International Journal of Computer and Electrical Engineering, Vol.2, No.6, December, 2010
1793-8163

Forecasting Network Traffic Load Using Wavelet Filters and Seasonal Autoregressive Moving Average Model

Akhter Raza Syed, Aqil Burney S.M. and Badar Sami *Members IACSIT and IEEE*

***Abstract*—A computer network's performance can be improved by increasing servers, upgrading the hardware, and gaining additional bandwidth. Another important issue to increase network performance is to use an appropriate mathematical model and using that model for forecasting the network load at peak hours. We analyze a network traffic time series of Internet requests made to a workstation. This series exhibits a long-range dependence and self-similarity in large time scale and exhibits multi-fractal in small time scale. With the growing demand of using computer networks, there is increasing demand to explore new techniques for forecasting. In this paper we used the wavelet filters based on multi-resolution analysis along with the Seasonal Autoregressive Moving Average (SARIMA) models for forecasting network traffic volume. We used the Daubechies 4 (db4) wavelet filters for compression of time series and SARIMA modeling for forecasting the series. Our proposed methodology is based on the theory that most of the network traffic signals are always having the noise and unnecessary details not required for further analysis hence forecasting directly from such signals may prove computationally intensive and prone to error. The denoised and compressed signals using wavelet filters will be comparatively more smoothed and helpful for speedy and accurate forecasts. We compare our results with simple SARIMA methodology and conclude that our proposed method give better and accurate forecasts. Using wavelet based SARIMA model, we believe that with the right choice of mother wavelet, moving towards a wavelet based forecasting system would significantly improve the forecasting efficiency.**

***Index Terms*—forecasting, network traffic, SARIMA, Wavelet.**

I. Introduction

Modeling and forecasting of network traffic data presents a number of challenges in recent paradigm due to the volatility of data. There are various methods used for forecasting time series including moving averages (MA), autoregressive moving averages (ARMA), autoregressive integrated moving averages [3], Fourier transform [2], artificial neural networks (ANN), and fuzzy logic [5][17]. A recently developed technique of wavelet has attracted the attention of researchers and is a new and rapidly growing area of research within finance, business, statistics, mathematics, computer science and engineering [15][8]. Using the wavelet transformation (WT), a multiresolution representation of a traffic signal is possible which breaks the signal into its shifted and scaled versions [14]. This breaking up of signal is used for smoothing of time series to differentiate what is a signal and what is noise. This filtered data is then further used to search time series models as possible candidates for forecasting. These models may be standard AR, ANN and fuzzy, etc to produce forecast that best approximate the mean and variance of actual traffic. Figure 2 and Figure 3 shows the internet traffic load over server aggregated half hourly and three hourly. In this paper, we presented an application of wavelets for forecasting network traffic data collected from the server of University of Karachi's High Speed Fibre Optics LAN with Wireless Computing Support (HSFB LAN) (see Fig. 1) using wavelet based seasonal autoregressive moving average model. We use wavelet filters to reduce the contamination of this signal and estimate models from SARIMA for forecasting. A comparison with Box and Jenkins SARIMA model is also given in this paper [3]. We prove that models using wavelet filters produce better forecasts as compared to unfiltered original time series. The rest of the paper is organised as follows. In section 2, we describe the wavelet properties that are most important for forecasting even non-stationary time series. In section 3, we give the brief introduction of Box and Jenkins SARIMA models; next we discussed our proposed model in section 4 and compare its forecasts with the forecast obtained from SARIMA model directly. In the last section, we give our conclusions.

II. Wavelet analysis

In this section, we will focus to describe some of properties of WT related to forecasting of time series data. There are commonly two types of wavelets decompositions [6][7][8][16], the continuous wavelet transforms (CWT) and discrete wavelet transforms (DWT). The first one transforms a time series over entire real axis where as second over discrete sets of points. The choice of selection between them depends on the task in hand and computational considerations. A DWT decomposition using Mallat's pyramid algorithm in multiresolution framework [11][12][13][14] is faster than CWT and many researchers

Manuscript received May 10, 2010. This work was supported by Higher Education Commission of Pakistan (HEC). We are also thankful to Mr. Zain Abbas who gives us full support in collection of network traffic traces

Akhter Raza Syed is the research fellow at Department of Computer Science University of Karachi also teaching there as visiting faculty e-mail: razasyed717@msn.com.

