

Neural Network Based Traffic Prediction for Wireless Data Networks

Gowrishankar*

*Department of Computer Science and Engineering, B.M.S College of Engineering,
Visvesvaraya Technological University,
Bangalore, Karnataka 560 019, India*

P.S. Satyanarayana

*Department of Electronics and Communication Engineering, B.M.S. College of Engineering,
Visvesvaraya Technological University,
E-mail: pss.ece@bmsce.ac.in
www.vtu.ac.in*

Received: 21-02-2008

Revised: 06-06-2008

Abstract

In a wireless network environment accurate and timely estimation or prediction of network traffic has gained much importance in the recent past. The network applications use traffic prediction results to maintain its performance by adopting its behaviors. Network Service provider will use the prediction values in ensuring the better Quality of Service(QoS) to the network users by admission control and load balancing by inter or intra network handovers. This paper presents modeling and prediction of wireless network traffic. Here traffic is modeled as nonlinear and non-stationary time series. The nonlinear and non-stationary time series traffic is predicted using neural network and statistical methods. The results of both the methods are compared on different time scales or time granularity. The Neural Network (NN) architectures used in this study are Recurrent Radial Basis Function Network (RRBFN) and Echo state network (ESN). The statistical model used here in this work is Fractional Auto Regressive Integrated Moving Average (FARIMA) model. The traffic prediction accuracy of neural network and statistical models are in the range of 96.4% to 98.3% and 78.5% to 80.2% respectively.

Keywords: Traffic flow, Time series, QoS, Prediction, FARIMA and Neural Networks.

1. Introduction

In recent past, due to the technological breakthroughs in the field of wireless system, has enabled the pervasive acceptance and deployment of wireless network. Still, wireless system is not a preferred choice to the wired counterpart and the reason being that wireless network will not ensure guaranteed QoS due to the unpredictable behavior of network traffic. The unpredictability is for various factors such as user mobility, arrival pattern and diversified network requirement of user application.

This leads to nonlinear and time varying traffic flow between wireless network infrastructure like wireless Access Points (AP)/Base Station (BS) and set of wireless devices using the network infrastructure. Now it is increasingly becoming important to provide guaranteed QoS to the users of the network that can be achieved by one or more remedial measures like admission control and load balancing by intra or inter network handovers [1]. For these measures accurate and timely forecasting of network traffic is the critical factor. The remedial measure in turn improves the QoS

* Department of C. S. & E., B.M.S. College of Engineering, Bull Temple Road, Hanumthta Nagar, Bangalore – 560019, Karnataka, India, gowrishankar.cse@bmsce.ac.in, www.bmsce.ac.in.

Gowrishankar, P.S. Satyanarayana

of a wireless network and hence the wireless network will be more reliable

The nucleus of any prediction mechanism is to monitor the past and present behavior of the system and then establish statistical relationship between set of inputs to the set of outputs over a given time scale or time granularity. This relationship is either linear or nonlinear and time invariant (stationary) or time variant (non-stationary). The wireless network traffic is a nonlinear and non-stationary system [2]. The number of users and their network requirement is nonlinear due application characteristics and user behavior [3]. The time varying behavior of the traffic is due to terminal mobility and network characteristics [4]. The network applications are classified into four classes: conversational, streaming, interactive and background.

The first two classes' of network bandwidth requirement is relatively constant and predictable. The third class of network bandwidth requirement is based on the user behavior in the form of mouse clicks. The fourth and the last class of network bandwidth requirement is always on the state of the network [5]. Number of user and the amount of traffic generated by them has no correlation. Hence the network traffic is nonlinear to the number of users served by the wireless network [6]. The number of users served by each AP or BS is dynamic due to the user preference, load balancing, inter/Intra network handover, admission control policy and network characteristics like fading and time-variant noise [4][7]. Neural networks are the efficient methods to model, evaluate and predict the behavior of nonlinear and non-stationary systems [8].

In this work, first traffic time series is extracted from realistic wireless trace [9]. The missing values of time series are imputed using Radial Basis function Network (RBFN). Then traffic time series is modeled and predicted using two different neural network predictors and finally results are compared with statistical model.

The rest of the paper is organized as follows: Section 2 presents review of the related works in the field of network traffic prediction and estimation. In section 3, time series extraction and imputation technique is described. In section 4, wireless network and its traffic modeled from perspectives. In section 5 the results of NN based and statistical predictors are validated with the set measurement trace. Conclusions are presented in Section 6.

