

Modeling Video Traffic in The Wavelet Domain

Sheng Ma and Chuanyi Ji

Department of Electrical, Computer, and Systems Engineering

Rensselaer Polytechnic Institute, Troy, NY 12180

e-mail: shengm@ecse.rpi.edu, chuanyi@ecse.rpi.edu

Tel: (518)276-8424, (518)276-6534

Abstract— A significant discovery from this work is that although video traffic has complicated short- and long-range dependence in the time domain, the corresponding wavelet coefficients are no longer long-range dependent in the wavelet domain. Therefore, a “short-range” dependent process can be used to model video traffic in the wavelet domain. In this work, we develop such wavelet models for VBR video traffic.

The strength of the developed wavelet models includes (1) it provides a unified approach to model both long-range and short-range dependence in video traffic simultaneously, (2) it has the ability to reduce the temporal dependence so significantly that the wavelet coefficients can be modeled by either independent or Markov models, and (3) the model results in a computationally efficient method on generating high quality video traffic.

Key words: wavelet, long-range dependence, short-range dependence, traffic modeling, VBR video traffic.

Topics: video networking, B-ISDN and ATM, admission control.

I. INTRODUCTION

Since VBR compressed video traffic is expected to be one of the main loading components in future B-ISDN and wireless networks, accurate modeling of the VBR traffic will be crucial to many important applications such as controlling the Quality of Service, effectively allocating network resources and designing buffer/capacity of networks. Numerous studies have been conducted on traffic modeling and performance analysis, see for example [9][6][12][11][21] and references therein.

One of the significant statistical properties of VBR video traffic has been found to be the co-existence of the so-called long-range dependence (LRD) and the short-range dependence (SRD) in the video trace [2][10][6]. Roughly speaking, this means that the auto-correlation function of the video traffic behaves similarly to that of long-range dependent processes such as Fractional Gaussian Noise process[19] at the large lags, and to that of short-range dependent processes such as DAR processes[6] at the small lags. The long-range and the short-range dependence embedded in video traffic results from scene changes, and suggests a complex behavior in the time domain[14]. This complex temporal behavior makes accurate modeling of video traffic a challenging task. In other words, using either a long-range dependent or a short-range dependent process alone would not do a good job on modeling the video traffic.

Ideally, a good traffic model needs to be (a) accurate enough to characterize pertinent statistical properties in the traffic, (b) computationally efficient, and (c) feasible to be used for the analysis needed for network

design. The existing models which have been developed to model both the long-range and the short-range dependence include FARIMA models[10][11], Transform-Expand-Sample (TES) modeling [18], scene-based models[12] and the Markov Modulated Processes[1]. A common feature of all these methods is that they model both LRD and SRD in the time domain. Among these methods, the scene-based modeling[12] and the Markov Modulated models[1] provide a physically interpretable model to include both the long-range and the short-range dependence. However, due to the dynamic and stochastic nature of the video traffic, it is difficult to accurately define and segment video traffic into different states of a Markov model. TES model is fast but too complicated to be used for analysis. The rest of the methods all suffer the computational complexity too high to be used for generating a large volume of synthesized video traffic[10]. A more computationally efficient method based on Fast Fourier Transform has been proposed [20] to model Ethernet traffic in the Frequency domain. Another method based on Markov models has been proposed to model the frequency components of video traffic[15]. Both methods suggest that interesting properties of either Ethernet or video traffic could be investigated in the Frequency domain. However, none of the methods are yet able to capture the long-range and the short-range dependence simultaneously,

Therefore the question remains open on how to develop a computationally efficient model which can capture both the long-range and short-range dependence in the video traffic. In this work, we will tackle this problem by developing a new method based on wavelets. Instead of modeling the video traffic directly in the time-domain, we model the statistical properties of wavelet coefficients in the wavelet domain.

Why do we choose wavelets? It has been shown in [13][23][8] that wavelets can provide compact representations for a special class of long-range dependent processes, the Fractional Gaussian Noise (FGN) processes. This is because the self-similar (deterministic) structure of wavelet bases naturally matches the (statistical) self-similar structure of the long range dependent processes. Then the wavelet coefficients can be modeled by simple statistics corresponding to the “short-range dependence” alone in the wavelet domain. In this work, we will show that a simple wavelet model based on independence assumptions is capable of capturing both the long-range dependence for more general LRD processes and the short-range dependence as

well, and thus provides a parsimonious and unified model to capture both the long-range and the short-range dependence in video traffic. Furthermore, since computational complexity of wavelet transforms and inverse transforms are in the order of N , our wavelet models can rapidly generate synthesized video traffic of length N with a computational complexity $O(N)$, and thereby provide one of the most efficient methods to synthesize high quality video traffic.

In this paper, we will first investigate why it is more efficient to model the video traffic in the wavelet domain, and what statistical properties of wavelet coefficients are pertinent to capture the long-range and the short-range dependence. We will then develop wavelet models based on these statistical properties. The method will be tested extensively using both well-known processes, which are either long-range or short-range dependent, and three real video traces.

Theoretical analysis on the buffer loss probability will be carried out for the FGN work load when the work load is a Fractional Gaussian Noise process. We will show that the buffer loss probability of the independent wavelet model only differs from that of FGN by a negligible amount. Then the analysis suggests that wavelet coefficients can be well-modeled as independent random variables in the wavelet domain.

