Asian Journal of Information Technology 15 (21): 4291-4298, 2016
ISSN: 1682-3915

Traffic Modeling for Next Generation Networks

[1]P.I. Liji and [2]S. Bose
[1]Anna University, Chenna, India
[2]Department of Computer Engineering, Anna University, Chennai, India

Abstract: Next Generation Wireless Network is Internet Protocol(IP) based that supports any time anywhere service and provide "Always Best Connected" (ABC) state. For real time multimedia service (eg: Video streaming, video conferencing, IPTV and online gaming) high Quality of Service (QoS) should be guaranteed with minimum delay, packet loss and jitters etc. Since NGNs is a ubiquitous wireless communication system seamless mo-bility need to be there to have a continuous service. Different technologies are integrated in to the NGNs and the frequent handover will be there with different technologies and therefore QoS need to maintain. Traffic in the network can be modeled in such a way that it should reduce the delay packet loss and jitters in the NGN environment. Conventional traffic modeling used to model the networks Poisson or Markovian process is inappropriate to model traffic in the networks. The time series Fbm, FGN or FARIMA can be used as traffic model to model the bursty traffic in the network. The FARIMA (p, d, q) is the best fit to model long range dependency as well as the short term dependency traffic in network and predit the future traffic from the present and past traffic. If we are able to predict the traffic in advance resource (mainly bandwidth) can be allo-cate in advance. Network performance like queue length, queue delay, loss probability etc can be anaylsed. Expermental result shows the reduction in buffer size and packet loss so as maintaine QoS in the network.For the queueing analysis and simulation has been done with large deviaton tech-niquie which represent the rare events occur in the most likely way.

Key words: Self similarity, long range dependency, FARIMA, QoS, India

INTRODUCTION

Next generation networks are IP based infrastructure that supports heteroge-neous access technology. The network would have a service provider which is equipped with multiple interfaces in the network. User can operate in cellular network technology and get handed over to a satellite based network and back to a fixed wireless network, depending upon the network coverage and prefer-ence of charging (Adas and Mukherjee, 1995). The service provider which is equipped with multiple interfaces (WiMAX, WLAN, GPRS etc) in the network has seamless mobility. The seamless mobilty in the networks will provide frequent handover from one technology to another. The real time application in the protocol level have stri-gent Quality Of Service (QoS) parameters (like minimum delay, mininum packet loss) will degrade performace of the network.If resource is allocated (mainly bandwidth) effiently, performance in the network can be improved so as to main-tain QoS given by the Service Level Agreement (SLA). Modeling the traffic can improve the resource allocation in the network (Leland *et al*., 1994). Conventional method used to models the traffic are Poission,Markoivn model or time series models like AR, ARMA, ARIMA models (Paxso and Floyd, 1995; Garrett and Willinger, 1994). These models are inappropriate to model the bursty traffic in high speed network with high bandwidth requirement and high variability. The high variabilty exhibits burstiness (peak rate/mean) for aggre-gated traffic can be described as a stochastic process with self similarity prop-erty. The persiestence at a wide range of time scale in arrival traffic input can be charaterised by property the Auto Correlation Function (ACF) will never di-verge to zero for large lag ie long range dependency; spectral density increases as with frequency.

Traffic modeling in the next genration networks should timely deliver real-time packets while minimizing packet losses virtually error-free transmission of non-real time packets good average delay and throughput performance of non-real time trace utilizing the bandwidth left unused by realtime trace fair usage of a channel for non-real-time trace among mobiles and low latencies for non-real-time packet transmissions and real-time connection setup and in handling handoff requests. The real time traffic having high variabilty and burstness over

