IP Traffic Forecasting Using Focused Time Delay Feed Forward Neural Network

Shahzad Ahmed

Abstract--There have been a number of methods presented by various researchers for traffic prediction, some of which involve modeling the problem of traffic prediction as a time series. It has been observed that Artificial Neural Networks (ANN) perform better than statistical methods for time series forecasting. The network performance and complexity varies with the choice of algorithm used. Back propagation (BPNN) has been used to predict IP traffic with a fair degree of accuracy but as the prediction interval increases and the inputs change drastically the forecasting accuracy suffers [9]. This paper discusses the use of Focused Time Delay Feed Forward Neural Network architecture to predict IP traffic patterns and overcome the short comings of back propagation neural networks when used for traffic forecasting along with improvements to the BPNN by using additional inputs like holidays and maintenance downtimes.

***Index Terms*— ANN, Focused Time Delay Feed Forward Neural Networks, IP traffic forecasting.**

I. INTRODUCTION

IP network traffic forecasting is an important area of study that is rightfully receiving considerable research attention [1, 2, 3]. It has both small and large scale applications. In production networks, network planning and re-structuring of the network relies heavily upon the forecasted traffic patterns of the network. In has been observed that in many organizations decisions like planed down time and maintenance windows are focused around the time of minimum network usage so that the affect on the network usage and services could be minimized. It may also be used for anomaly detection where the past network trends are analyzed to detect any malicious behavior [4].

Pakistan Revenue Automation Pvt. Ltd, 9th Floor Customs House Karachi Pakistan, Schehzad_ahmed@yahoo.com

This can be done by analyzing the past network traffic trends and predicting the future traffic usage. Manually the IP traffic prediction is performed by the network administrators intuitively, but with the involvement of complex factors such as holidays, momentous events, service types and number of users the accurate prediction of traffic becomes difficult [3]. The Artificial Neural Network (ANN) based traffic prediction techniques can come in handy when it comes to predicting sufficiently accurate IP traffic patterns of the network [2, 3, 8, 9]. These papers use a number of methods ranging from statistical models to neural networks based models employing back propagation each of which has its pros and cons. This paper is an effort to envisage traffic prediction model using the feed forward time delay for its better performance when predicting time series.

II. TRAFFIC FORECASTING

IP traffic prediction is manually done intuitively by network administrators, analyzing the past traffic patterns and growth of number of users [3]. However, advance applications of traffic prediction require some accurate as well as reproducible prediction models. IP traffic prediction techniques are getting more attention as research endeavors to find new and improved methods of predicting the traffic flow. With the advent of *Next Generation Networks* (NGN), integrating voice video and data on IP, greater emphasis has been placed on accurate modeling of predicted IP traffic so that bandwidth guarantees (*QoS)* could be provided for real time data. Traffic prediction is also used to detect misuse, intrusion and anomalies in IP network [1, 2].

A number of IP traffic prediction models have been proposed employing different statistical techniques such as Time Series Forecasting (TSF) as well as Artificial Neural Network based forecasting which predict the behavior based upon historical data. The TSF approaches can be divided into univariate and multivariate, depending if one or more variables are used. Multivariate methods are likely to produce better results, provided that the variables are correlated. It is observed however that ANN based multivariate methods give better results outperforming the statistical methods [2]

As the nature of Internet traffic is non-linear and self-similar the accurate prediction of projected traffic is a difficult task [5]. This leads to self similar prediction models employing ANN. In a comparison of different types of neural networks for IP traffic forecasting, Rutka et al [6] observe that Radial Base Function Networks (RBFN) employing self-similarity are computationally expensive, missing performance goals due to complex calculation and long calculation periods. Rutka et. al in [6] suggest that Multilayer Perceptron Networks (MLPN) could be a better choice and their performance could be improved by increasing the number of layers while there is no need to increase the number of neurons in each layer. Taeho et. al in [7] propose that improvements to the ANN based prediction methods involve the Virtual Term Generation (VTG) which improve the generalization performance of the ANN through the generation of derived training patterns from the original ones and training the ANN using these *artificial training patterns* in addition to the original patterns This again is a computationally expensive method and is not suitable for real time data. Zhao et. al in [8] discuss the use of Wavelet Neural Networks with decomposition of the traffic sequences into different frequency components and training the ANN using these sub components. This effectively improves the prediction accuracy through the replacement of sigmoid function with wavelet basis function. These methods are more accurate but are complex to compute. Satsri et. al in [9] present a rather simpler method that uses ANN with back propagation algorithm which is an effective forecasting tool, because it has a capability of pattern recognition and could relate factors that effect the result, it is reasonable to approximate the unknown function from some known input/output patterns which is very powerful for future trend forecasting.