The format of the paper is as follows. In Section 2, we will provide background knowledge. In Section 3, we will investigate statistical properties of wavelet coefficients, and develop our algorithm to generate the wavelet model for the long-range and the short-range dependent processes, respectively. In Section 4, we will provide statistical properties of wavelet coefficients for the real video traffic, and derive our algorithms based on these properties. We will give theoretical analysis on the buffer loss probability using the wavelet model on Fractional Gaussian traffic in Section 5, and then conclude the paper.

II. BACKGROUND

A. Long-Range versus Short-Range Dependence

Roughly speaking, Long-Range Dependence (LRD) can be considered as a phenomenon that current observations are significantly correlated to the observations that are farther away in time. This phenomenon is of particular interest to traffic modeling, since it has been discovered recently that both the Ethernet traffic[14] and video sources[2][10] possess the long-range dependence.

One formal definition [10] of a long-range dependent stationary process can be described as that the sum of its auto-correlation function $r(k)$ over all lags is infinite¹. This implies that the auto-correlation $r(k)$ decays asymptotically as a hyperbolic function of k , i.e., $r(k) \sim O(k^{-(2-2H)})$ for $k \geq 0$. H ($0.5 < H < 1$) is the so-called Hurst parameter, which is an important quantity used to characterize the LRD. Examples of such long-range dependent processes include the (asymptotically) second-order self-

similar process, Autoregressive Integrated Moving Average process (ARIMA), and the (exactly) second-order self-similar process, Fractional Gaussian Noise (FGN) process. FARIMA(p, d, q) is a fractional differentiation of an Autoregressive Moving Average (ARMA(p, q)) process, where p and q represent the orders of the ARMA(p, q) process and d ($0 < d < 0.5$) is a differentiation degree. The Hurst parameter H of FARIMA(p, d, q) equals to $0.5 + d$. For a positive p or q , FARIMA(p, d, q) is a mixture of both long-range and short-range dependent process. Because FARIMA(p, d, q) has $p + q + 3$ parameters, it is much more flexible than FGN in terms of simultaneously modeling of both SRD and LRD in the traffic. More detailed explanations on FGN and FARIMA processes can be found in [2].

The short-range dependence, on the other hand, can be characterized by an exponential decay of the autocorrelation function, i.e., $r(k) \sim \rho^k$ ($-1 < \rho < 1$). Examples of short-term dependent random processes include Autoregressive (AR) processes and Autoregressive-Moving-Average (ARMA) processes.

B. Wavelet Transformation

The model we will develop on video traffic is based on wavelets. Wavelets are complete orthonormal bases which can be used to represent a signal as a function of time[4]. In $L^2(R)$, discrete wavelets can be represented as

$$\phi_j^m(t) = 2^{-j/2} \phi(2^{-j}t - m), \quad (1)$$

where j and m are positive integers. j represents the dilation, which in turn characterizes the function $\phi(t)$ at different time scales. m represents the translation in time. Since such wavelets are obtained by dilating and translating a single function $\phi(t)$, $\phi(t)$ is called the mother wavelet. Moreover, all base functions, $\phi_j^m(t)$, have the same shape as the mother wavelet and therefore self-similar with each other.

A discrete-time process $x(t)$ can be represented through its inverse wavelet transform

$$x(t) = \sum_{j=1}^K \sum_{m=0}^{2^{K-j}-1} d_j^m \phi_j^m(t) + \phi_0, \quad (2)$$

where $0 \leq t < 2^K$. K is a positive integer, which characterizes the limited resolution in time. ϕ_0 is equal to the average value of $x(t)$ over $t \in [0, 2^K - 1]$. d_j^m 's are wavelet coefficients, which can be obtained through the wavelet transform

$$d_j^m = \sum_{t=0}^{2^{K-1}} x(t) \phi_j^m(t). \quad (3)$$

The mapping between $x(t)$ and its wavelet coefficients can be shown to be one to one.

The mother wavelet we choose in this work is the Haar wavelet. It is defined as

$$\phi(t) = \begin{cases} 1, & \text{if } 0 \leq t < 1/2, \\ -1, & \text{if } 1/2 \leq t < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

¹Please see [14][10] for other definitions and properties of the LRD.

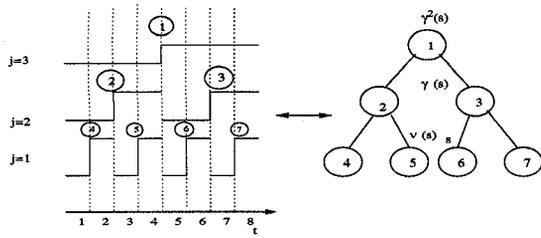


Fig. 1. Left figure shows the Haar wavelet basis functions. Right figure illustrates the corresponding tree diagram and two types of operations. $\gamma(s)$ is defined to be the parent node of node s . $\nu(s)$ is defined to be the left neighbor of node s . The number in the circle represents the one dimension index of the wavelet functions. s , $\nu(s)$ and $\gamma(s)$ also represent the one dimension index of wavelet coefficients.

The corresponding wavelet, $\phi_j^m(t)$, is a scaled and shifted $\phi(t)$ with the support over $[m2^j, (m+1)2^j - 1]$. The Haar wavelet based functions have been illustrated in Figure (1) for $N = 8$. The left figure of Figure (1) illustrates the wavelet base functions. The right figure of Figure (1) shows its corresponding tree diagram[3]. The number in the circle represents the one dimension index of the corresponding wavelet base function. We define $\gamma(s)$ and $\nu(s)$ (see illustration in the tree diagram) to be the node index of the parent of a node s and the node index of the left side neighbor of a node s , respectively.