Corresponding Author: P.I. Liji, Anna University, Chenna, India

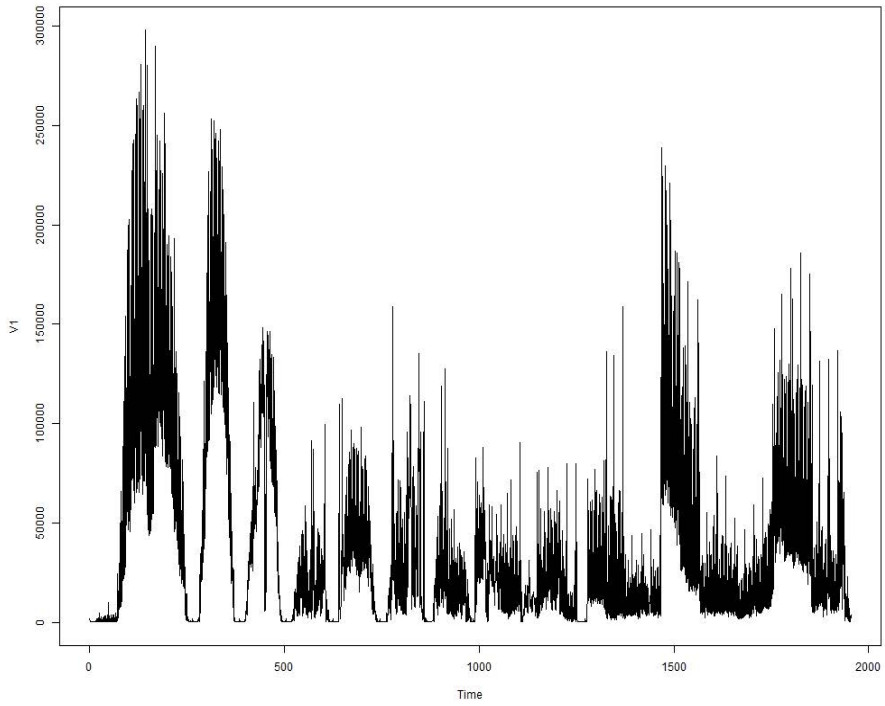

Fig. 1: Burst traffic

wide range of time scale and traditional traffic modeling like Poisson or Markovian process smoothed by averaging over a long enough time scale as shown in Fig. 1. The failure of the poisson model may results in the underestimation of the traffic burstiness and have greater impact in network performance including larger queuing delay and packet loss rate.

To model the traffic the packet arrived in the network can be considered as a stochastic process with time series $X_r(t)$ which is continues time or discrete time series and traffic can be viewed as a process defined by a set of packet arrival times t_1, t_2 ... The traffic in a network can be analyzed as Poisson pro-cess with inter arrival time in the network A_n with an exponetial parameter $f\tau$ arrival rate and this is given by $P(A_{nst}) = 1 - \exp(^{-\lambda t})$ is as an identical independent process (iid) random variable with mean arrival rate $1/\lambda$ and mean service $1/\mu$. The packet interarrival times in the next generation networks are described by marginal distributions with heavy tail rather than that of the ex-2 ponential. Aggregate number of packets and bytes in time exhibit correlations over large time scales (i.e., long-range dependece) and self-similar scaling prop-erties. Poisson arrival processes are quite limited in their burstiness, especially when multiplexed to a high degree. Wide-area traffic is much burstier than Poisson models predict, over many time scales. This greater burstiness has im-plications for many aspects of congestion control and traffic performance. The real VBR traffic analysis exhibits persistence with burstiness and nonstation-ary properties (i.e., this will not have a constant mean and variance). Poisson processes which memeoryless process lose their burstiness and flatten out when time scales are changed as shown in Fig. 2 and this model is inappropriate for the real time VBR traffic. The video VBR traffic in the network has both Long Range Depedency element (LRD) and Short Range Dependency (SRD) element so self-similarity model like fractional browian motion or fraction gaussian model are ineffeicent to model VBR traffic in the network .We are using FARIMA an asymptotic self similarity to model the traffic in the network. The three variable in the model can model both SRD and LRD. FARIMA can predit the future traffic from the present and past information. The queue simulation can be done using large derivation.

