Modeling and Predicting Long-range Dependent Traffic with FARIMA Processes

Fei Xue
Dept. of Information Engineering
The Chinese University of Hong Kong
Shatin, New Territory, Hong Kong

Abstract

FARIMA processes provide a unified approach to characterize both short-range and long-range dependence in network traffic. In this paper, we study how to construct FARIMA models for measured traffic and propose a three-step modeling procedure. Experiment results on actual traffic verify our suggested procedure and support that FARIMA model can capture the property of real traffic in a flexible and parsimonious way. Based on the linear prediction with FARIMA models, we propose one traffic predictor with compensation function, which has potential application in dynamic resource allocation.

1. Introduction

Recent studies of high quality measurements have shown that traffic in high-speed network exhibits self-similarity, i.e. long-range dependence [1,2]. Long-range dependence is characterized by an autocorrelation function that decays hyperbolically as the lag increases and produces many new features, which are drastically difference from more tradictional short-range dependent models such as Markovian and ARMA process. Two theoretical models which exhibits long-range dependence are Fractional Gaussion Noise (FGN) processes[3] and fractional autoregressive integrated moving average (FARIMA) processes[4,5].

Formulating effective traffic model is considered essential for network design, traffic management and performance evaluation. Also, accurate prediction of traffic levels will be crucial to dynamic bandwidth allocation, traffic control and network planning. Therefore, a desirable traffic model should be (1)flexible enough to parsimoniously capture the statistical properties in the traffic, (2) mathematically tractable for queueing analysis and simulation, and (3) feasible to be used for prediction needed in the context of traffic control.

FARIMA process has proved useful in the analysis of time series with long-range dependence and been extensively used in various applications recent years[6,7,8]. It is more flexible and parsimonious than FGN process with regard to the simultaneous modeling of both long-range and short-range dependence in the data. Due to its origins in time series analysis, FARIMA model can generate synthetic data easily and are amenable to be used in traffic prediction. For the detailed discuss on the application of FARIMA processes to self-similar traffic research see [9].

This paper focuses on modeling and prediction long-range dependent traffic with FARIMA processes. It is organized as follows. In Section 2, we summary the basic mathematics of FARIMA processes. Then we study how to fit a FARIMA model to the data in Section 3, our experiment results are presented and discussed in Section 4. In Section 5, we study the traffic prediction with FARIMA processes and propose a predictor with compensation function. Section 6 is the concluding remarks.

2. FARIMA processes

FARIMA processes are the generalizations of standard ARIMA (p,d,q) processes when the degree of differencing d is allowed to be fraction in the range (-0.5, 0.5).

The FARIMA(p,d,q) process $\{X_t: t=\ldots-1,0,1,\ldots\}$, where p and q are non-negative integers and $d \in (-0.5, 0.5)$, is described by the difference equation

$$\Phi(B)\Delta^d X_t = \Theta(B)a_t \tag{2-1}$$

where

$$\Phi(B) = 1-\phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \tag{2-2}$$

and

$$\Theta(B) = 1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \tag{2-3}$$

are polynomials of order p and q, respectively, in the backward-shift operator B. $\Delta^d = (1-B)^d$ is the fractional differencing operator and $\{a_t: t=\ldots,-1,0,1\ldots\}$ is a white noise process with constant variance σ^2.

When $\Phi(z)$ and $\Theta(z)$ have all their roots outside the unit circle and no common roots, the process $\{X_t\}$ is stationary and invertible. In the case $0<d<1/2$, long-range dependence occurs and the process exhibits self-similarity with Hurst parameter $H = d + 0.5$.

FARIMA(0,d,0) process, also called fractionally differenced noise, is the simplest and most fundamental form of the FARIMA processes. The property of FARIMA(0,d,0) process is similar to FGN process, but from modeling purposes, the FARIMA(0,d,0) has some advantages over FGN because it can be extended to encompass a wide range of short-term behavior of a time series.

