is able to determine number of membership functions and fuzzy rules for the development of ANFIS model. Hybrid learning algorithm is used for finding the Gaussian membership function parameters. Each ANFIS is trained for 1000 epochs .

The input and output membership functions are Gaussian and linear. For input argument different data sets has been used. For one output, *genfis2* has been used for initial FIS generation and training of ANFIS. The membership function type and their numbers are given in table 2. The input membership function curves for the model based on performance criteria for ANFIS are shown in figure 2. The rule extraction method used for training ANFIS are specified in table 3. Table 4 shows the results of model parameters types and values used for training ANFIS.

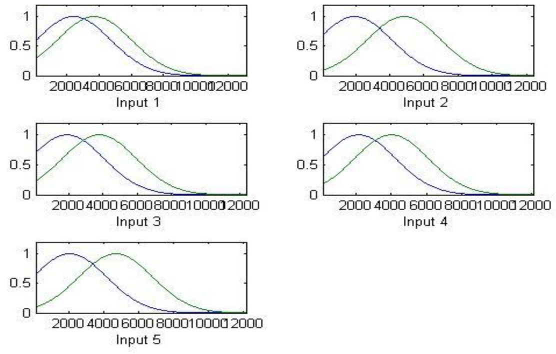

Fig:- 2 Input Membership function curves for the ANFIS model using subtractive clustering rule extraction method for 1-Step Ahead Model

Table 2: Parameters used in all the models for training ANFIS

Rule extraction method	Input MF type	Input partitioning	Output MF Type	Number of output MFs	Training algorithm	Training epoch number	Initial step size
Parameters used	Gaussian membership ('gaussmf')	variable	Linear	one	Hybrid learning	100	0.01

Table 3: Rule extraction method used for training ANFIS 1-Step Ahead Model

Rule Extraction Method	And method	Or method	Defuzzy method	Implication method	Aggregation method
Type	'prod'	'probor'	'wtever'	'prod'	'max'

Table 4: Types and values of parameters used for training ANFIS model for 1-Step Ahead Model

No. of Nodes	44
No. of linear parameters	18
No. of non-linear parameters	30
Total no. Of parameters	48
No. of training data pairs	2000
No. of testing data pairs	1000
No. of fuzzy rules	3

Figure 3 and 4 shows the comparative plots of observed and predicted Video Traffic Network Response data Prediction both for training and testing phases. The figures wisely demonstrate that (1) the model performance are in general accurate, where all data points roughly fall onto the line of agreement; (2) model using subtractive clustering is consistently superior in training phase than in testing phase.

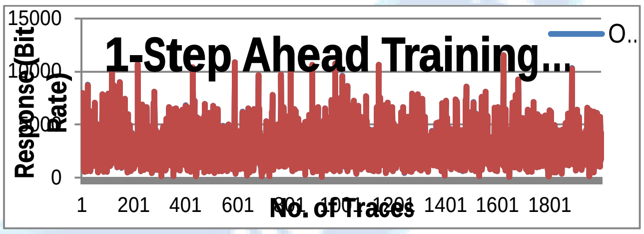

Fig.:- 3 Comparative plot of Actual and Predicted Video Traffic Network Response for Training Datasets

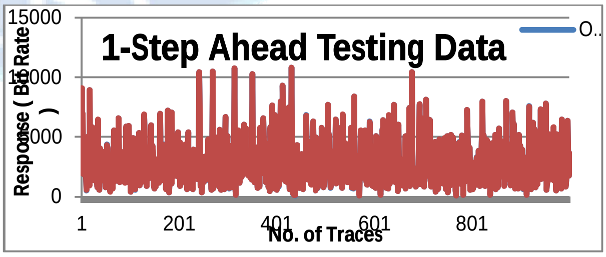

Fig.:- 4 Comparative plot of Actual and Predicted Video Traffic Network Response for Testing Datasets

4. Results and Discussions:
ANFIS model having five input variables are trained and tested by ANFIS method and their performances compared and evaluated based on training and testing data. The best fit model structure is determined according to criteria of performance evaluation. The performances of the ANFIS models for various step ahead predictions are shown in Fig. 5 and 6 and their best MAE and RMSE values based on radius of influence r=0.50, both for training and testing data are given in Table 5 below.