The motivation for using Haar wavelets is due to simplicity which results in computationally efficient methods for the wavelet transform and the inverse transform. In particular, when the Haar wavelets are used, the wavelet coefficients can be easily derived from Equation (3) as

$$d_j^m = 2^{-\frac{j}{2}} \left(\sum_{t=m2^j}^{(m+0.5)2^j-1} x(t) - \sum_{t=(m+0.5)2^j}^{(m+1)2^j-1} x(t) \right), \quad (5)$$

Since the two terms in the above expression correspond to the summation of the first and the second half of the time series $x(t)$, the wavelet coefficients can be computed very easily. As shown in [4], the computational complexity of the wavelet transform and the inverse transform is in the order of $O(N)$, where $N = 2^K$ is the length of the time series.

When $x(t)$ is a random process, which is of the interest to this work, the corresponding wavelet coefficients are random variables (please refer to [13][23] and references therein for theoretical details on wavelet representations of random processes). Due to the one-to-one correspondence between $x(t)$ and its wavelet coefficients, the statistical properties of the wavelet coefficients are completely determined by those of $x(t)$. Likewise, if the statistical properties of the wavelet coefficients are well specified, they can be used to characterize the original random process. This motivates our approach of using wavelets to model video traffic, i.e., to statistically model wavelet coefficients of the traffic in the wavelet domain.

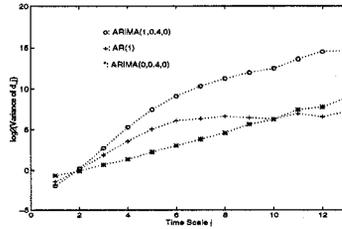


Fig. 2. Log 2 of Variance of d_j versus the time scale j

III. WAVELET MODELING OF LONG-RANGE AND SHORT-RANGE DEPENDENT PROCESSES

As the first step to investigate the feasibility and advantage of wavelets to model video traffic, we will first investigate wavelet representations on well-known long-term and short-term dependent Gaussian processes.

A. The Variances of Wavelet Coefficients

To understand whether and how wavelets capture the LRD and SRD, we plot in Figure (2) the variances of wavelet coefficients for three well-known processes: FARIMA(0, 0.4, 0), FARIMA(1, 0.4, 0), and AR(1)². ARIMA(0, 0.4, 0) is a long-range dependent process with Hurst parameter $H = 0.9$. AR(1) is a short-range dependent process, and FARIMA(1, 0.4, 0) is a mixture of the long-range and the short-range dependent process.

As observed from the figure, for FARIMA(0, 0.4, 0) process (LRD alone), the variance increases with j exponentially for all j . For AR(1) (SRD alone), the variance increases at an even faster rate than that of FARIMA(0, 0.4, 0) when j is small but saturates when j is large. For FARIMA(1, 0.4, 0) (a mixture of LRD and SRD), the variance shows the mixed properties from both AR(1) and FARIMA(0, 0.4, 0).

These results indicate that wavelets are capable of distinguishing the long-range from the short-range dependence.

B. The Correlation Structure of Wavelet Coefficients of LRD Processes

The correlation structure of (long-range dependent) FGN process has been investigated extensively in [13][23][8], and can be summarized as follows.

Theorem 1: (Kaplan and Kuo[13]) Let $x(t)$ be a FGN process with Hurst parameter H ($0.5 < H < 1$). Let d_j^m 's be the (Haar) wavelet coefficients of $x(t)$. Then

(1) for a given time scale j , d_j^m 's are i.i.d. Gaussian random variables with zero mean and variance $2^{j(2H-1)}(2^{2(1-H)} - 1)\sigma^2$, where σ is the variance of $x(t)$.

(2) for $(m_1 + 1)2^{j_1} - m_2 2^{j_2}$ large, where j_1, j_2, m_1 and m_2 are the dilation and the translation indices of two different wavelet coefficients respectively, the correlation between two wavelet coefficients is

$$E(d_{j_1}^{m_1} d_{j_2}^{m_2}) \sim O(|2^{j_1} m_1 - 2^{j_2} m_2|^{-2(1-H)}), \quad (6)$$

²The explicit expressions on the variances for FARIMA processes are too complicated to obtain. The variance of AR(1) process is given in [17].

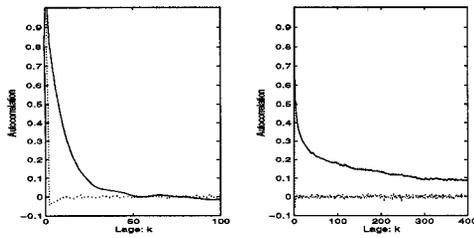


Fig. 3. Solid line: Autocorrelation coefficients of the original process. Dotted line: the normalized auto-correlation of wavelet coefficient, i.e., $\frac{E(d_1^m d_1^{m+k})}{\sigma_{d_1^m} \sigma_{d_1^{m+k}}}$. The left figure is for the AR(1) process. The right figure is for the FARIMA(0,0.4,0) process.

where $H' = 1 - H$.