2. Related Works

There are several research works on network traffic modeling and prediction of wireless network. Most of these works have used network traces for the study and these trace data are collected by three techniques, *SNMP* polling, *Syslog* or *tcpdump* from network sniffers [10]. The *SNMP* trace will specify total bytes transferred, number of bytes transferred with error and layer specific information of network [11]. The *Syslog* traces will enumerate the statistics of user behavior in terms of mobility and usage pattern of network. Sniffer will provide detail information about the size of packet, packet inter-arrival time and interface specific information of the network. The granularity of *SNMP*, *Syslog* and *tcpdump* trace is in terms of few seconds, one second and microseconds respectively. In [12], *tcpdump* and *SNMP* traces were used to study the variance of number of users, packet latency and handoffs between AP and laptops for the period of eight days. In [13], *SNMP* traces were used to characterize WLAN usage, user arrival and session duration, statistical model were proposed for network usage, arrival and session duration. In a study conducted at Dartmouth College [14], *SNMP*, *Syslog* and *tcpdump* traces were used to analyze the usage pattern and user distribution in a large wireless network. In [12], *tcpdump* trace is used to show the traffic characteristics from AP perspective and they also studied the behavior of traffic from Physical, Data link, IP and Transport Layer. Recently [15], *SNMP* trace is used to analyze the traffic characteristics like aggregate network load and its periodicity. Some works have also been reported in the field of network traffic prediction. The prediction process can be classified as long and short term. In long term prediction the process of prediction time granularity is in the order of week and month. In short term prediction the granularity is in terms of milliseconds, seconds, minutes and hours. The aggregated IP Backbone traffic is modeled for large time scale (weekly basis) and accurate forecasting for the period of six months by lower order ARIMA model [16]. In recent studies FARIMA process is used to model crucial features of self similar traffic like long range dependence (LRD) and short range dependence (SRD) of broadband internet traffic. Short term traffic forecasting in terms of millisecond to seconds of granularity were also reported in these works [17][18]. In recent past [15], Normalized ARIMA process and

historical means of recent traffic were used to forecast hourly and daily trend of wireless traffic. The perceptron neural network model with backpropagation learning algorithm is the used model in self similar traffic patterns [19]. Recently, Generalized Mean Neuron (GMN) model is used for forecasting of weekly internet traffic of a WAN router [20].

In the above works, the work is either limited to modeling of wireless traffic or forecasting the wireline or wireless network traffic (both short term and long term trends). In wireless network the forecasting granularity is limited to hourly trend of the traffic and due to the large spikes and non-stationary behavior of time series of traffic, the series is smoothed and converted in to stationary series by applying power transformation technique. In the proposed work the wireless traffic is modeled as set of random processes and for the purpose of forecasting the wireless traces were used to construct time series. These time series were used to forecast traffic for single step and multistep step for second to minute granularity.

3. Time Series Extraction and Imputation

Wireless trace contains set of events $x(n)$, logged at times t_n , for $n = 0, 1, 2, \dots, N$, where N is the number of events. The events may be packet transmitted from AP/BS to Mobile Terminal (MT) or vice versa and association and dis-association of MT with AP/BS due to the mobility of MT. The trace contains enormous data about user, user applications and the vital characteristics of network. The process of information extraction from the trace and represent this information in a time scale is termed as 'time series extraction' [21]. The time series generated from this process may have few missing terms. This is due to the minor flaw in the wireless measurement process of measurement tool. The data imputation technique will estimate the missing terms.

3.1. Time Series Extraction

The trace is a time stamped recorded of events and it is an output of measurement process. The trace need to be converted in to 'time series' with the regular time intervals. This will be used as an input to the traffic predictors for the purpose of prediction.

Suppose the time granularity of the time series or sampling interval of trace is $\Delta t = t_{k+1} - t_k$. Then $y(t)$ is a series of equally spaced observation of network traffic,

$\{y(t_1), y(t_2), \dots, y(t_k), \dots, y(t_n)\}$. An observation $y(t_k)$ corresponds to amount of traffic recorded in the trace at the k^{th} time interval i.e. $y(t_k) = \sum_k x_k$, $k = 1, 2, \dots, n$,

where x_k is the amount of traffic recorded in the trace during the time interval t_k , which is the sampling interval of trace. The sampling interval may be 0.1, 1, 10 or 60 seconds. Fig.1 shows time series representation of the wireless traffic for 1 second interval.

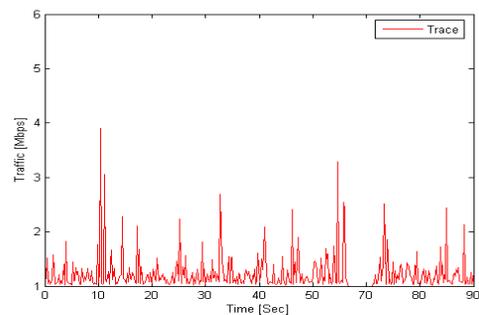


Fig. 1. Time Series of Traffic Trace.