FARIMA(p,d,q) process $\{X_t\}$can be regarded as an ARMA(p,q) process driven by fractionally differenced noise $\{Y_t\}$. From (2-1), we can obtain an explicit expression

$$X_t = [\Theta(B)/\Phi(B)]Y_t \tag{2-5}$$

where

$$Y_t = \Delta^{-d} a_t . \tag{2-6}$$

So, FARIMA(p,d,q) can be expressed by its two basic forms: ARMA(p,q) process and FARIMA(0,d,0) process, where the former describes the short-range behavior and the latter describes the long-range dependence.

3. Constructing a FARIMA model for time series

The applicability of FARIMA models is dependent on the ease with which they may be fitted to an observed time series. Selection and estimation of FARIMA model is more difficult than that of standard ARMA models. When both long- and short range correlation structures are present in the data, their behavior is hard to distinguish. Additionally, likelihood estimation techniques for FARIMA model present computational problems and often result in significant finite sample biases. However, methods for constructing ARMA models are now well established. FARIMA model is the extended version of ARMA model in nature, so it will be helpful to solve model building problems of FARIMA by applying the effective methods for ARMA processes.

We transfer the FARIMA problem to the ARMA problem by splitting FARIMA model into its “fractional differenced” and “ARMA” part. For a FARIMA process $\{X_t\}$, we can obtain from (2-1)

$$W_t = \Delta^d X_t \tag{3-1}$$

where

$$W_t = [\Theta(B)/\Phi(B)]a_t \tag{3-2}$$

So, if the differencing parameter d can be obtained in advance and the fractional differencing operator Δ^d can be implemented in practice, W_t, the fractionally differnced X_t, can be evaluated as an ARMA model.

The following is our proposed procedure to fit an FARIMA model to trace $\{X_t\}$:

Step 1. estimating fractional differencing parameter *d*

Hurst parameter H can be estimated by several methods such as Variance-time analysis, R/S analysis and periodogram-based analysis[9]. As remarked in Section 2, the strength of long-range dependence measured by d is the same as that by H upon $H = d + 0.5$. Then we can estimate parameter d from this relationship.

Step 2. fractional differencing on {X_t}

The calculation of {W_t} from {X_t} involves a finite approximation to the infinite sum in the definition of fractional differencing operator. The exact fractional differencing operator for the {X_t} series with mean value u is

$$\Delta^d (X_t - u) = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k (X_t - u) = \sum_{j=0}^{\infty} \varpi_j (X_{t-j} - u) \tag{3-3}$$

where

$$\varpi_j = \frac{(-1)^j \Gamma(1+d)}{\Gamma(1+j)\Gamma(1+d-j)}. \tag{3-4}$$

This involves the unobserved quantities X_0, X_{-1}, X_{-2},…. If these quantities are replaced by their mean value u, the operator can be calculated from the data. But the replacement of X_t, $t \le 0$ by u is unrealistic for a long-range dependent series, since if the X_t for small position t have a tendency to be greater than u, then it is likely that the X_t for small negative t would also be greater than u if they could be observed. So a compromise method is to replace X_t by u only for $t \le -M$, where M is chosen large enough for the assumption $X_t = u$ for $t \le -M$ to be reasonable. This yield the operator

$$\Delta_M^d (X_t - u) = \sum_{j=0}^{t+M-1} \varpi_j (X_{t-j} - u) \tag{3-5}$$

In practice, a high-order AR(p) model for {X_t} is built as auxiliary model which is used to estimate the presample values by "backcasting" method discussed in [10].

After using the operator Δ_M^d with estimated parameter d on {X_t}, we can obtain {W_t}.

Step 3. model identification and parameter estimation for FARIMA model

Now we can analyze {W_t} as an ARMA(p,q) model using conventional method in order to obtain the unknown order p, q. We select the best one from the candidate models based on the information criterions such as AIC (Akaike Information Criterion) and BIC(Bayesian Information Criterion) . The residuals from the ARMA(p,q) model fitted to { W_t } are used for diagnostic checking selected model by available methods such as Portmeanteau test[10].

Once the order p and q of are selected, we can estimate all the parameters of the desired FARIMA model by maximum likelihood algorithm proposed in [11].

4. Modeling measured traffic with FARIMA processes

In order to test the feasibility and performance of FARIMA models, we have applied the above procedure to obtain FARIMA models for two measured traffic traces (pAug.TL and pOct.TL) from Bellcore Lab and two traces (C1003 and C1008) from CERNET of China[12].