We recall that the corresponding temporal auto-correlation of Fractional Gaussian Noise decays at a rate $O(|k|^{-2(1-H)})$, where k is the lag between two samples and $0.5 < H < 1$. This rate leads to a divergent summation of the auto-correlations. The above theorem indicates that the wavelet transformation has changed the long-range-dependence in the time domain so significantly that the summation of the correlation of wavelet coefficients converges to a constant. This is because the correlation changes from mean revert ($0.5 < H < 1$) in time domain to mean avert ($0 < H' < 0.5$) in the wavelet domain. Figure (3) illustrates how drastic the reduction is by comparing the autocorrelation function of the original $ARIMA(0, 0.4, 0)$ process to the normalized auto-correlation function³ for d_1^m and d_1^{m+k} , where $ARIMA(0, 0.4, 0)$ is an asymptotically self-similar process and is very similar to a FGN process.

C. The Correlation Structure of Wavelet Coefficients of SRD processes

For short-range dependent processes, there exist no previous results on the explicit correlation structure of wavelet coefficients for discrete processes except the bounds for some of the continuous random processes[5]. Therefore, we derive the correlation wavelet coefficients given in a theorem below.

Theorem 2: Let $x(t)$ be a zero mean wide-sense-stationary (discrete) Gaussian process with the auto-correlation $r(k)$, where $r(k) = \sigma^2 \rho^{|k|}$ with $|\rho| < 1$, k is an integer and σ^2 is the variance of $x(t)$. Let d_j^m 's be the (Haar) wavelet coefficients of $x(t)$. Then

- (1) for a given time scale j , d_j^m 's are Gaussian random variables with a zero mean and a variance $\sigma^2(1 + \frac{2\rho}{1-\rho} - \frac{3\rho}{(1-\rho)^2 2^{j-1}}) + O(\rho^{2^{j-1}})$.
- (2) for $m_1 2^{j_1} - (m_2 + 1) 2^{j_2} > 0$, $E(d_{j_1}^{m_1} d_{j_2}^{m_2}) = 2^{\frac{-j_1-j_2}{2}} \rho^{m_1 2^{j_1} - (m_2+1) 2^{j_2}} (1 - \rho^{2^{j_2-1}})^2 (1 - \rho^{2^{j_1-1}})^2 \frac{\rho}{(1-\rho)^2} \sigma^2$.

The proof of the theorem can be found in [17].

This theorem shows that the correlation of wavelet coefficients decay exponentially as $|m_1 2^{j_1} - m_2 2^{j_2}|$, the short-

³It can be easily shown that the time series d_j^m for fixing j is stationary in terms of m . Therefore, the auto-correlation exists.

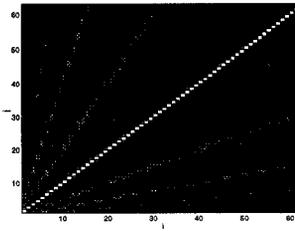


Fig. 4. Correlation Matrix of "FARIMA(0,0.4,0)".

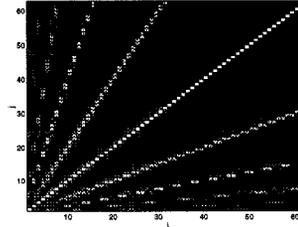


Fig. 5. Correlation Matrix of "AR(1)".

est distance between supports of two wavelet base functions, and the time-scale (dilation) index j increase. The rate of decay is even faster than the corresponding auto-correlation in the time domain. To illustrate the decay rate, we compare the (temporal) auto-correlation for an $AR(1)$ process with that of its wavelet coefficient for $j_1 = j_2 = 1$ in Figure (3). As shown in the figure, the correlation of wavelet coefficient decays very rapidly as k increases.

Combining Theorems 1 and 2, we can see that the wavelet transforms significantly reduce the temporal dependence so that the complicated mixture of the short- and the long-range dependence in the time domain may be sufficiently modeled by a "short-range" dependence process alone in the wavelet domain.

D. Empirical Studies on The Correlation Structure of LRD and SRD

What short-range correlations need to be captured among wavelet coefficients? Unfortunately, an answer to this question can not be provided by the theoretical results given in the previous section, since they only hold for $|(m_1 - 1) 2^{j_1} - m_2 2^{j_2}|$ large. To visualize the correlation structure, the correlation matrices of $ARIMA(0, 0.4, 0)$ and $AR(1)$ are plotted in Figures (4) and (5), respectively. A pixel (i, j) in the image of the correlation matrix represents the correlation between the i -th and the j -th wavelet coefficients, where i and j are the (node) one dimension index in the tree diagram (see Figure (1) for the tree diagram). The gray level is proportional to the magnitude of the correlation. The higher the magnitude of the correlation, the whiter the pixel in the image. The figure shows that besides the diagonal line⁴, there are 4 pairs of lines having "visible" correlations⁵. They correspond to the correlation between $\gamma^k(s)$ and s , where $\gamma(s)$ represents the parent of the node s (see Figure (1) for illustration of

⁴In order to have enough gray level to see more subtle details, the diagonal pixels, which is always 1, is set to 0.5.

⁵We only consider $K = 5$ which has only 5 level in the tree diagram.

the definition of $\gamma^k(s)$, and $\gamma^k(s)$ denotes the parent of the node $\gamma^{k-1}(s)$ with k being 1, 2, 3 and 4 counting from the diagonal line. From the figures, we can conclude that the most significant correlation is due to the parent-child relationship. Since the complicated correlation in the time domain actually concentrates on these types of correlations in the wavelet domain, we can use a parsimonious model in the wavelet domain to faithfully represent the original traffic.

E. Modeling The Correlation Structure

We propose several models to model the partial correlation among wavelet coefficients, ranging from the simplest to the most complex. The simplest model assumes that the wavelet coefficients are statistically independent. The most complicated model is a third order Markov model.