Table 5:- RMSE and MAE Values for Datasets after using ANFIS

Prediction	MAE		RMSE	
	Training	Testing	Training	Testing
One Step Ahead	0.004938	0.00507	9.79	9.93
Three Step Ahead	0.0053	0.005537	12.6	12.44
Five Step Ahead	0.006282	0.006114	13.81	13.66
Ten Step Ahead	0.004296	0.004276	11.3	11.12

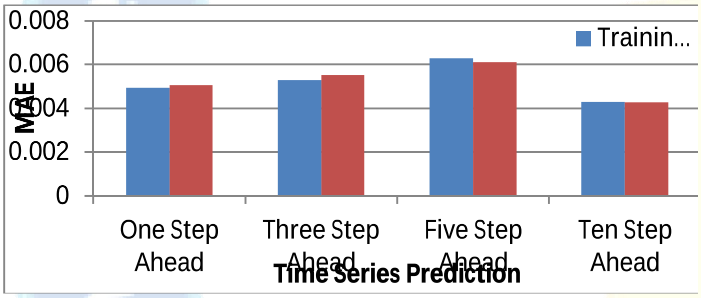

Fig. 5 : MAE Plot for Training and Testing data for various ahead prediction models

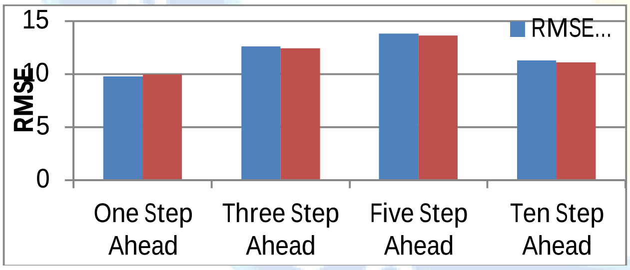

Fig. 6 : RMSE Plot for Training and Testing data

Further in order to evaluate the ability and competence of the model to forecast the network traffic error values MAE has been used. It gives average deviation of the predicted values in relation to measured data and can analyse long term performance of the models.
The absolute error of training and testing data sets for network traffic response output are shown in fig.7 and fig. 8 below.

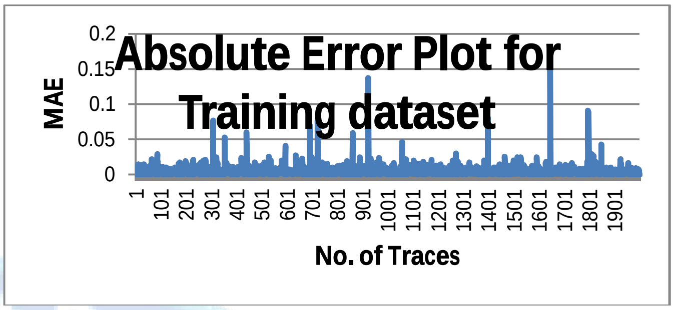

Fig. 7 :- Absolute Error plot for 1-step ahead model for Training Data

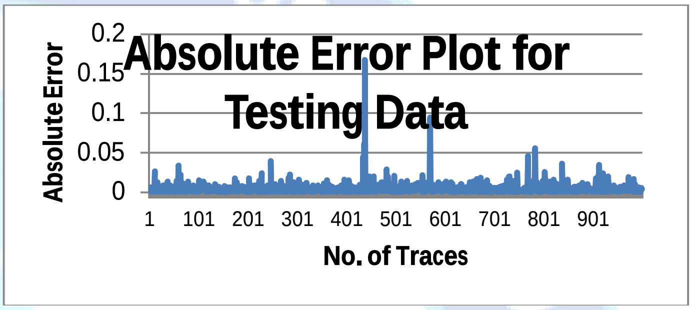

Fig. 8 :- Absolute Error plot for 1-step ahead model for testing data

From the investigation of the above results, it is seen that the network traffic response prediction model developed using ANFIS technique has been able to perform well. The comparative analysis of all the four models show that the ANFIS has been able to best predict the one step ahead prediction, as compared to 3, 5 and 10-step ahead prediction models. But further analysis shows that 10-step ahead model has shown improvement in terms of RMSE and MAE with respect to 3 and 5-step ahead models. Also from the perusal of the charts given in Fig. 5 and 6, it is seen that in all the models the RMSE and MAE values for training and testing datasets are almost same, which again shows that ANFIS model has not over-trained.