3.2. Time Series Imputation

The accuracy of the prediction system depends upon the choice of prediction model and the integrity or completeness of the collected trace [22]. The trace collected from the measurement may, sometimes, become incomplete due to the malfunctioning of measurement tools or human errors. In a prediction system, the future value is predicted using the past information i.e. If the predicted value of the predictor at time n is denoted by, $\hat{y}(t_n)$, then the predicted value of the time series can be represented by the function $\hat{y}(t_n) = f[y(t_{n-1}), \dots, y(t_{n-k})]$, where k is the number of previous steps. If some, $y(t_{n-j}), \forall 1 \leq j \leq k$, are missing from the input, the value of $\hat{y}(t_n)$ cannot predicted correctly by the predictor. There are two general approaches to deal with the problem of such type of missing data: They could be ignored or imputed (filled) with new values [23]. The first solution is applicable only when small amount of data is missing and it reduces the efficiency of prediction. Hence it is more constructive and practically viable to consider imputation.

Gowrishankar, P.S. Satyanarayana

Here NN system is used for the purpose of imputing the missing values of time series.

In NN, the problem of time series imputation is considered as a ‘function mapping’ technique [24]. The feed-forward NN has the ability to map any nonlinear and non-stationary function to any arbitrary degree of accuracy [25]. One such popular feed-forward NN is the Radial Basis Function Network (RBFN). It is a single, hidden layer feed-forward network. Each node in the hidden layer has a parameter vector called ‘center’. These centers are then compared with network inputs to produce radially symmetrical responses. These responses are scaled by ‘connection weights’ of the output layer to produce network output using Gaussian basis functions:

$$\hat{y}(t_n) = \sum_{i=1}^k w_i \exp\left(-\frac{\|y - \mu_i\|}{2\sigma_i}\right) \quad (1)$$

Where σ_i is the radius of the influence field of the hidden layer neuron, k is the number of hidden layer neurons, $y = [y(t_{n-1}), y(t_{n-2}), \dots, y(t_{n-k})]$ and μ_i are the input vector and prototype vector respectively. Fig.2 shows the imputation of missing values in the time series by RBFN.

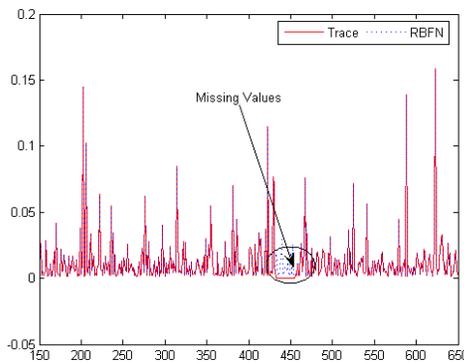


Fig. 2 . Time series Imputation by RBFN

4. System Model

4.1. Traffic model

In any stationary time series system, the series can be modeled as $y(t) = f(x_t)$, where $x_t = [y(t-1), \dots, y(t-m)]$. However this is not true in case of non-stationary time series, the entire series cannot be determined by single function $f(\cdot)$, instead by set of function f_1, f_2, \dots, f_k where k is the number of factors determine each elements of the time series, hence the non-stationary time series can be modeled as

$$y(t) = \sum_{i=1}^k p_i^t f_i(t) \quad (2)$$

$$\sum_{i=1}^k p_i^t = 1$$

Where $f_i(t)$ is a random process and hence the non-stationary time series can be modeled by a set of random process[26]. Since the traffic at AP or BS is a non-stationary time series, this can be determined by number of users and their arrival pattern, numbers of sessions(application) of individual user, session inter arrival and size of each session. All these are random processes [27].

$$Tr(t) = \sum_{i=1}^3 p_i^t f_i(t) \quad (3)$$

Where $f_1(t)$ is random process of number of users and the user arrival pattern, $f_1(t)$ is discrete value continuous time process User distribution either uniform or lognormal process and user arrival pattern is time varying Poisson process[10][27][28]. The $f_2(t)$ is random process of number of session and session inter arrival pattern, $f_2(t)$ is discrete value and continuous time random process, user session is Lognormal process or time varying Poisson process and session inter arrival is a BiPareto, Weibull, Markov Modulated Poisson Process(MMPP) or Time varying Poisson process[13][27][29]. The $f_3(t)$ is the size of each session, $f_3(t)$ is discrete value continuous time process and either BiPareto or Lognormal random Process[27][29].

4.2. FARIMA model

There is a well established theory for statistical modeling and forecasting of stationary time series by Auto regressive Moving Average (ARMA) Process, as the order of Auto Regression(AR) value increases the system can capture the seasonal trend of the time series and this leads to better approximation of stationary time series and the Moving Average(MA) value tries to remove effect of unknown initial value of the series[30]. The combination of AR and MA process will lead to high prediction accuracy of stationary time series with removal of dependency of series by unknown initial values. However, in case of non-stationary time series such approximation will not be possible by simple ARMA process due to the fact that time series represents only one realization of set of stochastic process. For the better forecasting and approximation the time series can be decomposed into permanent (non-stationary) and transitory (stationary) parts [31].