The technique presented by Satsri et. al in [9] is effective when compared with statistical models while forecasting only for a short time intervals. It does not scale for a longer span of prediction window and does not account for very fast change in the input data due to the usage of back propagation algorithm as the back propagation algorithm relies on the back propagation of error. We anticipate improving the prediction accuracy, the forecasting time window and the response of the prediction model to the changing inputs by employing time delay feed forward neural networks.

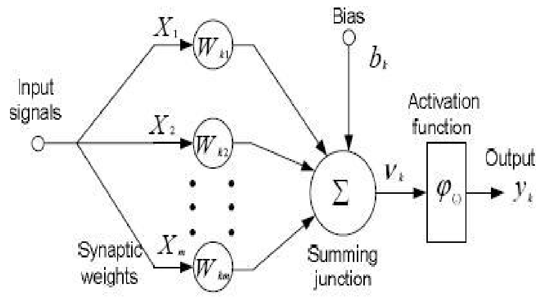


Fig. 1 ANN with Back Propagation Architecture

A. Back Propagation Neural Networks

This study uses the same back propagation model arrangement used by Satsri et. al in [9] for the sake of comparison, comprising of a single layer which involves three stages:

i) The feed forward of input training pattern.
ii) The back propagation of associated error.
iii) Adjustment of weights.

In back propagation arrangement a multilayer perceptron has an input layer of source nodes and an output layer of neurons (i.e., computation nodes); these two layers connect the network to the outside world. In addition to these two layers, the multilayer perceptron usually has one or more layers of hidden neurons, which are so called because these neurons are not directly accessible. The hidden neurons extract important features contained in the input data. The training is usually accomplished by using a *back propagation (BP) algorithm* that involves two phases [10].

During the forward phase the free parameters of the network are fixed, and the input signal is propagated through the network layer by layer. The forward phase finishes with the computation of an error signal.

$$e_i = d_i - y_i \quad (1)$$

where *di* is the desired response and *yi* is the actual output produced by the network in response to the input **x***i*.

During this second phase, the error signal *ei* is propagated through the network in the backward direction, hence the name of the algorithm. It is during this phase that adjustments are applied to the free parameters of the network so as to minimize the error *ei* in a statistical sense. The back-propagation learning algorithm is simple to implement and computationally efficient in that its complexity is linear in the synaptic weights of the network. However, a major limitation of the algorithm is that it does not always converge and can be excruciatingly slow, particularly when we have to deal with a

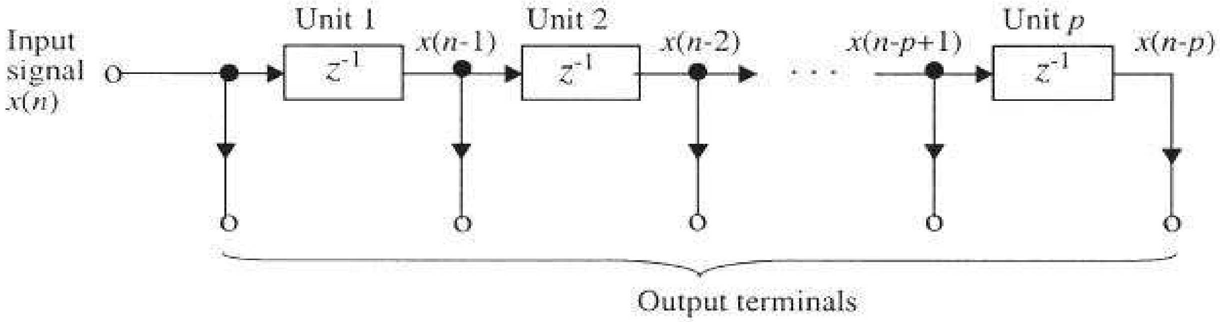


Fig. 2 Ordinary tapped-delay line memory of order p.

difficult learning task that requires the use of a large network. [11]. A complete iteration of this algorithm can be written

$$x_{k+1} = x_k - \alpha_k g_k \qquad (2)$$

where x_k is a vector of current weights and biases, g_k is the current gradient and α_k is the learning rate[12].