MATERIALS AND METHODS

Self similarity, long range dependency and heavy tailed distribution: The real time traffic have significant variance (burstiness) on a wide range of time scale.The burstiness in wide range can be represented stastically using self similarity charaterstic (Leland *et al*., 1994; Erramilli *et al*., 1996). The self-similarty commonly used to capture the fractal behavior of traffic model which is a ubiquitous

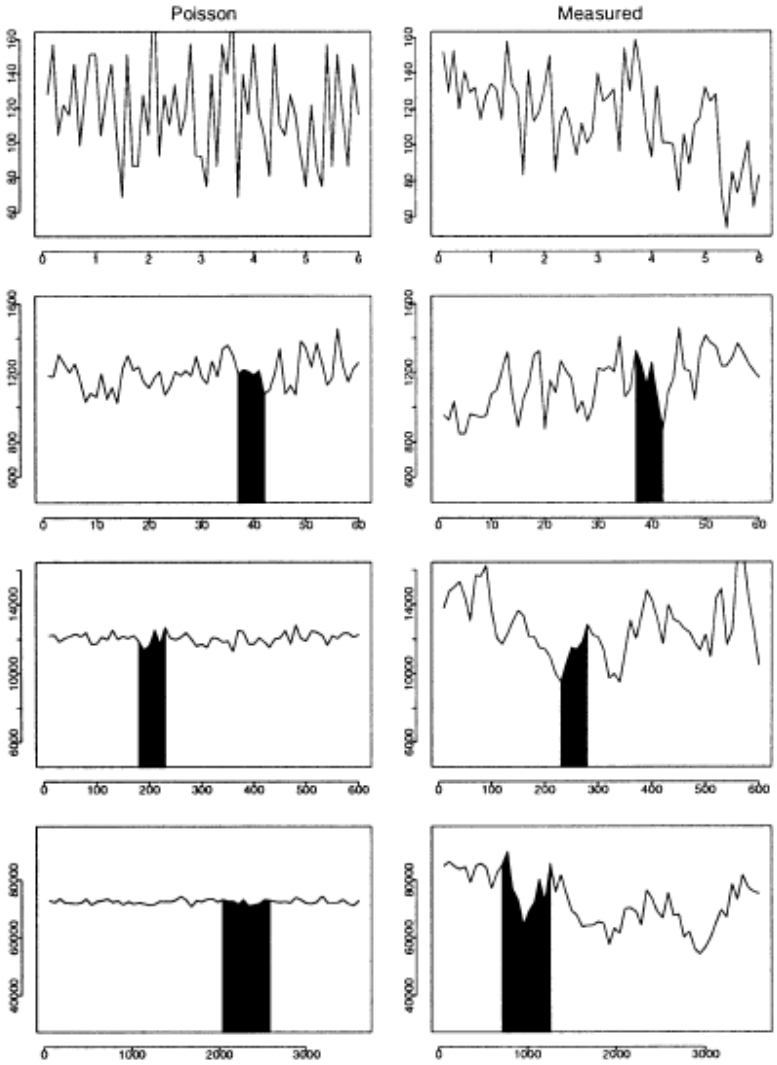

Fig. 2: Example for ethernet traffic for self similarity model

pheneomenon in next gen-eration traffic networks (Adas and Mukherjee, 1995) and the network traffic behaviour can be analysed as a stochastic time series data in distributional sense. For real time multimedia traffic in next generation network shows burstiness in wide range of time scale as shown in Fig. 2. High variability and persistence in the packet arrivial exhibits both correlation and burstiness. This can be represented as self similarity pro-cess. The inter arrival time to the network can be represented as stochastic time series process X(t) with constant mean μ_x, variance σ_x^2 and the autocorelation function (acf) $\rho(k) = E[X_t - \mu](X_{t+k}-\mu]/E[(X_t-\mu^2]$, k = 0, 1, 2. The se-ries X(t) when aggregated over m non overlapped block size $x^m = (X_k^m : k = 1, 2, 3, ...)$ as $P(k) \approx k^{-\beta} L_1(k)(1)$ where $0 < \beta < 1$ and L is slowly varying at infinity as $\underline{\text{Lim}}\, L(tx)/L(x)$ (x) = 1. For an aggregated traffic with m non overlapping block size can be represented as $X_k^{(m)} = 1/m (X_{km-m+1}+...+X_{km})$. The process is exactly (second order) self similar with X^m if the $\rho^m(k) \rightarrow \rho(k)$, $k \geq 0$ for all m = 1,2,...k. Fractional Browian Motion is an example of exactly (second order) self similar model. Asymptotically self similar as $\rho^m(k) \rightarrow \rho(k)$ as $m \rightarrow \infty$and FARIMA time series model is an example of this model. The main features of self similarity is the variance of the sample decrease very slowly, the autocorrelation function decay hyperbolicaly (LRD) rather than exponential, the spectral density obey power law behaviour.defines the property of so-called Long-Range Dependence (LRD). High speed traffic in the NGN shows persistence and slowly decaying correlation and this can be described as long range depe-dency. The

long range dependence or long memory, the autocorrelation function of a stochastic time series decays slower than an exponential in time. The aggre-gated traffic with wide range of time scale exhibit long-range dependence with the acf asymptotically equal to: X (mt) = m^HX(t).