Aqil Burney S.M. is the Chairman and Professor at Department of Computer Science University of Karachi. He is also the project director of HSFB LAN burney@uok.edu.pk

Badar Sami is the Assistant Professor at Department of Computer Science University of Karachi also he is deputy project director of HSFB LAN e-mail: badarsami@uok.edu.pk

have considered DWT more appropriate.

A. Decomposition and Reconstruction of Time Series

Let $f(t)$ be a time series signal defined in $L^2(R)$ space, which denotes a vector space for finite energy signals, where 'R' is a continuous number system. The WT of $f(t)$ is defined as

$$CWT_\psi f(a,b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt \qquad (1)$$

where $\psi_{a,b}(t) = |2|^{\frac{1}{2}} \psi\left(\frac{t-b}{a}\right)$.

Here $\psi(t)$ is the base function often called mother wavelet, $a,b \in R,\ a \neq 0$ are the scale and translation parameters. If we take the discrete values for dilation and translation parameters by considering $a = a_0{}^m$ and $b = na_0{}^m b_0{}^m$, where a_0 and b_0 are fixed values and $a_0 > 1$ and $b_0 > 0$, $m,n \in Z$ and Z is the set of positive integers. Then the discrete mother wavelet will be as under:

$$\psi_{m,n}(t) = a_0{}^{\frac{m}{2}} \psi\left(\frac{t - na_0{}^m b_0}{a_0{}^m}\right) \qquad (2)$$

and the corresponding DWT is

$$DWT_\psi f(m,n) = \langle f, \psi_{m,n} \rangle = \int_{-\infty}^{\infty} f(t)\psi_{m,n}(t)dt \qquad (3)$$

It is also possible to reconstruct a signal f using wavelet transform through:

$$f = \sum_m \sum_n C_{m,n}\psi_{m,n} \qquad (4)$$

where $C_{m,n}$ are the wavelet coefficients which are calculated by the inner product of

$$C_{m,n} = \langle f, \psi_{m,n} \rangle \qquad (5)$$

The important point is that, equation 4 gives one-on-one correspondence of original signal in terms of its wavelet coefficients which means that data compression and high storage can be achieved by discarding certain coefficients that are in significant and can not contribute towards the reconstruction of original signal [6][7][8].

B. Data Compression using wavelet based Multiresolution Analysis (MRA)

In this section we will briefly discuss very important numerical application of wavelets i.e. data compression using MRA. The wavelet multiresolution analysis (MRA) is the process of decomposition of a discrete signal into approximate and detailed signals at each time scale [11], through a series of scaling functions $\phi_{j,k}(t)$ and wavelet functions $\psi_{j,k}(t)$, where $k \in Z$, using pyramid algorithm. These scaling and wavelet functions are obtained by dilating and translating the mother scaling function ϕ and mother wavelet function ψ as:

$$\phi_{j,k}(t) = 2^{-\frac{j}{2}} \phi(2^{-j}t - k)$$

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - k)$$

Its wavelet transform is as follows:

$$f(t) \cong \sum_k s_{J,k}\varphi_{J,k}(t) + \sum_k d_{J,k}\psi_{J,k}(t) + \sum_k d_{J-1,k}\psi_{J-1,k}(t) + \cdots \sum_k d_{1,k}\psi_{1,k}(t) \qquad (6)$$

The sums with coefficients $s_{J,k}$ and $d_{J,k}$ in equation 6 show the approximation at the coarsest scale and details at all scales respectively. It means the wavelet function is scaled and shifted along the axis and therefore the signal, too.