Model 1: Independent model, which only models the mean and the variance of wavelet coefficients. d_j^m 's are chosen as independent Gaussian variables with zero mean and variance σ_j estimated from data at each j .

Model 2: First order Markov model, which models the correlation between $\gamma(s)$ and s , i.e., the parent-child relationship. This can be implemented as follows.

$$d_{s_j} = a_j d_{\gamma(s_j)} + b_j w_j, \quad (7)$$

where a_j and b_j are the parameters to be determined from data. w_j is Gaussian noise with zero mean and unit variance. $\gamma(s_j)$ represents the parent of s_j , where s_j is the one dimension index of a wavelet coefficient (node) in the j -th time-scale (See Figure (1) for illustration of the one dimension index of a wavelet coefficient.).

Model 3: Third order Markov model, which models the correlation among $\gamma(s)$, $\gamma^2(s)$, $\nu(s)$ and s . In the (correlation matrix) graph, this model can match the first two strongest lines as well as the (barely visible) line near the diagonal, which represents the neighboring relationship. Please refer to [17] for details on the algorithm.

In terms of the complexity of the aforementioned models, Model 1 only requires one parameter (the variance) σ_j at each level j ⁶. Models 2 and 3 have two and four parameters at each level, respectively.

F. An Algorithm on Generating Wavelet Models for Gaussian Processes

The models on the correlation structure can now be included in an algorithm to obtain wavelet models for a Gaussian process.

Let $\hat{x}(t)$ be the trace of length N from a Gaussian process.

Algorithm 1

1. Perform wavelet transform on $\hat{x}(t)$ to obtain \hat{d}_j^m 's (wavelet coefficients of $\hat{x}(t)$).
2. Estimate the required parameters in the selected wavelet correlation model (Section 3.5) from \hat{d}_j^m 's.

⁶The mean of d_j^m can be shown to be zero for a stationary process.

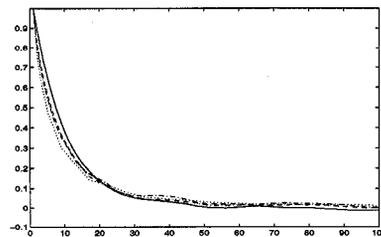


Fig. 6. Sample Correlations. “.”: AR(0.9); “—”: Model 3; “- -”: model 2; “-·-”: model 1.

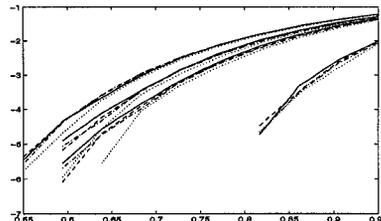


Fig. 7. Buffer Response. x-axis: utilization. y-axis: $\log_{10}(\text{Overflow probability})$. “.”: AR(0.9); “—”: Model 3; “- -”: model 2; “-·-”: model 1. The normalized buffer size is 0.1, 0.5, 1, 10 from top down.

3. Generate coefficients d_j^m from the wavelet correlation model with previously computed parameters for all m and j .
4. Do inverse wavelet transform using the new wavelet coefficients (d_j^m 's) obtained at the previous step to get the synthesized traffic in the time domain.

To estimate the computational complexity of the algorithm, notice that the computational complexity of the wavelet transform (Step 1) and the inverse transform (Step 4) is $O(N)$ respectively. The computational complexity of Steps 2 and 3 is also $O(N)$ ⁷. Then the total computational cost of the algorithm is $O(N)$.

G. Experimental Results

The algorithm is first used to evaluate the performance of our correlation models. Sample paths of length 2^{16} are generated from an AR(1) process with Gaussian noise and the AR parameter to be 0.9. We plot the sample autocorrelation and buffer response in Figures (6) and (7), respectively, for all three correlation models presented in Section 3.5. We observe from the figure that the simplest wavelet model which neglects the dependence in the wavelet domain performs reasonably well. The models which capture more correlations among wavelets only improve the performance slightly.

To further test the performance of Model 1 which generates independent wavelet coefficients, we generate the corresponding wavelet models for FARIMA processes. A sample path of length 2^{16} is generated from FARIMA(0, d , 0) (LRD alone) for $d = 0.2, 0.3, 0.4$, respectively, using a simulator in SPLUS. Such a sample path provides a time series $\hat{x}(t)$ to be used by Algorithm 1. A synthesized time series from the wavelet model is then generated by Algorithm 1

⁷ Assume that $O(1)$ time is needed to generate one Gaussian random variable.

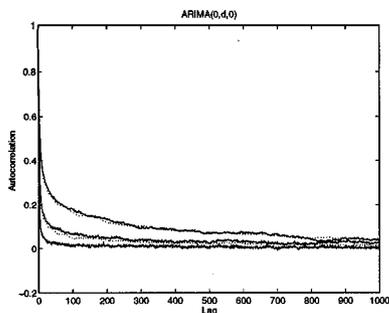


Fig. 8. Autocorrelation functions for FARIMA(0,d,0) (solid lines) and Algorithm 1 (dotted lines). $d=0.2, 0.3$ and 0.4 from the bottom up.