Self similarty along with Long memeory process or long range dependency process can be measured by Hurst parameter. The value of H, 0.5<H>1 mean the process is having high self similarity with long range dependency. Major cause of persistence is due to heavy tail packet arrival traffic in the net-works and this can degrade the queueing performace.So the QoS parameter such as packet delay, lost packet and jitter depend on the queueing performance. The random variable in the bursty traffic will show a heavy tail distribution. P [Z > x] ~ $cx^{-\infty}$ This we can find using the probabilty density function.The self similarity have a hyperbolically decaying covariance function of the following form ρ(k) ≅ k^{2H-2}L(t) as k→∞‡ where L(t) is the function slowly variable at infinity as this will show the process long range dependenc $\underline{Lim}_{t\to 0}$ L(tx)/L = 1‘ the covariance function is nonsummable a‡” the series formed by sequential values of the covariance function diverges as Σ_K ρ(k) = ∞. A covariance stationary process X is called asymptotically (second-order) self-similar with self similarity parameter capturing the persistence phenomena observed empirically in many modern highspeed networks. High speed traffic in the network exhibit perisestence due to high variability and correlation ie the arriving current packet value strongly depends upon the past values of a stochastic process. The empiri-cal marginal distributions are typically not Gaussian, they tend to be skewed to the right and the tail behaviour of the marginal distribution can be accu-rately described using the heavy-tailed distributions. So, the model proposed is the FARIMA with stable innovation. For an infinite variance process the d value depends on H-1/α where α is tail index of heavy tailed distribution. In high speed networks multimedia traffic the packet interarrival time can be described as a marginal distribution with an heavy tailed rather than an expo-nential. The probabilty density function of this traffic is normally skewed to the left (lowest values). Due to burstiness in the traffic, this will show high variabilty and whose activity period follows heavy tailed distribution may be caused by LRD. The Heavy-tailed distributions characterise long-memory processes with strong time-dependence structures that vanish very slowly. The heavy tailed distribution is asymtotically self similar and the this asymtotically self similar x^m = xasm→∞ this define the LRD and has the asymptotic form. Infinite variabilty in the networks shows high burstiness and the major cause of bursti-ness heavy tailed distribution. The $\rho(x) \sim c_p k^{-\alpha}$where α = 0, 1 and H = 1-1/α. A random variable X follows a heavy-tailed distribution (with tail index α):