III. BOX-JENKINS APPROACH

A stochastic model, in which the current value is expressed as a finite, linear aggregate of previous values of the process and a shock, is an autoregressive (AR) process. The p^{th} order autoregressive process is

$$\phi(B)Y_t = a_t \qquad (7)$$

The model contains p+2 unknown parameters μ, ϕ_1, ϕ_2, … ϕ_p, σ_a^2. In model (7) $\phi(B)$ is an autoregressive operator, which is defined as $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p$. A finite moving average process of order q is given by the relation

$$Y_t = \theta(B)a_t \qquad (8)$$

It contains q+2 unknown parameters μ, θ_1, θ_2, … θ_q, σ_a^2 which have to be estimated before using the model. $\theta(B)$ is a moving average operator which can be defined as $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q$. To achieve great flexibility in fitting of actual time series, it is sometimes advantageous to include both autoregressive and moving average terms in the model. This leads to the mixed autoregressive moving average (ARMA) model. An ARMA(p, q) represent p^{th} order autoregressive and q^{th} order moving average model.

$$\phi(B)Y_t = \theta(B)a_t \qquad (9)$$

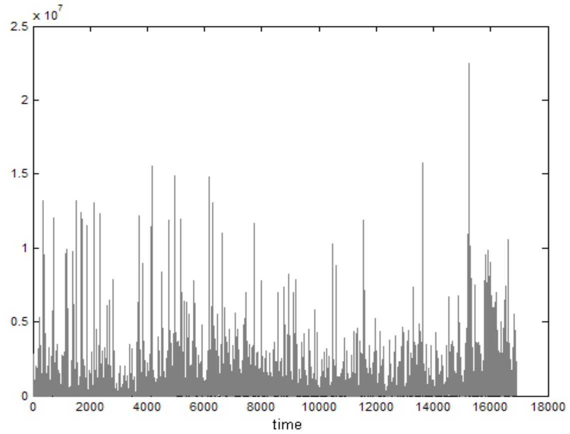

Fig. 2 Network Traffic over a serer aggregated ½ hourly

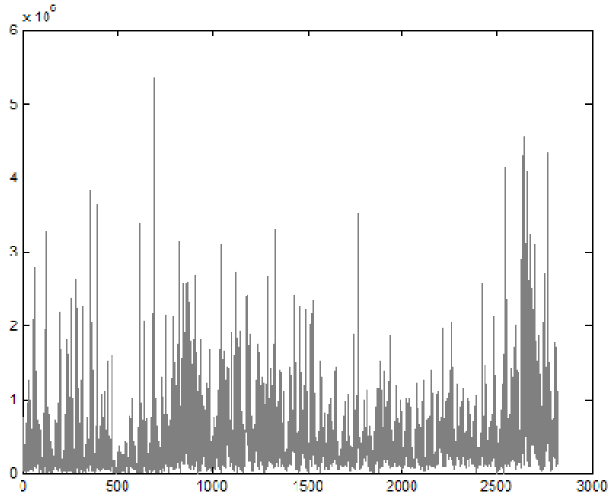

Fig. 3 Network Traffic over a serer aggregated 3 hourly

Many series actually encountered in business or industry exhibits non-stationary behavior and do not vary about a fixed mean. Models that describe such homogeneous non-stationary behavior can be obtained by supposing some suitable difference of the process to be stationary. These models are referred to as autoregressive integrated moving average (ARIMA) processes. The ARIMA(p, d, q) process is

$$\phi(B)\nabla^d Y_t = \theta(B)a_t \qquad (10)$$

Further considerations about ARIMA process can be found in [3][4]. In the real life many series exhibits periodic behavior with period ' s ', when similarities in the series occur after s basic time intervals. For the monthly series like the international airline passengers, the basic time interval is one month and the period of seasonality is $s = 12$ months. The ARIMA model defined above is not suitable in this situation another model known as SARIMA is defined. The seasonal autoregressive integrated moving average model of simple order (p,d,q) and seasonal order $(P,D,Q)_s$ is defined as.