B. Focused Time Lagged Feed Forward Neural Networks

The back propagation algorithm for ANN is concerned basically with an approximation of the systems being static without any dynamics. *Time* is an essential dimension of learning. We may incorporate time into the design of a neural network implicitly or explicitly. A straightforward method of implicit representation of time1 is to add a *short-term memory structure* in the input layer of a static neural network (e.g., multilayer perceptron). The resulting configuration is sometimes called a *focused time lagged feed forward network (TLFN)* [11].

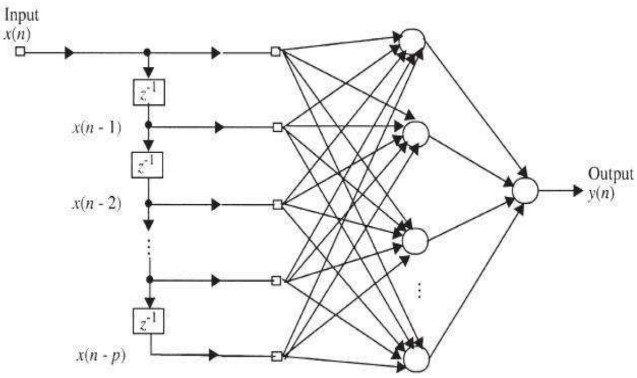


Fig. 3 Focused time-lagged feedforward network (TLFN)

Tapped-Delay-Line (TDL) Memory is the most commonly used form of short-term memory. It consists of *p* unit delays with (*p* + 1) terminals as shown in the figure 2, which may be viewed as a single input–multiple output network. Figure 3 shows a focused TLFN network using the combination of a TDL memory and multilayer perceptron. In Figs. 2 and 3, the unit-delay is denoted by *z*-1. The *memory depth* of a TDL memory is fixed at *p*, and its *memory resolution* is fixed at unity, giving a *depth resolution constant* of *p*.

Focused TLFNs using ordinary tapped-delay memory or gamma memory are limited to stationary environments. To deal with non stationary dynamical processes, we may use distributed TLFNs where the effect of time is distributed at the synaptic level throughout the network. The training of a distributed TLFN is naturally a more difficult proposition than the training of a focused TLFN. Whereas we may use the ordinary back-propagation algorithm to train a focused TLFN, we have to extend the back-propagation algorithm to cope with the replacement of a synaptic weight in the ordinary MLP by a synaptic weight vector. This extension is referred to as the temporal back-propagation algorithm [13].

III. EXPERIMENTS

A. Theory & Methodology

The traffic prediction methodology first involves the gathering of traffic patterns. The data was gathered from one of the core routers of a production network involving multiple high bandwidth links for a web hosting service. The links involved active-standby fiber optic links for high redundancy. The traffic pattern data was gathered using SNMP interface statistics polling [14]. The SNMP statistics are then propagated to the network monitoring server. We used cacti open source network monitoring

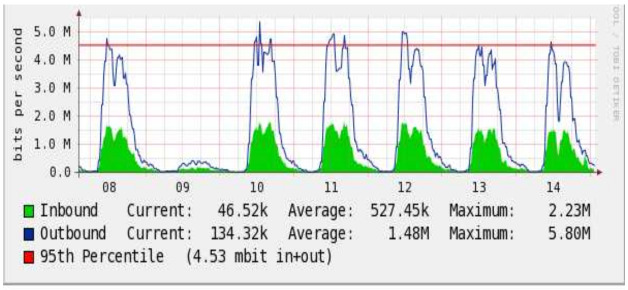


Fig. 4 Weekly Monitored Traffic Pattern 30mins Average

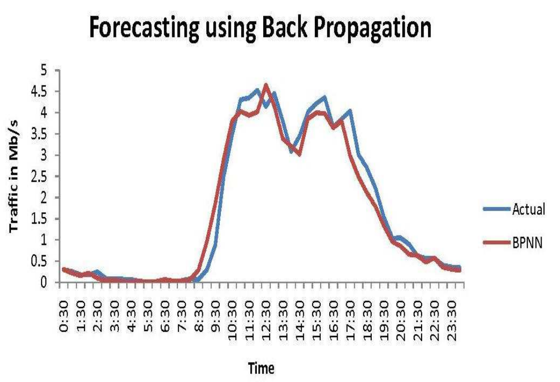
solution [15] running rrdtool to graph the traffic [14, 16]. The rrdtool supports exporting the data into convenient csv format which was used to extract the input data for our test setup. The traffic patterns were used as an input time series to the neural network and using that input time series the future traffic patterns were the output of the neural networks, both back propagation and focused time delay feed forward networks.