4.1 fitting FARIMA models for real traffic

In the experiment, we preprocessed measured traces to get four data sets. Each data set has 2000 samples, and each sample represents the arrival packet number during 0.01 second interval.

Figure 1 and Figure 2 show the autocorrelation function and spectral density of the trace "pAug.TL" respectively. The autocorrelation function does not become negligible even at a high lag and the spectrum density is very large at the frequency near zero. All these characterizations indicate the traffic is long-range dependent. The estimated Hurst parameter of the trace is 0.8, namely d = 0.3, which is estimated by the periodogram-based analysis. An auxiliary AR(8) model is built for fractional differencing. The autocorrelation function and spectral density of the fractional differenced series are shown in Figure 3 and 4 respectively. From these results it can be concluded that fractional differencing operator Δ_M^d defined in (3-5) has taken effect on removing the long-range dependent behavior of traffic trace. Then we fitted an ARMA(p,q) model to the fractional differenced series and chose ARMA(1,1) as the best one. The fitted model was also subjected to diagnostic checks and test of goodness of fit. Finally, a FARIMA(1, 0.3, 1) model was fitting to the "pAug.TL" trace by the maximum likelihood algorithm.

The experimental results on modeling these traffic traces with FARIMA processes are briefly summarized in Table 1.

4.2 discuss on the experimental results

Further experiments have been made to compare FARIMA model with other models such as AR and FGN in description ability. Table 2 presents the fitted models for pAug.TL.

Figures 5 to Figure 7 show the autocorrelation functions of the trace pAug.TL and the fitted models in each case. As seen from Figure 5, the correlation behavior of trace sample and FAIMA model are very similar and the difference is insignificant from the statistical view. But FGN and AR model can't have the same good fit as FARIMA model in the respect of autocorrelation function. We define the mean squared error (MSE) of the autocorrelation function as

$$MSE(m) = \frac{\sum_{i=1}^{n} (ACF_m(i) - ACF_s(i))^2}{n} \tag{4-1}$$

where $ACF_{\mathrm{m}}(i)$ and $ACF_{\mathrm{s}}(i)$ indicate the autocorrelation at the ith lag of model m and the actural trace respectively. From the MSE vales listed in Table 3, we can find that FARIMA is the best one among them. In addition, FARIMA model has less parameter than the AR model for the same traffic trace. So all these results confirm our previous views that FARIMA model can capture the long-range and short-range correlation structures in a flexible and parsimonious way.

An interesting finding in our experiments is that the order p, q of the fitted FARIMA models will be constant for some long time. As an example, we studied five consecutive traffic data sets aug1-aug5 which is obtained from trace pAug.TL. Each data set has 2000 samples and each sample represents the packet arrival during 0.01 second. So the whole duration of the five data sets is 100 second. We fitted FARIMA model to each data set. The results are given in Table 4. The orders of the five FARIMA models are same, i.e., p=1 and q=1. We also have the similar observation on other measured traces that the p, q values do not change during a long periods. This finding may guide us to simplify model fitting procedure and reduce the computational complexity for FARIMA modeling.

In Step 2 of our proposed modeling procedure, an auxiliary model is used to estimate the presample values. For the consecutive traffic trace, the presample values are just the previous traffic data. So we can use the previous data instead of "backcasting" values for the operator Δ_M^d , which will make fractional differencing more easily implemented.

5. Prediction long-range dependent traffic with FARIMA process

Traffic predicting can be useful in dynamic resource allocation, traffic management and control schemes. The tradition models often ignore the correlation and burstiness properties of network traffic, so prediction based on them can't be accurate and efficient enough. In contrast, FARIMA processes have great promise to predict traffic behavior due to the ability to capture both short-range and long-range correlation structures simultaneously.

5.1 linear predictor of an FARIMA process

Let $\{X_t\}$ be the causal invertible FARIMA*(p,d,q)* process, we can write

$$X_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} \tag{5-1}$$

and

$$a_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \tag{5-2}$$

where $\sum_{j=0}^{\infty} \psi_j B^j = \Theta(B)\Phi^{-1}(B)\Delta^{-d}$ and $\sum_{j=0}^{\infty} \pi_j B^j = \Phi(B)\Theta^{-1}(B)\Delta^{d}$.