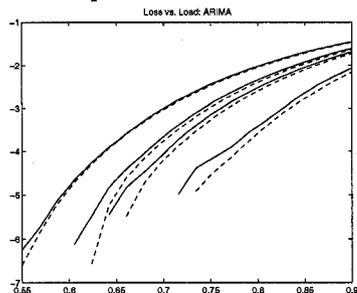


Fig. 9. The vertical axis: $\log_{10}(\text{Loss Rate})$. The horizontal axis: work load. The solid lines: FARIMA(0,0.4,0). Dotted lines: Algorithm 1. The normalized buffer size: 0.1, 0.5, 1, and 10 from the top down.

using Model 1. The original and the synthesized sample paths are both used to test the performance of the wavelet model.

The autocorrelation functions estimated from the sample paths are plotted in Figure 8. The loss rate is estimated from the sample path from the original FARIMA process and the wavelet model, we plot the log of loss rate versus the work load for the normalized buffer size 0.1, 0.5, 1, and 5, respectively. The results show that the auto-correlation function and the loss rate due to the wavelet model are very close to those due to the actual FARIMA process.

IV. WAVELET MODELING OF VIDEO TRAFFIC

A. The Video Sources

Two video traffic sources are chosen to test our algorithms: (1) JPEG coded “Star Wars” at frame level [10], and (2) MPEG-I coded videos at frame level. The trace “Star Wars” [10] is obtained by applying JPEG-like encoder to each of 171,000 frames at an interval of $\frac{1}{24}$ second per frame of the 2-hour movie of “Star Wars”⁸. This source is used to test our model, since the source provides rich variations in terms of scene changes. The second data set is the MPEG coded video source. The source is constructed from MPEG-I encoded video sequences created by Rose [21]. We choose a video called “Jurassic Park”, which has 53332 frames at an interval of 40ms. A standard MPEG encoder generates three types of compressed frames: I, P and B. I frames are compressed using intra-frame coding only, while, P and B frame, in addition to

intra-frame coding, allow using motion compensation techniques. As a result, I frames are the largest in size, followed by P frames and B frames. In addition, many MPEG encoders have the same Group-of-Picture (GOP) pattern. The GOP pattern in “Jurassic Park” consists of 12 frames as *IBBPBBPBBPBB*.

B. Modeling the Video Sources

An extension to Algorithm 1 is needed to include the following important properties of video traffic: (1) marginal probability density functions of wavelet coefficients at each time scale for non-Gaussian wavelet coefficients, and (2) a mechanism to deal with the periodic structure in MPEG-I video at the frame level.

Since the generated wavelet coefficients by Algorithm 1 actually have a Gaussian marginal density function, a transformation is needed on the generated d_j^m 's so that the resulting d_j^m 's can have the same empirical density function as that of the data. Such a transformation can be done easily through a method described in [11] with little computation.

The MPEG-coded video traffic possesses periodic structure due to the repeated GOP patterns. To deal with the periodic structure, we note that the interframe redundancy is reduced by using P and B frames. This results in I, P, B frames with significantly different statistical characteristics. Therefore, I, P, and B frames should be distinguished from one another at the frame level for modeling. Above the GOP level, no such distinction is needed. Algorithm 1 can be modified for MPEG traffic by taking the advantage of wavelet modeling applicable at different time scales. Specifically, we can use Algorithm 1 directly to model the MPEG video traffic above the GOP level. Below the GOP level⁹, we treat wavelet coefficients with the same relative position in a GOP pattern as a group, and assume those wavelet coefficients have the same statistical properties. We can then model different groups of wavelet coefficients by choosing different parameters in the correlation models.

C. Experimental Results

To evaluate both the performance and the computational efficiency of the wavelet models (Algorithm 1), we apply the algorithm to obtain wavelet models for two video sources, and to generate synthesized video trace. Models 1 (independent wavelet coefficients) and 2 (Markov-dependent wavelet coefficients) have been used, and shown comparable performance. The results are thus only presented for Model 2. FARIMA models are used to model the sources as well for comparison. In particular, FARIMA models are chosen to have 25 AR terms and 20 MA terms. A maximum-likelihood algorithm provided by SPLUS [22] is used to determine all 45 parameters for the FARIMA models using the video trace data. Furthermore, the Gaussian marginal distribution of the FARIMA process is shaped to be the same as that of trace data by the same method

⁹We assume the GOP pattern is periodic as in most cases. If the length of GOP pattern is not a power of 2, we will add zero size frames to the end of each GOP to make it a power of 2.

⁸Detailed description on the “Star Wars” trace can be found in [10].

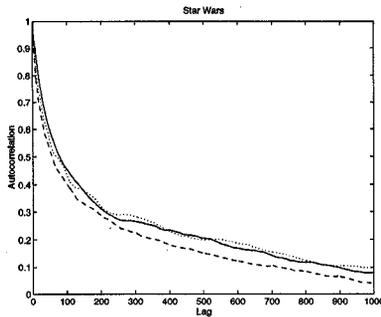


Fig. 10. “-”: Autocorrelation of “Star Wars”; “- -”: FARIMA(25,d,20); “..”: Algorithm 2

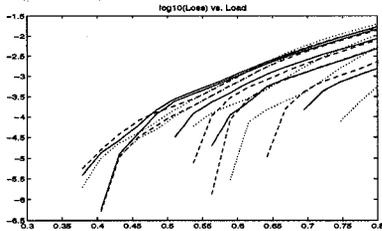


Fig. 11. Vertical axis: $\log_{10}(\text{Loss Rate})$; horizontal axis: work load. “-”: the single video source; “- -”: FARIMA(25,d,20); “- - -” Algorithm 2. The normalized buffer size: 0.1, 1, 10, 30 and 100 from the top down.

described in Section 4.2 to make fair comparisons. Synthesized video traffic from both wavelet and FARIMA models are used to obtain sample auto-correlation functions in the time-domain, and to estimate the buffer loss rate. Since the FARIMA(25, d , 20) gives very poor results on the MPEG source¹⁰, the results are not given in corresponding figures.