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_S^D Y_t = \theta_q(B)\Theta(B^s)a_t \qquad (11)$$

IV. Wavelet Based Seasonal Autoregressive Moving Averages Model

As discussed in previous section wavelet can decompose a time series signal into well localized components that allow for zooming in of interesting features. The important details can be enhanced and unimportant details or noise can be left during reconstruction. We use these properties of wavelet in our proposed model for forecasting. In our proposed model, we decompose and smooth the signals using wavelet filtering techniques. We compress the signal while retaining more than 99% energy of original time series using global threshold method [9][10] in which magnitudes of all those wavelet coefficients which are more than a threshold level are set to zero. Hence we can write the equation 6 as

$$f(t) \cong S_J(t) + \sum_{i=1}^{J} D_i(t) \qquad (12)$$

which can be further reduced using partial discrete wavelet transform (PDWT) to

$$f(t) \cong S_{J_0}(t) + \sum_{i=1}^{J_0 \le J} D_i(t) \qquad (13)$$

The equation (13) will now be our reduced model on the basis of which we will compress our time series signals. The compression ratio will be the ratio between the length of reduced model and length of original signal. For example 20% compression ratio means we select 20% of the coefficient for compression of a signal and discard 80%. If the compression ratio is high we will better approximate the true signal.

V. Experimental Results and Evaluation

We applied the algorithm over the network traffic collected from University of Karachi High Speed Fibre Optics Local Area Network with Wireless Computing Support (HSFB LAN). The traffic over the server accumulated hourly. We collected large amount of data but few traces are mentioned in the table 1. We generated a large amount of data through simulation that is published in [1]. In this section we experimentally demonstrated the superiority of our model over traditional SARIMA model. We model and predict the network traffic load over the server, using SARIMA and our proposed technique. After the iterative testing of different wavelet families, we selected db4 wavelet, the most suited to our application. We compressed the original signal while using 20% of the transformed coefficients with db4 wavelets which retained more then 99% energy of the original time series. The compressed signal allows us to trade off a small amount of precision for very large gains in speed. The compression process removed most of the noise from signal. First, we estimate the models that are used for forecasting of original and filtered time series using the criteria of minimum sum of squared error (SSE), Akaike information criterion (AIC) and maximum adjusted R^2 (see table 2 and 3). We forecast the network traffic load using $(2,0,3)(1,0,1)_{12}$.

We used the SARIMA models for our comparative study of WBSARIMA for compressed signals. Results given in table 2, 3 and 4 show that the described model gives minimum SSE, AIC, BIC and maximum value of adjusted R^2, and has given much better forecast than linear SARIMA Models. This also shows the forecast evaluation statistics using original and compressed models.

TABLE I: A SEGMENT OF INTERNET TRAFFIC LOAD AT UNIVERSITY OF KARACHI SERVER ACCUMULATED HOURLY FROM 28-JUNE 2009 TO 2-JULY 2009

	28jun	29jun	30Jun	1Jul	2Jul
12AM	665	7437	674	165	707
1AM	2098	5564	1090	144	1238
2AM	382	1888	3374	157	3472
3AM	175	5523	2155	187	386
4AM	123	2873	226	202	161
5AM	136	750	383	157	160
6AM	133	369	815	136	203
7AM	11964	263	12921	13035	3243
8AM	70317	774	59770	72500	44571
9AM	123800	6852	79015	103083	83806
10AM	73451	8721	95836	60090	109697
11AM	138350	8532	96870	48112	125634
12PM	130300	10946	116448	78622	109987
1PM	109862	14940	39936	116994	50057
2PM	97955	15209	143111	96952	28406
3PM	32700	13277	87622	96247	20717
4PM	71029	9783	37975	92675	11207
5PM	46870	8273	32745	61327	8393
6PM	39363	5747	33177	58215	6640
7PM	13906	3611	21590	21597	1347
8PM	9250	2014	10705	7395	598
9PM	2068	381	5042	2565	401
10PM	5452	583	402	1072	419
11PM	2246	654	182	1438	340

VI. CONCLUSION

In this paper, we model and forecast Network traffic load over the University of Karachi's server using wavelet filters based SARIMA models. To achieve our goal, we decompose our original time series using best suited db4 wavelet filters and then reconstruct it using 20% best featured coefficient which compressed and denoised the signal. This reconstructed signal retains more then 99% energy of the original signal. We fitted models based on SARIMA to the original and compressed series. Model estimated on the bases of compressed series gave best short term forecast as compared to original SARIMA model. We believe that moving towards a wavelet based forecasting system would significantly improve the forecasting efficiency of time series data. We also conclude that models based on compressed time series using wavelets also improved the speed and efficiency of forecasted models as compared to traditional time series models.