From the theorems on linear prediction [13], the *h*-step-ahead predictor $\hat{X}_t(h)$ of an FARIMA process is

$$\hat{X}_t(h) = -\sum_{j=1}^{\infty} \pi_j \hat{X}_t(h-j) \tag{5-3}$$

and

$$\hat{X}_t(h) = \sum_{j=h}^{\infty} \psi_j a_{t+h-j} \tag{5-4}$$

with the mean squared error

$$\hat{\sigma}_t^2(h) = E(X_{t+h} - \hat{X}_t(h))^2 = \sigma^2 \sum_{j=0}^{h-1} \psi_j^2 \ . \tag{5-5}$$

Equation (5.3) is the most convenient one for calculation of $\hat{X}_t(h)$. It can be solved recursively for *h*=1,2,3,…, using the conditions $\hat{X}_t(n) = X_{t+n}$ for *n*<0. This linear predictor converges much more slowly to the mean value than that based on ARMA process.

5.2. traffic prediction with compensation function

In the prediction-based resource allocation schemes, if the predicted value is less than the observed one, some packets will be dropped and this effect will downgrade network performance seriously. To alleviate this problem, a method was suggested based on ARMA models [14]. Here we extend this idea to FARIMA models and proposed a predictor with compensation function.

Let $\{a_t\}$ be the residuals from the fitted FARIMA model, the *h*-step-ahead compensation predictor $\hat{X}_t^u(h)$ with parameter *u* is given by

$$\hat{X}_t^u(h) = \hat{X}_t(h) + \xi_u \tag{5-6}$$

where $\mathrm{P}[a_t \le \xi_u] = u$, $0.5 \le u < 1$. $\hat{X}_t(h)$ is the linear predictor defined in (5-3) and ξ_u is the 100*u*th percentile of the distribution of $\{a_t\}$.

Our experimental results indicate that the residuals $\{a_t\}$ approximately follow the zero-mean Gaussian distribution. As an example, Figure 8 shows the probability density of the residuals from fitted FARIMA models on trace pAug.TL. This finding is in good agreement with the views expressed in [15]. Under this condition, we can obtain that

$$\xi_u = z_u \times \sigma \tag{5-7}$$

where z_u is the uth quantile of a standard Gaussian distribution and σ is the standard error of the $\{a_t\}$.

From the definition of this predictor, we can expect that for a given u, about 100u% forecasts will be over the observed series, and 100(1-u)% forecasts will be below the observed ones. So the choice of parameter u will depend on the quality of service of various applications. For the loss-sensitive applications the parameter u should be very close to one. But if the cost of losing packets is not high, the u can be smaller. This predictor may have potential application in dynamic bandwidth and buffer allocation strategies for connectionless service—especially if they are required to reserve bandwidth in advance.

6. Conclusion

In this paper, we studied how to model and predict long-range dependent traffic with FARIMA processes. The key of our three-step modeling procedure is to apply the well-established methods for ARMA to solve model building problem of FARIMA. The experiment on four measured traffic traces proved that it is more easily implemented in practice. This modeling procedure can also be used for analysis other long-range dependent time series in a wide range of fields. Our experiment results support that FARIMA process is capable of simultaneous modeling long-range and short-range behavior of network traffic. We shown that it is possible to simplify the fitting procedure and reduce the computational complexity for developing a model. Our traffic predictor is based on the linear prediction with FARIMA and has potential application in dynamic resource allocation.

Future work includes applying FARIMA model for measurement-based network control.

REFERENCE

[1]. W.Leland, et al., “On the self-similar nature of Ethernet traffic (extended version),” *IEEE/ACM Transactions on Networking* , 1994, 2(1):1-15.

[2]. Jan Beran, Robert Sherman, et al. “Long-range dependence in Variable Bit Rate Video Traffic,” *IEEE Transactions on Communications* 43(2/3/4): 1566-1579 ,1995

[3]. B.Mandelbrot and J.Van Ness, ”Fractional Brownian Motions, Fractional Noises and Applications.” *SIAM Review*, 10:422-437, 1968

[4]. C.W.J.Granger and R.Joyeux, “An introduction to long-memory time series models and fractional differencing”, *Journal of .Time Series Analysis*. 1: 159-29, 1980.