$$y(t) = y^p(t) + y^k(t) \quad (4)$$

Here $y^p(t)$ denotes permanent and $y^k(t)$ is transitory component. Such decomposition will help in determining the Long Range Dependence(LRD) and Short Range Dependence(SRD) properties of the time series separately[30]. A non-stationary process $y(t)$ is transformed into stationary process by differencing the process d times and the parameter d determines the LRD properties of the process.

$$(1-L)^d = \delta + x(t) \quad (5)$$

Where L is backward shift operator, $\delta = [y(t) - y(t-1)]$ and $x(t)$ is ARMA process, the parameter of ARMA process p and q determines SRD properties of the process. The original non-stationary process is denoted as $ARIMA(p,d,q)$. The accuracy of non-stationary time series prediction lies on the estimation of differencing term d . If the degree of differencing d is allowed to take non-integral value and the range is between $-1/2 < d < 1/2$, then such process is called as $FARIMA(p,d,q)$ process[30]. When the non-stationary time series is modeled by $FARIMA(p,d,q)$ process, the $FARIMA(p,d,q)$ model can be used to predict the future steps of time series[27].

$$\hat{y} = \sum_{j=1}^{n-1} \pi_j y_{n-j}$$

$$\pi_j^n = \pi_{j+n-1} - \sum_{i=1}^{n-1} \pi_i \pi_j^{n-1}$$

$$\pi_j = \frac{(-1)^j \Gamma(1+d)}{\Gamma(1+j)\Gamma(1+d-j)} \quad (6)$$

Where $\Gamma(\cdot)$ is a gamma function and \hat{y} is the predicted value of the time series from $FARIMA(p,d,q)$ model.

4.3. Neural network model

Feed-forward NN has the ability to map any nonlinear and non-stationary function to an arbitrary degree of accuracy [25]. One such popular feed-forward network is the RBFN. It is a single hidden layer feed-forward network. Each node in the hidden layer has a parameter vector called as center. These centers are used to compare with network input and produce symmetrical radial response. These responses are scaled by connection weights of the output layer and then produce network output, where Gaussian basis function is used and given by.

$$\hat{y} = \sum_{i=1}^n w_i \exp\left(-\frac{\|y - \mu_i\|^2}{2\sigma_i}\right) \quad (7)$$

Here σ_i is the dimension of the influence field of the hidden layer neuron, y and μ_i are the input vector and prototype vectors respectively. RBFN has achieved considerable success in nonlinear function prediction but the performance of RBFN is less satisfactory for the nonlinear and non-stationary function prediction [32]. The Recurrent Radial basis function network considers the time as an internal representation and the non-stationary aspect of nonlinear function can be obtained by having self-connection on the input neuron of sigmoidal firing function and the recurrent weights are in the range $[-1, +1]$, with normal distribution. This is a

Gowrishankar, P.S. Satyanarayana

special case of locally recurrent, globally feed-forward neural network [33]. The RRBFN output for Gaussian basis function is

$$\hat{y} = \sum_{i=1}^n w_i \exp \left(- \frac{\sum_{j=1}^m (y^j - \mu_i^j)^2}{\sigma_i} \right) \quad (8)$$

Here \hat{y} is the predicted value of the time series, j and i are the number of neurons in the input layer and hidden layer respectively. The architecture of RRBFN is shown in Fig.3.

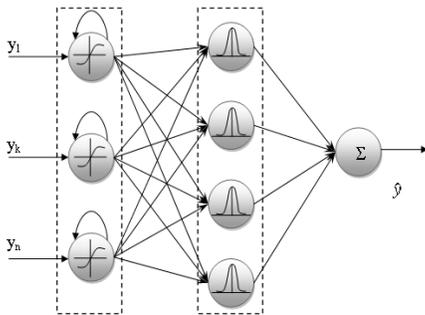


Fig. 3 . Recurrent Radial basis Function Network

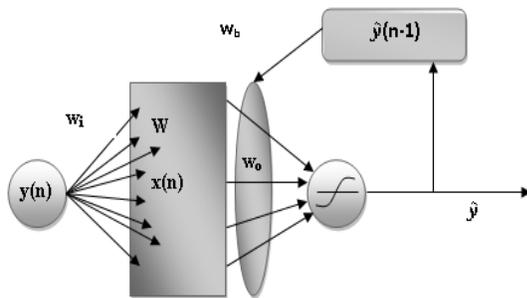


Fig. 4. Echo State Network

The ESN is a class of Recurrent NN [34]. It has two functional units: a dynamical system with a rich set of dynamics followed by a memory-less output readout unit shown in Fig 4. The dynamical system is also called as ‘reservoir’. It consists of large number of neurons that are randomly interconnected and self-connected. These connections are fixed. The biggest advantage of this network is in its learning algorithm. The speed of learning is unique among all RNNs. During the training process only weight of memory-less output function is changed through offline linear regression process or by online methods like Recursive Least Square (RLS) method. The general state of reservoir is given by

$$x(n) = \varphi(w_i y(n)) + wx(n-1) + w_b \hat{y}(n-1) \quad (9)$$

Where $\varphi(.)$ is a sigmoidal activation function, $y(n)$ is a current input vector, $x(n-1)$ is the state of reservoir at time step $(n-1)$, $\hat{y}(n-1)$ is the output of ESN at previous time step, w_i, w and w_b are the input, reservoir and feedback weight vectors respectively. The output of ESN at time step n is given by.