TABLE II: EVALUATION FOR ORIGINAL SERIES

R-squared	0.8574910	Mean dependent var	30.987668
Adj. R-squared	0.8256782	S.D. dependent var	15.348765
S.E. of regr.	5.6876543	Akaike criterion	5.7654773
Sum sq. resid	4328.9871	Schawarz criterion	6.4567755
Log Likelihood	-188.56785	F-statistic	72.883902
Durbin-Watson	2.0198765	Prob(F-statistic)	0.000123

TABLE III: EVALUATION FOR COMPRESSED SERIES

R-squared	0.8987654	Mean dependent var	32.487665
Adj. R-squared	0.9123456	S.D. dependent var	12.468915
S.E. of regr.	5.6435673	Akaike criterion	6.7426073
Sum sq. resid	689.9871	Schawarz criterion	2.3456773
Log Likelihood	-398.56785	F-statistic	121.1313
Durbin-Watson	1.9238765	Prob(F-statistic)	0.000000

TABLE IV: COMPARISON OF FORECASTS

Time Jul 3rd	Original data	Forecast (original series)	Forecast (compressed series)
09 AM	93467	96789	94367
10 AM	109709	1123456	108765
11 AM	74398	74567	74987
12 Noon	71234	78765	72432
01 PM	62345	69876	62789
02 PM	69345	72345	70345

REFERENCES

[1] Akhter Raza Syed, Aqil Burney, S.M., Afzal Saleemi, "Simulation and analysis of Self-Similar Network traffic using WBFGN " proceedings page 148-152 2009 of 2nd International Conference on Environmental and Computer Science (28-30 Dec, ICECS 2009) IEEE Xplore
[2] Blackman, R B and Tukey, J W (1958), The Measurement of Power Spectra. New York: Dover.
[3] Box, G E P and Jenkins, G M, (1976), Time Series Analysis Forecasting and Control 3rd Edition, Pearson Education Inc.
[4] Box, G E P, W G Hunter, and J S Hunter, Statistics for Experimenters, Wiley, New York, 1978.
[5] Burney, S M Aqil, Tahseen, J and Cemal, A, (2004), "Levenberg – Marquardt Algorithm for Karachi Stock Exchange Share Rates Forecasting", International Journal of Computational Intelligence, Vol. 1. No. 2, pages 168-173.
[6] Daubechies, I (1988), "Orthonormal Bases of Compactly Supported Wavelets", Communication for Pure and Applied Mathematics, 41, pages 909-96.
[7] Daubechies, I (1989), "Orthonormal Bases of Wavelets with Finite Support- Connection with Discrete Filters in Wavelets, Time Frequency Methods and Phase Space", First Int. Wavelets Conf, Marseille, Dec. 1987, pages 33-66. Springer- Verlag.
[8] Daubechies, I (1992), Ten Lectures on Wavelets. Philadelphia, Pa.: SIAM
[9] Donoho, D L, Johnstone, I M, Kerkyacharian, G, Picard, D (1995), "Wavelet shrinkage: Asymptopia?" J. Royal Statist. Society, Series B, 57, pages 301-369.
[10] Johnstone, I M and Silverman, B (1994), "Wavelet Threshold Estimates for Correlated Noise", Technical Report, Stanford University
[11] Mallat, S G (1989a), "Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(R)$", Transactions of the American Mathematical Society, 315, pages 69-87.
[12] Mallat, S G (1989b), "A Theory for Multiresolution Decomposition", The Wavelet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11, pages 674-93.
[13] Mallat, S G (1989c), "Multifrequency Channel Decomposition of Images and Wavelet Models", IEEE Transactions on Acoustics, Speech, and Signal Processing, 37, pages 2091-110.
[14] Mallat, S G (1998), A Wavelet Tour of Signal Processing. San Diego: Academic Press.
[15] Mayer, Y (1992), Wavelets and Operators. Cambridge University Press.
[16] Percival, D B and Walden, A T (2002), Wavelet Methods for Time Series Analysis, Cambridge University Press
[17] Puyin, L and Hongxing, (2004), "Efficient Learning Algorithms for Three-Layer Regular Feed Forward Fuzzy Neural Networks", IEEE Transactions on Fuzzy Systems, Vol. 15, no. 3, pages 545-558