$$P(X > x) \approx cx^{-a}$$

The infinite variance is given by the tail of the distribution that is given by ∝ and the value is in between 1<α<2. The heavy tailed distribution is stable α and this is sum of Gaussian iid infinte variance random variables generalized by the central limit theory:

$$?(k) \approx k^{-\beta} L_1(k), k \to \infty, 0<\beta<1$$

Several studies shows that the long-range dependence may be caused by heavy tailedness of certain traffic characteristics like file size and web application. Both long range dependency and self similarity are associted with heavy tail behaviour. The heavy-tailed distribution is the root cause of LRD and self-similarity. Different mathematical models are their to represent the self similarity, long range depedency and heavy tailed distribution.

Indroduction to time series model: A time series is a set of observations generated sequentially in time and this can be consider as stochastic process with equispaced N dimesional random variable x_1, x_2, x_3, ..., x_N with a probabilty density function $p(x_1, x_2, x_3, ..., x_N)$ (Basu and Mukherjee, 1996). The process can be either stationary or nonstationary. Stationary process means the marginal distribution of the process does not change with time. We can find stationary in the series using Dickey-Fuller (Basu and Mukherjee, 1996). Mostly VBR traffic signal non-stationary with varying mean and variance. If a time series is nonstationary then the auto correlation function will never cut off nor die quickly but will slowly decay to zero. Time series models like ARMA and ARIMA processes are inherently short-range dependent models, incapable of parsimoniously capturing the persistence phenomena observed empirically in many modern high speed networks ARMA and ARIMA processes have autocorrelation functions which decay geometrically in the lag, namely, (n)$r^n$for some 0 <r<1 as n →α. Traffic in the network can either be stationary or nonstatinonary, stationary the mean value of the process remain in equilibrumi about the a constant mean level we can find stationarty in the series tatking the charatistic equation of the polynomial. Mostly multimedia data will exibit nonstationary behaviour and this we can find by verify

polynomial if the value greater than unit circle then the series is a stationary if we differcate d times the series will become a stationary series and we can fit the series.

RESULTS AND DISCUSSION

Fractional arima time series model: FARIMA is an extention of classic ARMA model and this is flexible to model both short term and long term correlation structure of a series (Leland *et al*., 1994). Both lower frquency componets as well as higher freqency can be fitted using one or two parameters. FARIMA is a class of long memory that can explicitly account for persistence to incorporate the long term correlation in the data. A FARIMA (p, d, q) process X = X(k), k∈Z FARIMA process is a standard given by ARIMA process degree of differencing d being a real number with (0, d, 0) sequence showing the long-range dependence is generated 1/2<d <1/2. For sta-tionarity and invertibility, it is assumed that all roots of $\Phi(p)\Theta|(q) = 0$ are outside the unit circle and |d| <0.5 where p = q = 0 is known as fractionally differenced white noise. Real time multimedia traffic measurements found the co-existence of both long range and short range dependency The FARIMA(p, d, q) time series X = X(k), k ∈Z is define as:

$$\phi(B)\nabla^{d}X_t = \theta(B)a_n\ (-0.5, 0.5) \quad (1)$$

Where:

$$\phi(B) = \phi_1(B) + \phi_2(B)^2 + \ldots + \phi_p(B)^p$$

$$\phi|(B) = \theta_1(B) + \theta_2(B)^q + \ldots + \theta_q(B)^q$$

$$\nabla^{d}_{x_t} = (1 - B)x_t = x_t - x_{t-1}$$

$$\nabla^{d}_{x_t} = (1-B)^{d}x_t$$

$$\rho(K) = \frac{d(1+d)\ldots(K-1+d)}{(1-d)(2-d)\ldots(K-d)}, k=1,2\ldots$$

FARIMA process have correlation ρ_k which behave asymptotically as k^{2d-1}. Parameter estimation is the first step in fitting an FARIMA For a series if the value of 0 < d < 0.5 the series is nonstationary this will not have a common mean and variance. FARIMA (0, d, 0) process is called fractionally differenced noise and this can represent the wide range of LRD we have to make the series stationary so the series become ARMA process and we can fit the series using the ARMA model. The autocorrelation function in FARIMA is given by:

Table 1: Parameter estimated for FARIMA

Model	ϕ	θ	d
FARIMA (2, 0.46.1)	(0.88, 0.01)	0.46	0.77

$$\rho_k = \Gamma(1 - d)\Gamma(k + d)/\Gamma(d)\Gamma(k + 1 - d)$$

If d>0, the hyperbolic decay of the correlation as lag increases, this the hyperbolic decay will show asymptotically self similar:

$$\phi(B)(1 - B)^{d}x_t$$

FARIMA (p,d,q) where p, q ∈ N ∪ {0} and d ∈ R For long range dependency ,p = q = 0 Let d ∈ R, $\{X_t\}_{t\in Z}$ is FARIMA where $\Delta^{d}X_t = \in_t \Phi(B)\Delta^{d}X_t = \Theta(B)a_t$ estimate $\Delta^{d}x_t = a_t$ where a_t is an iid with mean 0 and variance $\sigma^2{}_a$ LRD in a traffic model is by simply allowing the input into a queue to have heavy-tailed characteristics FARIMA is two step estimatioican procedure d value is estimated usingR/S method and parameter for p, q are estimated using normal ARMA fitting method

Identification and estimation of farima: For the time series both short term as well as the long term dependence are observed. Hyperbolic decay of the correlation function. Select parameter p and q using goodness of test (AIC value should be minimum) estimate the parameter values Φ(B), Θ(B) forp, qvalue using maximum likelihood algorithm (Basu and Mukherjee, 1996). Calculate value for d using R/S method (11). FARIMA (2, 0.46, 1) with ϕ_1, ϕ_2 and θ_1. Estimate the parameters of the FARIMA model Table 1.

Forecasting and bandwidth allocation: In the VBR traffic dynamic bandwith allocation is needed. The traffic in the network has either mean value or peak value. The bandwidth in the network should be a value between this peak and mean value. It is difficult to estimate the effective bandwidth in the network. Using FARIMA model, we can predict the resource needed for the incoming traffic and allocate that resource to the particular application. For this, we have to buffer the mean square error value with a biase value.If the traffic is SRD their is no significant effect in the queue only a white noise will buffered (Garrett and Willinger, 1994; Adas and Mukherjee, 1995).

Procedure for forecasting is as follows calculate the residuals from the fitted model and obtain the 100uth percentile of the distribution of the residuals. Call it ε_u obtain the minimum mean square error forecast for X_{t+1}, based on the FARIMA model. Call it $X_t(1)$. The forecast for the 100uth percentile for X_{t+1} is $X_t(1)+\varepsilon u$.