5. Conclusion
Here the applicability and capability of ANFIS techniques for time series analysis of network traffic response data prediction has been investigated. The studys has been carried out using MATLAB simulation environment. The data has been taken from the video stream files of Telecommunication Networks Group, Technical University of Berlin, Germany. From the analysis of the above results, given under heading Results and Discussions, it is seen that the network traffic prediction model developed using ANFIS technique has performed well. The analytical analysis of all the four models show that the ANFIS has been able to best predict the one step ahead prediction, as compared to 3, 5 and 10-step ahead prediction models.

References:
[1]. Parvinder S. Sandhu, Sunil Khullar, Satpreet Singh, Simranjit K. Bains, Manpreet Kaur, Gurvinder Singh, "A Study on Early Prediction of Fault Proneness in Software

Modules using Genetic Algorithm", World Academy of Science, Engineering and Technology, 2010, pp. 648-653.
[2]. Ahmet Okutan, et. al., (2012), "Software defect prediction using Bayesian networks", Empir Software Eng (2014) 19:154–181
[3]. Mrinal Singh Rawat, et. al.,(2012), "Software Defect Prediction Models for Quality Improvement: A Literature Study", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 2,
[4]. CHIU, S., (1994), "Fuzzy Model Identification based on cluster estimation", Journal of Intelligent and Fuzzy Systems, 2 (3), pp 267–278.
[5]. Jang, J-S. R., (1993), "ANFIS-Adaptive-Network Based Fuzzy Inference System", IEEE Transactions on Systems, Man and Cybernatics, 23(3), pp 665-685.
[6]. Ming Zhang and Yanhong Lu, (2015), " Adaptive Network Traffic Prediction Algorithm based on BP Neural Network", International Journal of Future Generation Communication and Networking Vol. 8, No. 5 (2015), pp. 195-206.
[7]. Manish, P. Ganvir, Dr. S.S.Salankar, (2015), "Time Series Forecasting of Packet Loss Rate Using Artificial Neural Network Based on Particle Swarm Optimization", International Journal of Engineering Research and General Science Volume 3, Issue 2, pp. 466-472.
[8]. C. Narendra Babu and B. Eswara Reddy, (2015), "Performance Comparison of Four New ARIMA-ANN Prediction Models on Internet Traffic Data", Journal of Telecommunication and Information Technology, pp. 67-75.
[9]. Manish R. Joshi et. Al.,(2012) "A Review of Network Traffic Analysis and Prediction Techniques", pp. 1-22.
[10]. Samira Chabaa, Abdelouhab Zeroual, Jilali Antari, (2010), " Identification and Prediction of Internet Traffic Using Artificial Neural Networks", J. Intelligent Learning Systems & Applications, 2010, 2, 147-155.
[11]. Andreas Petlund, Pal Halvorsen, Pal Frogner Hansen, Torbjorn Lindgren, Rui Casais, Carsten Griwodz, (2012), "Network Traffic from Anarchy Online: Analysis, Statistics and Applications", ACM 978-1-4503-1131-1/12/02, pp. 1-6.
[12]. Anukool Lakhina, et.al., (2005), "Mining Anomalies Using Traffic Feature Distributions", SIGCOMM'05, August 21–26, 2005, Philadelphia, Pennsylvania, USA., pp.1-12.
[13]. Sun Guang, (2013), "Network Traffic Prediction Based on the Wavelet Analysis and Hopfield Neural Network", International Journal of Future Computer and Communication, Vol. 2, No. 2, April 2013.
[14]. Han Song, Luying Gan (2015), "The Research on the Prediction of the Network Traffic Based on the Improved IAC-Gray Method", CHEMICAL ENGINEERING TRANSACTIONS, VOL. 46, pp. 1297-132.
[15]. Tao Peng and Zhoujin Tang,(2015), "A Small Scale Forecasting Algorithm for Network Traffic based on Relevant Local Least Squares Support Vector Machine Regression Model", Appl. Math. Inf. Sci. 9, No. 2L, 653-659.