[5]. J.R.M.Hosking, “Fractional differencing,” *Biometrika*, 68: 165-176, 1981.

[6]. D.J.Noakes, et al. “Forecasting annual geophysical time series time series”, *International Journal of Forecasting*, 4:103-115,1988

[7]. F.X.Diebold and G.D,Rudebush, ”Long memory and persistence in aggregate output”, *Journal of Monetary Economics*, 24:189-209,1989

[8]. Y.W.Cheung, “Long-memory in foreign exchange rate”, *Journal of Business and Economic Statistics*, 11:93-101, 1993

[9]. F.Xue, “Modeling Analysis and Performance Evaluation for self-similar traffic.” Ph.D. Dissertation, Tianjin University, June 1998

[10]. G.Box, G. Jenkins “Time series: forecasting and control” revised ed., San Francisco: Holden-day, 1976.

[11]. Xu J.Haslett, et al., “Space-time modeling with long-memory dependence: Assessing Ireland’s Wind Power Resource” *Applied Statistics* 38(1): 1-50 1989.

[12]. F.Xue, Y.Shu, et al. “Analysis of Self-similarity and Queueing Performance for CERNET Traffic”, Proc.4th CERNET Academic Conf. pp.88-93, Xi’an, China, 1997

[13]. P.Brockwell, R.Davis, Time Series: Theory and Methods, 2nd ed. New York, Springer Verlag, 1991

[14]. S.Basu and A.Mukherjee, “Time Series Models for Internet Traffic”, *IEEE INFOCOM96*, San Francisco, California, 1996

[15]. G.Addie, “Traffic Will be more Gaussian in Future”, *ATNAC’96*, 1996

Table 1 Fitted FARIMA models of CERNET and Bellcore traces

Trace	FARIMA(p,d,q)		
	p ($\phi_1, \phi_2, \ldots, \phi_p$)	d	q ($\theta_1, \theta_2, \ldots, \theta_q$)
Paug.TL	1(-0.1714523)	0.2949911	1(-0.373878)
Poct.TL	1(-0.1511854)	0.2520007	1(-0.251823)
C1003	1 (0.3040983)	0.1978055	0
C1008	1(-0.04399369)	0.183442	0

Table 2: Fitted AR, FGN model of pAug.TL

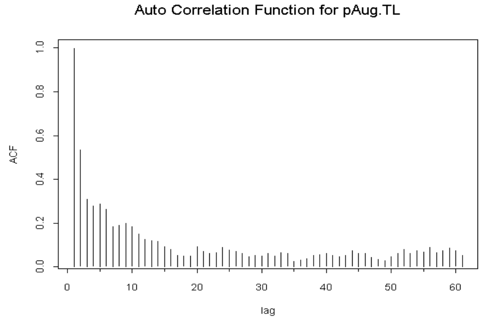
Model	parameters
AR(8)	ϕ_1=0.4840741, ϕ_2= -0.0335644, ϕ_3=0.0692946 ϕ_4=0.0813773, ϕ_5=0.0663615, ϕ_6=-0.0509957 ϕ_7=0.0478156, ϕ_8=0.0471420
FGN	H=0.8, Mean=3.8995, Variance=7.35707

Table 3: MSE values of fitted FARIMA, AR, FGN models of trace pAug.TL

Model	MSE value
FARIMA(1,d,1)	0.0006031
AR(8)	0.0026151
FGN	0.0398209

Table 4: Fitted FARIMA models of five data sets of trace pAug.TL

Data set	ϕ_1	θ_1	d
Aug1	-0.1714523	-0.373878	0.2949911
Aug2	-0.1143173	-0.305762	0.2783244
Aug3	-0.1853628	-0.329502	0.3515382
Aug4	-0.1763141	-0.306262	0.2813467
Aug5	-0.1599587	-0.280055	0.2995875