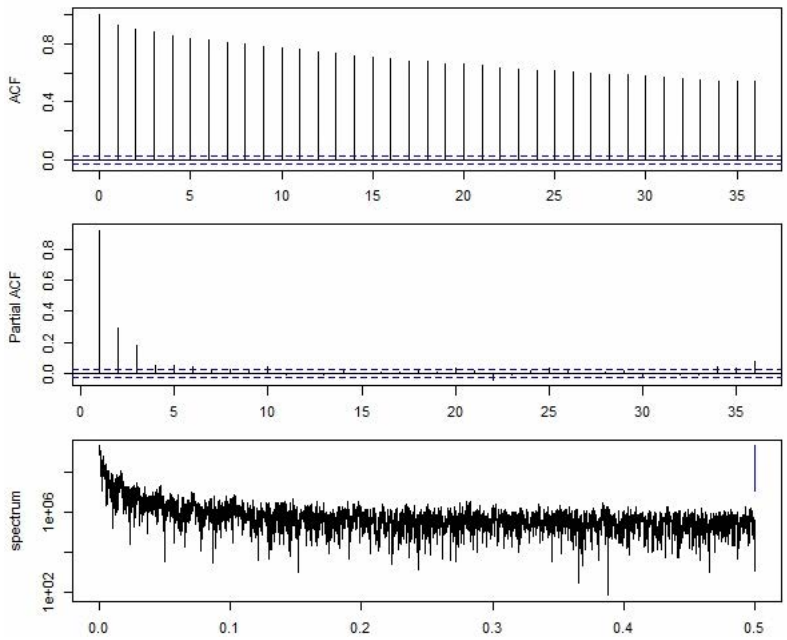

Fig. 3: Autocorrelation function and spectral density for VBR traffic

$$X_u(t) = \widehat{X}(t) + \varepsilon_\mu$$

where low bandwidth is needed and in other very high bandwidth is needed to satisfy the QoS.Predicting the future behaviour of the traffic for the present and past traffic and we can allocating the bandwidth for the incoming traffic. A dynamic bandwidth allocation strategy predict bandwidth requirement for the future traffic using linear prediction with minimizes the mean square error. Reserving bandwidth equal to predicted value only prediction error should be buffered. The predicted value will be white noise or short memory. FARIMA parameters are estimated and used to predict traffic in next generation net-works. The FARIMA parameters are estimated from the historical traffic data. Estmate the nonlinear parameter which minimize the to predict the traffic forecast h step ahead as shown in Fig. 3. Predicting the behaviour of the traffic form the present holt-winters filtering value and allocating buffer space and bandwidth:

$$X_u(t) = \widehat{X}(t) + \varepsilon_\mu$$

The real time traffic behaves bursty (e.g. its variance varies over time). So linear prediction of time series will not give a accurate result. We can add a biase value with this so both peak and mean value can be represented by this prediction. So bandwidth can dynamicaly allocated. The k step ahead prediction can be defined as X_{t+K} mean square error $E = ((X_{N+h)} - X_h)\ X_h = E(X_{N+h)}/X_N, X_{N-1})$. Prediction of traffic and dynamic allocation minimize the cell loss rate and delay if the predicted value is more, extra bits can be transmitted from the buffer. Since, the errors resemble noise or at most short memory then smaller buffers, less delays and higher utilization are expected when compared to traditional model (Fig. 4) (Garrett and Willinger, 1994). The tail behavior of queues for the buffered traffic decays exponentially. The effective bandwidth for queue simulation depends on:

$$P_r(Q > B) \approx e^{-\alpha B^{2-2H}}$$

Traffic model will evaluvate the behaviour of the aggregated traffic and pre-dict the queuing performance so that resource can be allocated. If the traffic is overloaded in the link this will result in congestion and packet loss and this will degrade the strigent QoS parameter. To overcome this call admission con-trol should establish neither on peak nor average bandwidth allocation using effective bandwidth. This effective bandwidth is a function of traffic charateristic. The input to the buffer is the error series which is short white noise and will not have the persistence. The analysis of Queueing system work load process is assumed to have an associate large derivation principle (Fig. 5).

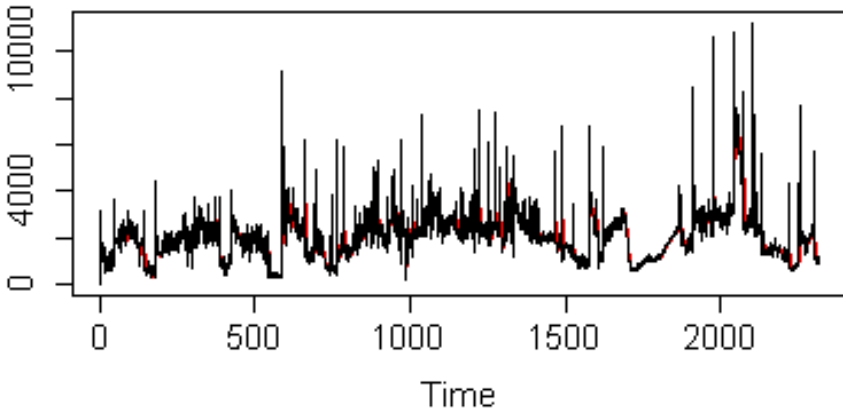

Fig. 4: Obsered trace and fitted forecast (Holt-Winters filtering)