$$\hat{y} = \varphi_o(w_o x(n)) \quad (10)$$

Here, $\varphi_o(.)$ is the linear or sigmoidal activation function of the output neuron, w_o is the output weight vector and $x(n)$ is current state of the reservoir.

5. Traffic Prediction

The traffic prediction process involves two major phases, the first phase is the time series extraction of data traffic and second phase is neural network and statistical model based time series prediction of wireless data traffic.

For the purpose of time series extraction of data traffic, the wireless trace available at CRAWDDAD data repository is used [9]. The original dataset contains the traffic trace of 476 wireless APs for the period of 11 weeks [14]. Here the traffic prediction is limited to single AP for the time granularity of one second, ten second and one minute. The original trace file is filtered to obtain the traffic of single AP and for this *SNMP* trace is used to extract the detail traffic from *tcpdump*

file. The filtered traffic trace is further processed to obtain amount of traffic in bytes per unit time. The filtered trace contains traffic between AP and set of wireless users for the period of one day. This traffic is combined and arranged to form a time series in to the different time scales such as 1 millisecond, 1 second and 1 minute. Once the time series is extracted from network trace, the traffic is forecasted for different time steps and the steps refers to the unit of time .The filtered traffic has to be normalized for the purpose of n-step ahead prediction[35].

The Neural network based traffic prediction involves training and testing of RRBFN and ESN predictor. The first phase extracted traffic time series used for training and testing of the predictors. The training and testing samples are randomly picked from the sample size of 1000. The RRBFN network has three layers: input, hidden and output. Here there are 300 neurons in the input layer with sigmoidal activation function and with the recurrent connections, the range of recurrent weights are -1 to +1. The hidden RBF layer has 175 neurons with RBF activation and output layer has single neuron with linear activation. The ESN network has 350 neurons in the reservoir with 75% of recurrent connection in the range of weights between -1 to +1. Input weights are in the range of -0.40 to +0.40 and recurrent weights are in the range of - 0.6 to +0.6. Spectral radius is set to 0 and all the layer neurons have sigmoidal activation function and the predicted step sizes are 1 and 10 for both RRBFN and ESN.

In FARIMA based traffic prediction, the first step is to build $FARIMA(p,d,q)$ model to describe wireless trace and the next step is to predict the network traffic for different time scale using FARIMA model . The process of building FARIMA model involves four stages [17]. First stage is to preprocess the normalized traffic to obtain zero mean time series. Next is to approximate the value d , where $d = H - \frac{1}{2}$, the value H refers to the Hurst Parameter, the Hurst Parameter is estimated by Variance- time plot of the normalized trace. Then the exact value of d can be obtained by fractionally differencing the $y(t)$ to obtain $x(t)$ of equation 3, where $x(t)$ is ARMA model of the original time series. Finally ARMA process parameter p and q , where in initially p and q values should start from smallest value, the best (p,q) value is obtained by model identification and diagnostic checking process[30].The model parameters for various time granularity is listed

in Table 1. From the estimated FARIMA process and current and past values of time series, the future steps of time series are predicted. The predicted step sizes are 1 and 10.

Table 1. FARIMA Model Parameters

Trace	p	q	d	Mean	Variance
1 (Sec.)	4	2	0.0178	0.0126	0.00067
10 (Sec.)	4	2	0.1865	0.1246	0.0036
1 (Min.)	4	2	0.3457	0.4101	0.0065

Fig. 5 shows the 10 step traffic prediction from RRBFN, ESN and FARIMA model, for 1 second time granularity. The 1 step traffic prediction for 10 second time granularity is shown in Fig. 6. The Fig. 7 shows 10 step predictions for the time granularity of 1 minute.