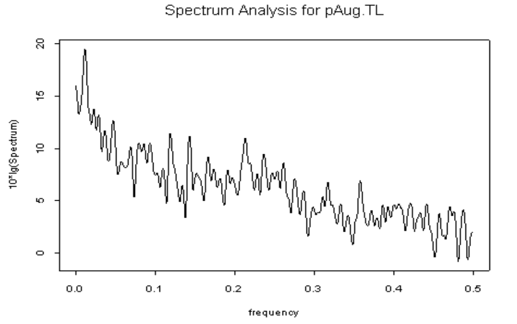
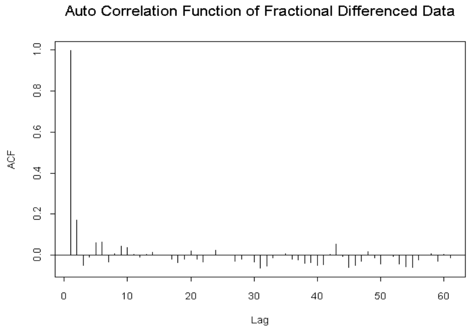


Figure 1: Autocorrelation function of trace pAug.TL **Figure 2: Spectrum of trace pAug.TL**



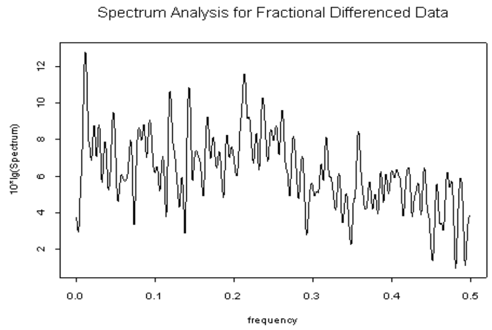
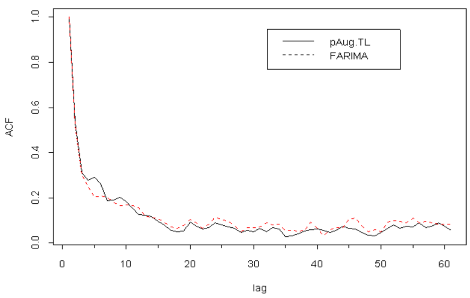


Figure 3: Autocorrelation function of trace pAug.TL fractional differencing

Figure 4: Spectrum of trace pAug.TL after after fractional differencing



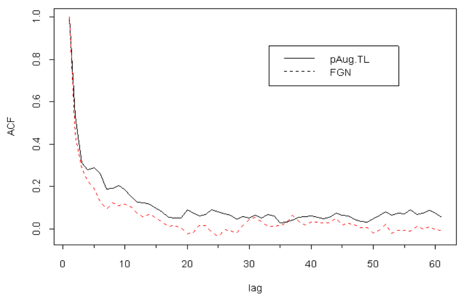
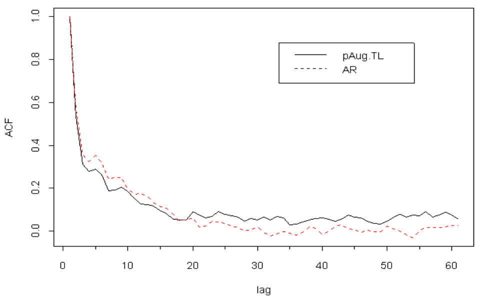


Figure 5: Autocorrelation functions of trace p**A**ug.TL and its FARIMA model

Figure 6: Autocorrelation functions of trace pAug.TL and its FGN model



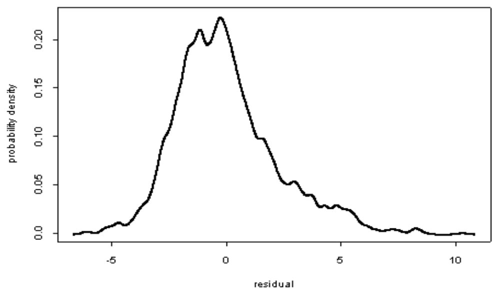


Figure 7: Autocorrelation functions of trace pAug.TL and its AR model

Figure 8: the probability density function of the residual from fitted FARIMA of pAug.TL