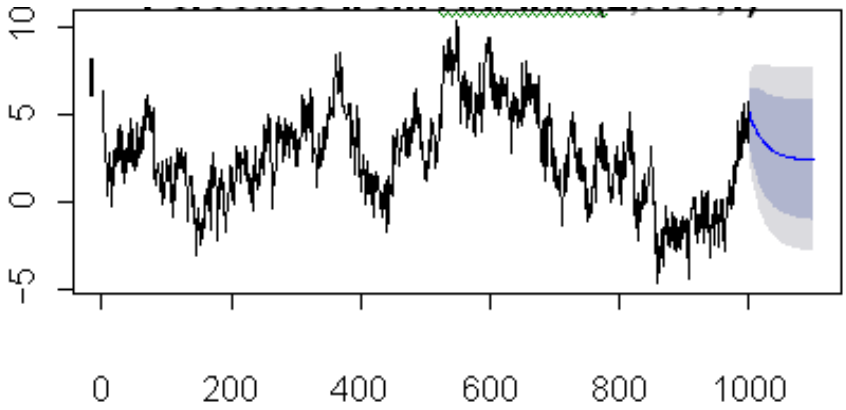

Fig. 5: Forecast h step ahead (Forecasts from ARFIMA (2,0.36,1))

CONCLUSION

Real time VBR bursty network traffic is complex in next generation network as it exhibits strong dependence and self-similarity, models of time series such as Poisson and Markov processes are not appropriate for its modeling.Maintaining high utilization of the bandwidth is the objective for efficient traffic manage-ment which include call admision control policing, scheduling, buffer manage-ment and congestion control etc. The high variable and highly correlated VBR traffic in the network can be modeled with self similarity models like FARIMA or FGN. Mostly VBR traffic in the network exhibits both short range and long range dependency along with heavy tailed distribution can be modeled with FARIMA time series model with parameter (p, d, q). FARIMA time series traf-fic model can predict traffic and allocate bandwidth dynamically. This model is flexible enough to parsimoniously capture the statistical property of traffic can allocate bandwidth on demand. The tail behaviour of the queue can be analysed with traffic model. Performance in the network can be improved with the analysis of QoS parameter like packet loss, delay and variances.

REFERENCES

Adas, A. and A. Mukherjee, 1995. On resource management and QoS guarantees for long range dependent traffic. Proceedings of the IEEE 14th Annual Joint Conference on Computer and Communications Societies Bringing Information to People, April 2-6, 1995, IEEE, New York, USA., ISBN:0-8186-6990-X, pp: 779-787.

Basu, S. and A. Mukherjee, 1996. Time Series Models for Internet Trac. San Francisco Publisher House, California, USA.,.

Erramilli, A., O. Narayan and W. Willinger, 1996. Experimental queueing analysis with long-range dependent packet traffic. IEEE. ACM. Trans. Networking, 4: 209-223.

Garret, M. and W. Willinger, 1994. Analysis, modeling and generation of self-similar VBR video traffic. Proceedings of the ACM/SIGCOMM'94, October 1994, ACM, London, UK., pp: 269-280.

Leland, W.E., M.S. Taqqu, W. Willinger and D.V. Wilson, 1994. On the self- similar nature of ethernet trac (extende version). IEEE. ACM. Trans. Networking, 2: 1-15.

Paxon, V. and S. Floyd, 1995. Wide-area trac: The failure of Poisson modeling. ACM. IEEE. Trans. Networking, 3: 226-244.