B. Data collection and Preprocessing

The first step was to gather the IP traffic data of 6 months. The gathered data was sampled at an interval of 2 hours. The data was then divided into training and testing data in ratio of 65% and 15% while 10% of the data was used for validation. Pre-processing techniques were applied and some of the noisy data was removed and interpolated. For test purposes using the 6 month traffic data single day traffic was forecasted. The forecasting was performed using all the three methods i.e. using back propagating, back propagation with additional processing (excluding the holiday and maintenance periods) and focused time delay feed forward neural networks.

C. Forecasting using Feed Forward Back Propagation Network

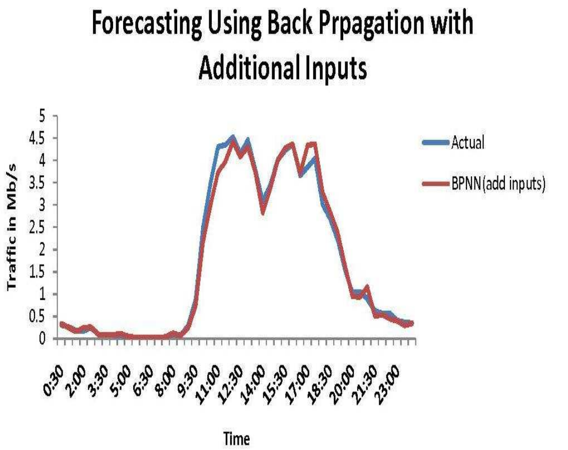
For the sake of comparison first the feed forward back propagation setup of [9] was recreated and testing was performed. The three stage network



consisted of a single input time series as the feed forward input training pattern, the back propagation of the associated error, and adjustment the weights. The input layer consisted of 10 neurons with tan-sigmoid as the transfer function with Levenberg-Marquardt back propagation algorithm (LMBP). The test was conducted with a momentum of 0.01 and default learning rate of 0.7.

D. Forecasting using Feed Forward Back Propagation Network with additional Inputs

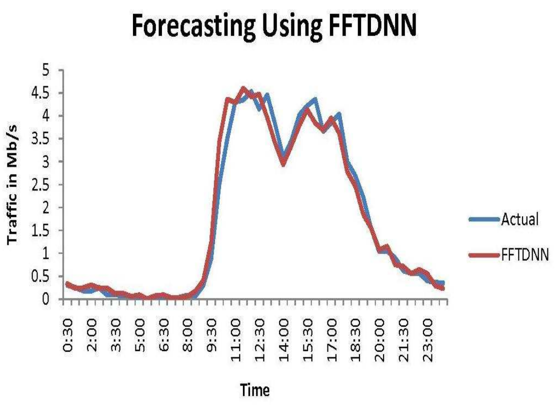
In this setup we used additional inputs to the back propagation network. These inputs involved the factors that may have an impact over the predicted traffic patterns. The inputs involved holiday indicators along with the traffic patterns, network maintenance downtime, bandwidth upgrade and network outage in upstream network beyond our

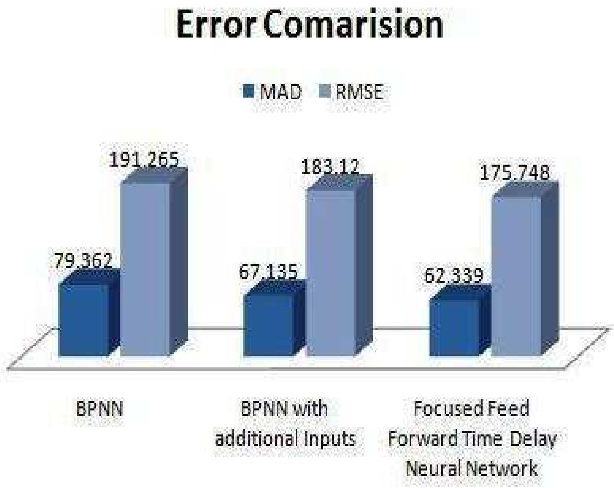


control. In order to study the impact of this upon the overall accuracy of the forecasting the test were also conducted by removing the data associated to these instances Independent tests were conducted using a combination of these factors while to study the actual underlying influencing factor.