The auto-correlation functions of FARIMA(25, d , 20) and of the wavelet model are plotted in Figures 10 and 12 for the single “Star Wars” source and the “MPEG” source, respectively. As can be seen, the wavelet models have consistently a better match to the auto-correlation functions than the FARIMA models, especially for single sources. The reason is probably due to the fact that the wavelet model can match the marginal distribution at different time scales. The results on the loss rate of a single buffer due to each source are given in Figures 11 and 13, respectively. It can be seen that the wavelet model has a loss rate comparable to that of FARIMA(25, d , 20) at the small buffer size, and a better performance at the large buffer size.

As for the computational time, it takes more than 5-hour CPU time on a SunSPARC 5 workstation for an FARIMA(25, d , 20) model to estimate its parameters from the data and to generate synthesized video traffic of length 171,000. It only takes 3 minutes on the same machine for Algorithm 2 to complete the same tasks. The computational complexity to generate synthesized video traffic of length N is $O(N^2)$ for an FARIMA model, and only $O(N)$ for a wavelet model.

¹⁰This is due to the fact that the MPEG source is more complicated, and therefore, a more complex FARIMA model is needed.

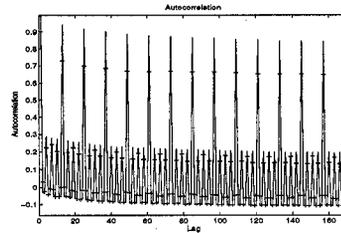


Fig. 12. Autocorrelation. “-”: “Jurassic Park”; “+”: Wavelet Model

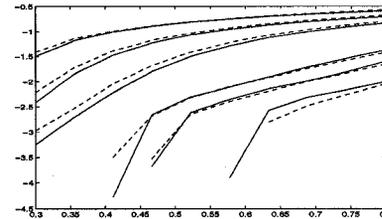


Fig. 13. Buffer Response. x-axis: utilization. y-axis: $\log_{10}(\text{Overflow probability})$. “-”: “Jurassic Park”; “- -”: Wavelet Model. The normalized buffer size: 0.1, 0.5, 1, 10, 30, 100 from the top down.

V. ANALYSIS ON THE BUFFER OVERFLOW PROBABILITY

In this section, we will provide theoretical analysis on the buffer overflow probability of our wavelet model. For simplicity, we will present here the analysis for a special case when the incoming traffic to the buffer is assumed to be the Fractional Gaussian Noise (FGN). It has been shown in Sections 3.4 that the independent wavelet model can model FGN process and other process with Gaussian marginal density functions. Since the independent wavelet model neglects dependence among wavelet coefficients, it is important to estimate the buffer overflow probability obtained for the wavelet model, and compare with that for the FGN process. To our knowledge, this has not been done previously, and is a necessary step to validate our basic idea on using wavelet models to model video traffic using the buffer loss rate as a performance measure.

Let $\hat{x}(t)$ be a FGN process which represents the actual traffic to the buffer. Let $x(t)$ be the wavelet representation of $\hat{x}(t)$, where the wavelet coefficients are obtained through Algorithm 1. Let μ and σ be the mean and the variance of $x(t)$, respectively. Suppose $x(t)$ is fed into a single queue at the beginning of the slotted time t for $0 \leq t < N$, where $T = 2^K$. Let C represent the capacity. Let B_N and \hat{B}_N be the buffer sizes at the end of the $(N - 1)$ -th time slot due to the synthesized traffic $x(t)$ by the wavelet model and the FGN process $\hat{x}(t)$, respectively. For the infinite buffer size, B_N and \hat{B}_N can be expressed as [7][19]

$$B_N = \sup_s \left(\sum_{i=1}^s x(N-i) - sC \right), \quad (8)$$

$$\hat{B}_N = \sup_s \left(\sum_{i=1}^s \hat{x}(N-i) - sC \right). \quad (9)$$

Then the probability of the buffer size over B , for the

wavelet model, can be given through a theorem below.

Theorem 3: The buffer overflow probability due to the synthesized (wavelet) traffic $x(t)$ satisfies

$$\begin{aligned} Pr(B_N > B) &\sim Pr(\hat{B}_N > B) \\ &\sim \exp\left(-\frac{(C - \mu)^2 \left(\frac{B}{C - \mu}\right)^{2(1-H)} \left(\frac{1-H}{H}\right)^{2H}}{2\sigma^2(1-H)^2}\right), \end{aligned}$$

where $N = 2^K$ is the length of the traffic with K being a positive integer. B can be expressed as

$$B = (C - \mu)2^\beta, \quad (10)$$

where β is assumed to be an positive integer k_0 for simplicity.

The proof of the theorem can be found in [17].

This result shows that the simple wavelet model which ignores the correlations among different wavelet coefficients has the same Weibull decay as that of FGN in terms of the loss rate given in [7][19]. In another word, wavelet models obtained by Algorithm 1 can faithfully model the FGN traffic. Therefore, this result, as the first step, shows the capability and performance of the wavelet models, and the feasibility for using the models for network design.