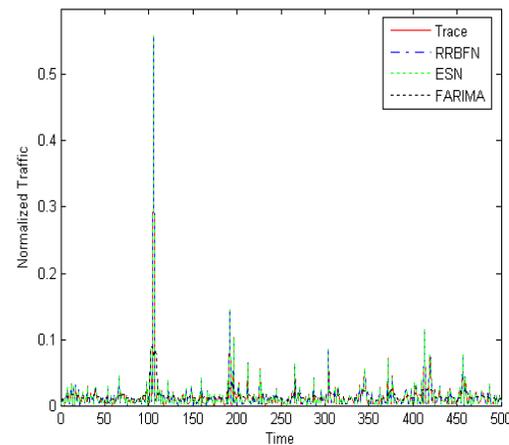


Fig. 5 . 10 – Step Traffic Prediction (1 –Second Interval)

Gowrishankar, P.S. Satyanarayana

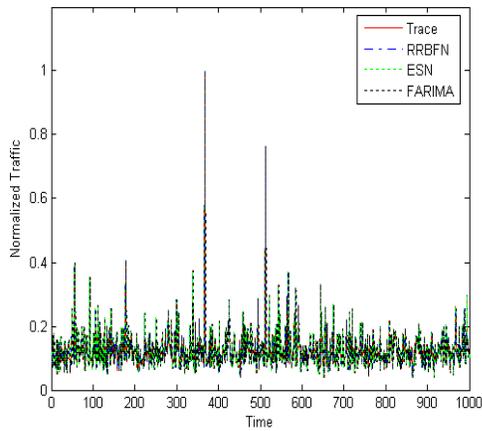


Fig. 6. 1-Step Traffic Prediction (10-Second Interval)

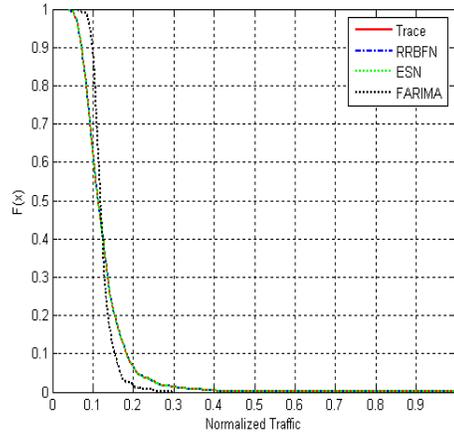


Fig. 8. 10-Step CCDF Plot (1-Second Interval)

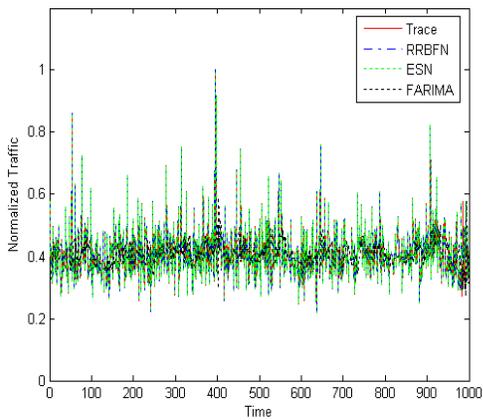


Fig. 7. 10-Step Traffic Prediction (1-Minute Interval)

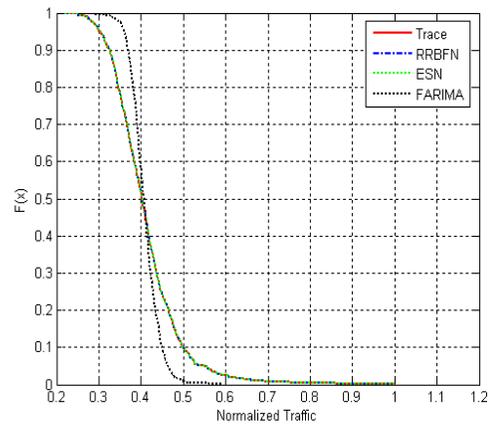


Fig. 9. 10-Steps CCDF Plot (1-Minute Interval)

Complementary-CDF (CCDF), plot for trace and predicted samples of 10 step of 1 second granularity shown in Fig 8. The Fig.9 is CCDF plot for 10 step prediction output of 1 minute granularity.

The Quantile-Quantile Plot (Q-Q Plot) RRBFN, ESN, FARIMA Models output and trace for 10 step of 1 minute granularity is shown in Fig 10.

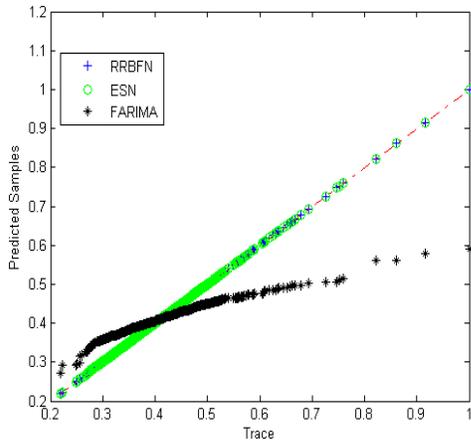


Fig. 10. 10-Step Q-Q Plot(1-Minute Interval)

Normalized Auto correlation Function (ACF) plot for trace and predicted samples for 10 – step ahead prediction of 10 seconds granularity is shown in Fig. 11.