E. Forecasting using Focused Time Delay Dynamic Feed Forward Neural Network

We used non-linear focused time delay feed forward neural network for the prediction of time series traffic pattern with a hidden layer of 10 neurons and an input delay of 1 to 5. This was used to evaluate the dynamic feed forward networks for predicting the time series data. The complete test setup involved the steps of normalizing the input data and converting it into time series and cascading it to the dynamic feed forward network. The data used to train the ANN was the same data that had been used in the back propagation setup.





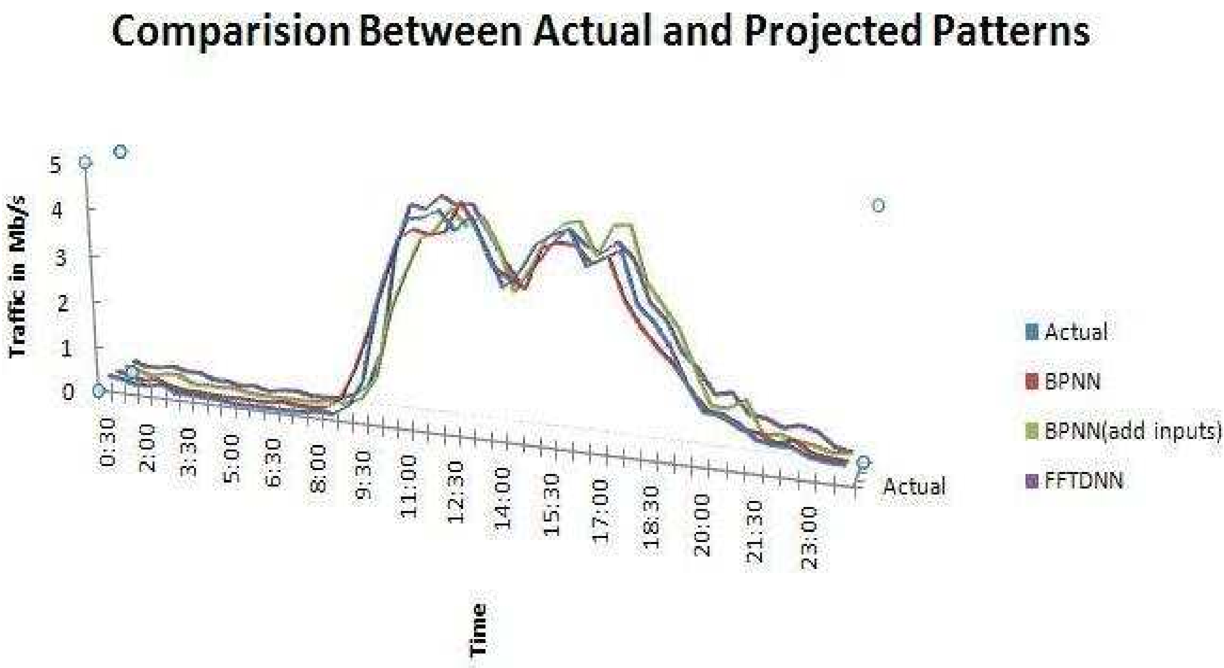
IV. RESULTS AND DISCUSSION

The error comparison of the different model test setups are shown in fig.

The results show that when additional inputs such as holiday data and maintenance outage is also applied to the input of the neural network the performance of prediction of data increases. This is due to the fact that the traffic patterns in real time are sensitive to the network usage which in turn is dependent upon whether the network resource is available in first place or not. The second aspect of this behavior is that network usage is pretty low in the event of holiday sessions.

Another aspect of viewing the same effect is that if we remove the input data which pertains to the time periods of network outages and also holidays or any other event that affects the normal behavioral pattern of network traffic, the prediction efficiency is increased. However this also implies that the predicted traffic only reflects the normal data and will not account for any abnormal patterns that may very likely are encountered in real world scenarios.

The focused time delay feed forward network as it uses no feedback elements, is faster to train and is a better performer when it comes to time series prediction problems like IP traffic prediction. The reason for its better performance may be as it is a dynamic neural network unlike back propagation it can adapt to the changing inputs better than the feed foreword back propagation hence it performs better when it is encountered with changing inputs.