We would like to point out that in addition to the assumption for FGN workload, the theoretical result is limited by two conditions: (a) the buffer size B_t is only considered at time $t = N$ instead of all t , and (b) the buffer full size B is assumed to be $2^{k_0}(C - \mu)$, which only represents a subset of all possible values for B . Extension of the results in a more general setting will be given in our future work.

VI. CONCLUSIONS

An important discovery from this work is that wavelet coefficients of the video traffic, which has complicated short- and long-range temporal dependence, are no longer long-range dependent in the wavelet domain. Therefore, the "short-range" dependent process can be used to model the video traffic in the wavelet domain. This opens up new possibilities for modeling, analyzing and controlling the long-range dependent video traffic.

In this work, we have developed wavelet models for VBR video traffic. The model provides new understanding on the co-existence of the long-range and the short-range dependence in the video traffic, and a unified approach to capture both LRD and SRD simultaneously. The good performance of the model has been obtained through extensive tests on two video sources using both the auto-correlation and the buffer loss rate as performance measures. The computational complexity for developing the model and for synthesizing a large volume of video traffic has shown to be $O(N)$, which is the lowest attained.

ACKNOWLEDGMENT

We would like to thank Olive Rose and Mark Garrett for making the video traces available, and would like to thank Marina Thottan and Xusheng Tian for helpful comments.

The support from National Science Foundation ((CA-REER) IRI-9502518 and ECS-9312594) is gratefully acknowledged.

REFERENCES

- [1] Allan T. Andersen and Bo Friis Nielsen. An application of superpositions of two state Markovian sources to the modeling of self-similar behavior. In *Proceedings of INFOCOM*, 1997.
- [2] Jan Beran, Robert Sherman, Murad S. Taqqu, and Walter Willinger. Long-range dependence in variable-bit-rate video traffic. *IEEE Transactions on Communications*, 43, 1995.
- [3] Kenneth C. Chou, Stuart A. Golden, and Alan S. Willsky. Multi-resolution stochastic models, data fusion, and wavelet transforms. *Signal Processing*, 34(3), 1993.
- [4] I. Daubechies. *Ten Lectures on Wavelets*. Philadelphia: SIAM, 1992.
- [5] Dijkerman, Robert W., and Ravi R Mazumdar. Wavelet representations of stochastic processes and multi-resolution stochastic models. *IEEE Transactions on Signal Processing v 42 n 7 Jul 1994 p 1640-1652*, 1994.
- [6] D.P.Heyman and T.V.Lakshman. What are the implications of long-range dependence for VBR-video traffic engineering? *IEEE/ACM Transactions on Networking*, 1996.
- [7] N.G. Duffield and N. O'Connell. "Large deviations and overflow probabilities for the general single server queue, with applications". DIAS-STP-93-30, Dublin Institute for Advanced Studies, 1993.
- [8] Patrick Flandrin. Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Transactions on Information Theory*, 38(2), 1992.
- [9] V. Frost and B. Melamed. Traffic modeling for telecommunication networks. *IEEE Communication Magazines*, 32, 1994.
- [10] Mark W. Garrett and Walter Willinger. Analysis, modeling and generation of self-similar VBR video traffic. In *Proceedings of SIGCOM*, 1994.
- [11] C. Huang, M. Devetsikiotis, I. Lambadaris, and A.R. Kaye. Modeling and simulation of self-similar variable bit rate compressed video: A unified approach. In *Proceedings of SIGCOM*, 1996.
- [12] Predrag R. Jelenovic, Aurel A. Lazar, and Nemo Semret. The effect of multiple time scales and subexponentiality in MPEG video streams on queueing behavior. *IEEE Journal on Selected Area of Communications*, 15, 1997.
- [13] L.M. Kaplan and C.J. Kuo. Fractal estimation from noisy data via discrete fractional Gaussian noise (DFGN) and the Haar basis. *IEEE Transactions on Information Theory*, 41(12), 1993.
- [14] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of ethernet traffic. *IEEE/ACM Transactions on Networking*, 2(1):1-15, February 1994. Available through ftp://ftp.bellcore.com/pub/wel/tome.ps.Z.
- [15] S.Qi Li and C-L Hwang. Queue response to input correlation functions: Discrete spectral analysis. *IEEE/ACM Transactions on Networking*, 1:317-329, 1993.
- [16] S. Ma and C. Ji. Wavelet traffic modeling, May, 1997. technical report.
- [17] S. Ma and C. Ji. Modeling network traffic in the wavelet domain, Oct. 1997. submit to JSAC.
- [18] B. Melamed, D. Raychaudhuri, B. Sengupta, and J. Zdepski. TES-based video source modeling for performance evaluation of integrated networks. *IEEE Transactions on Communications*, 10, 1994.
- [19] I. Norros. "A storage model with self-similar input". *Queueing System*, 16:387-396, 1994.
- [20] V. Paxson. Fast approximation of self-similar network traffic. Available through ftp://ftp.ee.lbl.gov/papers/fast-approx-selfsim.ps.Z, 1995.
- [21] O. Rose. Simple and efficient models for variable bit rate MPEG video traffic. Technical Report 120, Institute of Computer Science, University of Wuerzburg, July 1995.
- [22] Statistical Sciences, Seattle: StatSci, a division of MathSoft, Inc. *S-PLUS Guide to Statistical and Mathematical Analysis, Version 3.3*, 1995.
- [23] G.W. Wornell and A.V. Oppenheim. Wavelet-based representations for a class of self-similar signals with application to fractal modulation. *IEEE Transactions on Information Theory*, 38, 1992.