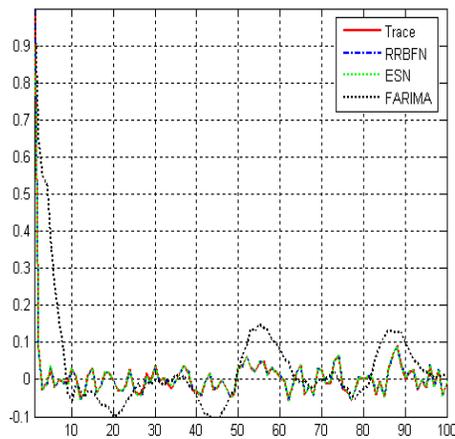


Fig. 11. 10-Step ACF Plot (10-Second Interval)

From CCDF plot, the functional behavior of trace and model output can be visually inspected. In Q-Q plot and ACF plot the first and second order statistical relationship between normalized trace and predicted output of neural and statistical models can be visually determined [27].

6. Conclusion

In this paper wireless network traffic for short time step is modeled as the time series and future values of time series are predicted by RRBFN, ESN and FARIMA models. The prediction accuracy of both neural network models is found to be superior for 1 second, 10 seconds and 1 Minute time granularity. The Prediction accuracy of RRBFN model for three time granularity ranges from 96.4% to 98.3%. Prediction accuracy of ESN model for three time granularity found to be 97.8%, to 98.1%. The prediction accuracy of FARIMA model of three time granularity seems to be less satisfactory compare to neural models. The prediction accuracy of FARIMA model is in the range of 78.5% to 80.2%. The FARIMA prediction accuracy will be slightly increased 1 minute granularity, this is due to the fact that the time aggregation makes the wireless trace exhibit high degree of self similarity, the estimated Hurst Parameter H from variance time plot is 0.8457. In a shorter time granularity time series behavior is nonlinear and non-stationary due to burstiness in the network traffic and user behavior.

In neural Network Models, ESN based predictor is slightly better in terms of prediction accuracy, but network takes more training samples and longer training time compared to RRBFN. During the training phase the training Mean Squared Error (MSE) for ESN and RRBFN are 0.01254 and 0.0162 respectively.

In general, the results show that neural network predictors are better than statistical models. The performance of FARIMA model is based on parameter estimation. The process of parameter estimation is trial and error and this affects the performance of FARIMA model to a large extent. The property of burstiness of wireless network traffic will be better captured by neural network predictors and the long range dependency is captured by adjusting the weights of the neurons during the training phase. In short range

Gowrishankar, P.S. Satyanarayana

dependency the input is confined by the recurrent architecture of the network.

Future work includes applying NN models for long term prediction of wireless traffic. Furthermore, the NN based predictors are used for traffic prediction and Bit Error Rate (BER) prediction. The NN based traffic and BER predictors are used in Multi Criteria Decision Making (MCDM) for inter network handover (Vertical handovers) in heterogeneous wireless networks

Acknowledgements

The authors wish to thank the editor and anonymous reviewers for their valuable comments and suggestions that helped to improve the presentation of this paper. This work is supported partly by Technical Education Quality Improvement Program (TEQIP) Government of India.