Akhter Raza Syed is a research fellow at University of Karachi, Pakistan. He received his M.Sc. Degree from University of Karachi in 1991. The M.Phil. work was converted to Ph.D. in 2006 and leading towards Ph.D. He is working as Assistant Professor in Education Department. Also he worked as visiting faculty member in Institute of Business Administration (IBA Karachi) from 2004 to 2009. He is also teaching as visiting faculty in department of computer science University of Karachi. He is a senior member of IACSIT

Dr. S. M. Aqil Burney is a Meritorious Professor and approved supervisor in Computer Science and Statistics by the Higher Education Commission of Pakistan. He is also the Director & Chairman at Department of Computer Science, University of Karachi. His research interest includes AI, Soft Computing, Neural Networks, Fuzzy Logic, Data Mining, Statistics, Simulation and Stochastic Modeling of Mobile Communication System and Networks and Network Security, Computational Biology and Bioinformatics.

Badar Sami is faculty member at Department of Computer Science, University of Karachi. Apart from teaching, research and management, he had also been Deputy Project Director of HSFB LAN. His research interests include Data/Text Mining; Localization; Systems Integration, Social, Ethical and Professional issues in IT.

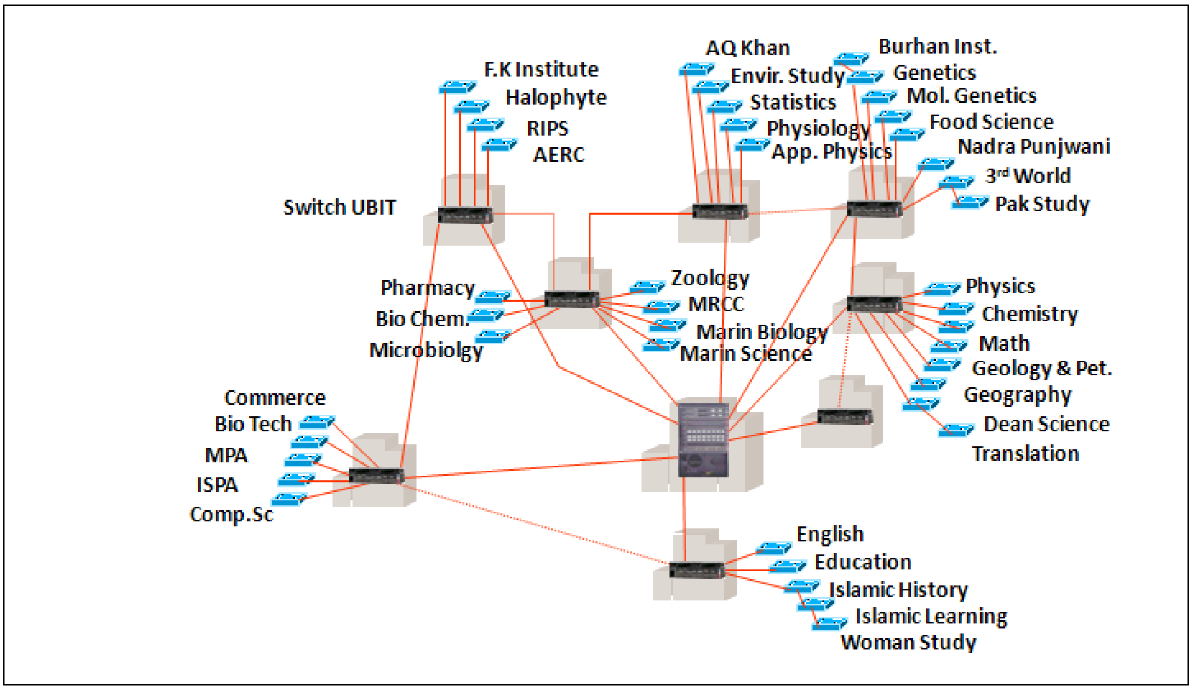

Fig1. High Speed Fibre Optic LAN with Wireless Computing Support (HSFB LAN) at University of Karachi