Model	MAD	RMSE
BPN	79.362	191.265
BPN with additional inputs	67.135	183.120
Focused Time Delay Feed-Forward Network	62.339	175.748

V. CONCLUSIONS

This paper introduced various improvements to IP traffic prediction using back propagation neural networks. These involve usage of additional inputs like holiday session, maintenance time-outs and upstream service outage periods. These factors are environment dependent could vary from environment to environment however they have an impact over the traffic prediction accuracy of a neural network. If these parameters are taken out of the prediction equation of a neural network the system's accuracy might increase but the downside of this approach may be that the generalization ability of the neural network is reduced. Time delay dynamic feed forward neural networks look promising while predicting IP traffic patterns unlike back propagation they can adapt to the change in input and forecast better than back propagation network. Their behavior and usage for a long term prediction needs to be investigated in further research.

REFERENCES

[1] H. Tong, C. Li, and J. He. Boosting Feed-Forward Neural Network for Internet Traffic Prediction. In *Proc. Of the IEEE 3rd Int. Conf. on Machine Learning and Cybernetics*, pages 3129.3134, Shanghai, China, August 2004.

[2] Paulo Cortez, Miguel Rio, Pedro Sousa, and Miguel Rocha Topology Aware Internet Traffic Forecasting using Neural Networks. Artificial Neural Networks – ICANN 2007. pp 445-454

[3] K. Papagiannaki, N. Taft, Z. Zhang, and C. Diot. Long-Term Forecasting of Internet Backbone Trafic. *IEEE Trans. on Neural Networks*, 16(5):1110.1124, September 2005.

[4] J. Jiang and S. Papavassiliou. Detecting Network Attacks in the Internet via Statistical Network Traffic Normality Prediction. *Journal of Network and Systems Management*, 12:51.72, 2004.

[5] Bestavros A., Crovella M. E. Self-similarity in world wide web traffic: Evidence and possible causes// IEEE/ACM Trans. Networking, 1997. – Vol. 5. – P. 835–846

[6] G. Rutka. Neural Network Models for Internet Traffic Prediction // Electronics and Electrical Engineering. - Kaunas: Technologija, 2006. – No. 4(68). – P. 55–58.

[7] Taeho C. Jo"The Effect of Virtual Term Generation on the Neural-based Approaches to Time Series Prediction" Ottawa, Ontario, Canada, pp 516-520 June 2003.

[8] Qigang Zhao, Xuming Fang, QunzhanLi,Zhengyou He"WNN-Based NGN TrafficPrediction" pp230-234, 4 -8 April 2005

[9] S. Satsri S. Ardhan V. Chutchavong and O. Sangaroon "ANN based NGN IP Traffic Prediction in Thailand" International Conference on Control, Automation and Systems 2007 Oct. 17-20, 2007 in COEX, Seoul, Korea

[10] Rumelhart, D. E., G. E. Hinton, and R. J.Williams, 1986, "Learning internal representations by error propagation," in D. E. Rumelhart and J. L. McCleland, eds. (Cambridge, MA: MIT Press), vol. 1, Chapter 8.

[11] Nonlinear Dynamical Systems: Feed Forward Neural Network Perspectives by Simon Haykin et. al Wiley-Interscience (February 21, 2001) ISBN-13: 978-0471349112.

[12] MATLAB Neural Network Toolbox TM 6 User's Guide Howard Demuth et. al The Math Works TM October 2008.

[13] Wan, E. A., 1994, "Time series prediction by using a connectionist network with internal delay lines," in A. S.Weigend and N. A. Gershenfield, eds., *Time Series Prediction: Forecasting the Future and Understanding the Past* (Reading, MA: Addison-Wesley), pp. 195–217.

[14] Essential SNMP: Simple Network Management Protocol. By Douglas R. Mauro, Kevin J. Schmidt Published by O'Reilly, 2001 ISBN 0596000200, 9780596000202

[15] CACTI open source network monitoring http://www.cacti.net/documentation.php

[16] RRDTOOL open source time series graphing software http://oss.oetiker.ch/rrdtool/doc/index.en.html

[17] Artificial neural network based peak load forecasting using Levenberg-Marquardt and quasi-Newton methods Saini, L.M.; Soni, M.K. Generation, Transmission and Distribution, IEE Proceedings- Volume 149, Issue 5, Sep 2002 Page(s): 578 - 584