References

1. W.Xiuchao and A.L.Ananda, Link Characteristics Estimation for 802.11 DCF Based WLAN, in *Proc 29th Annual IEEE International conference on Local Computer Network*, 2004, pp. 302-309.
2. Y.Chan and K.Chandra, Time Series Models for Internet Data Traffic, in *Proc. International Conference on Local Computer Networks*, 1999, pp.164 -171.
3. H. Choi and J.Limb, A Behavioral Model of web Traffic, in *Proc. Seventh International conference on network Protocols*, 1999, pp. 327-334.
4. W.Hsu and A.Helmy, On modeling user Association in Wireless LAN Traces on University Campus, in *Proc. 2nd IEEE International workshop on wireless network measurement (WINMEE 06)*, 2006.
5. ETSI, Universal Mobile Telecommunication Systems (UMTS); Selection Procedures for the Choice of radio Transmission Technologies of the UMTS, *Technical Report TR 101 112 v. 3.2.0*, 1998.
6. J. Lee, S. Choi and H.W. Jung, Analysis of User Behavior and Traffic Pattern in a Large-Scale 802.11 a/b Network, in *Proc. 1st IEEE International workshop on wireless network measurement (WINMEE05)*, 2005.
7. J.Landman and P.Kritzinger, Delay Analysis of downlink IP Traffic on UMTS Mobile, *Performance Evaluation*, 62, (1-4), (2005), 68 -82.
8. Y.K.Tay, J.O.Kyong, K.Chiho and et.al., Artificial Neural Network for non-stationary time series, *Neurocomputing*, 61, (2004), 439 -447.
9. CRAWDAD: <http://crawdad.cs.dartmouth.edu> (November 25, 2007).
10. S.Rajeev, A.L.Ananda, C.C.Mun and et.al. *Mobile, Wireless and Sensor Networks: Technology, Applications and Future Directions*, (John Wiley & Sons, Inc 2006).
11. IEEE computer society, *LAN MAN standards committee IEEE 802.11 Management Information Base in IEEE std. 802.11-1999*, 1999
12. Y.Jihwang, Y.Moustafa. and A.Ashok, Characterizing the IEEE 802.11 Traffic: The Wireless Side, in *Proc. International workshop on wireless traffic measurement and modeling (WiTeMeMo'05)*, 2005
13. B.Anand, M.V.Geoffrey, P. Bahal and P.Venkatragan, Characterizing the user Behavior and network performance in public wireless LAN, in *Proc. ACM Sigmetrics*, 2002.
14. K.David and E.Kobby, Analysis of Campus wide wireless network, *Wireless Network*, 11, (2005), 115 -113.
15. P.Maria, S.Haipeng, R.Elias and et.al, Short-term traffic forecasting in a campus-wide wireless network, in *Proc 16th IEEE International symposium on Personal, indoor and Mobile Radio communication*, 2004, pp.1446-1452.
16. P.Konstantina, T.Nina, Z.Zhang and et.al, Long-term Forecasting of Internet Backbone Traffic, *IEEE Transaction on Neural Networks*, 16(5), (2005), 1110-1124.
17. S.Yantai, J.Zhigang, Z.Lainfang and et.al. Traffic Prediction Using FARIMA Models, in *Proc. IEEE International conference on communication*, 2, 1999, pp. 891-895.
18. M.R.Corradi, G.Garropo S.Giordano and et.al., Analysis of f-ARIMA processes in the modeling of broadband traffic, in *Proc. IEEE International conference on communication*, 3, 2001, pp.964-968.
19. H.Yousefizadeh, Neural Network Modeling of ON-OFF Source Models with Self Similar Characteristics, in *proc. 1st Workshop on Fractals and Self-Similarity, ACM SIGKDD*, 2002.
20. R.N.Yadav., K.K.Prem and J.Joseph, Neural network learning with generalized-mean based neuron model, *soft computing- A fusion of foundations, Methodologies and Application*, 10(3), (2006), 257-263.
21. S.A.Sara, Integrated Multiquadric Radial Basis Function Approximation Methods, *Computers and Mathematics with Applications*, 51(8) (2006), 1283-1296.
22. S. Chiewchanwattana, C.Lursinsap and C.H. Chu, "Time-series Data Prediction based on Reconstruction of Missing Samples and Selective Ensembling of FIR Neural Network", *9th International Conference on Neural Information processing*, 5, (2002), pp 2152-2156.
23. A.Farhangfar, L.A.Kurgan and W.Pedryez, "A Novel Framework for Imputation of Missing Values in Databases", *IEEE Transaction on Systems, Man and Cybernetics-Part A: Systems and Humans*, 37(5), (2007), 692-709.
24. H. Junninen, H.Niska, K.Tuppurainen, J.Ruuskanen and M.Kolehmainen, "Methods of Imputation of Missing Value in Time Series", *Atmospheric Environment*, 38(18), (2004), 2895-2907.

25. M.Li., G.Huang, P.Saratchandran and N.Sundarajan, Performance Evaluation of GAP-RBF Network in Channel Equalization, *Neural processing Letters*,22,(2005), 223-233.
26. R.Dahlhaus, Fitting Time Series Models to Non-stationary Processes, *The Annals of Statistics*, 25(1), (1997), 1 -37.
27. H.Felix, K.Merkouris P.Maria and et.al, Spatio-Temporal Modeling of Traffic Workload in a campus WLAN, *in proc. First International workshop on Technology and policy for Accessing Spectrum (TAPAS'06)*, 2006
28. M.George, H.Y.W.Strasky, Y.Yuan and L.Songwu, Characterizing flows in Large Wireless Data Network, *in proc. Mobicom' 04*. 2004.
29. J.N.Carl, S.Iraj, S.Wim, A compounded model for TCP connection arrival for LAN and WAN Applications, *Computer Networks*, 40(3) (2002), 319-337.
30. K.Gebhard and W.Jurgen, *Introduction to Modern Time Series Analysis*, (Springer-verlag 2007).
31. R.N.Charls and K.Heejoon, Spurious periodicity in inappropriately Detrended Time Series, *Econometrica*, 49,(1981),741 -751.
32. S.Haykin, *Neural Networks: A Comprehensive Foundation* (Person Education Inc 1999).
33. R.Zemouri, D.Racoceanu and N.Zerhouni, Recurrent Radial basis function network for time series prediction, *Engineering Applications of Artificial Intelligence*,16, (2003), 453-463.
34. H.Jaeger, Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication, *Journal of Sciences*, 304 (2004), 78-80.
35. F.Huifang and S.Yantai ,Study on Network Traffic Prediction Techniques, *in proc. Intentional Conference on Wireless communications, Networking and Mobile Computing*,(WiCom 2005),2005,pp.1041-1044