Threshold Autoregressive Models for VBR MPEG Video Traces

Bongseog Jang[1] and Charles Thompson[2]
Center for Advanced Computation and Telecommunications
University of Massachusetts, Lowell, MA 01854

Abstract- In this paper variable bit rate VBR Moving Picture Experts Group (MPEG) coded full-motion video traffic is modeled by a nonlinear time-series process. The threshold autoregressive (TAR) process is of particular interest. The TAR model is comprised of a set of autoregressive (AR) processes that are switched between amplitude sub-regions. To model the dynamics of the switching between the sub-regions a selection of amplitude dependent thresholds and a delay value is required. To this end, an efficient and accurate TAR model construction algorithm is developed to model VBR MPEG-coded video traffic. The TAR model is shown to accurately represent statistical characteristics of the actual full-motion video trace. Furthermore, in simulations for the bit-loss rate actual and TAR traces show good agreement.

Keywords: video modeling, MPEG, TAR, VBR, nonlinear model

1. Introduction

To better support video services on high speed and integrated networks an understanding of the characteristics of VBR video traffic is required. Video traces with low levels of scene activity have exponentially decaying temporal correlations with respect to time-lag [6,13]. Whereas video traces with non-uniform scene activity have frame sizes that change slowly over long time intervals. The autocorrelation function (ACF) for these types of video traces decay slowly or do not reach zero even for long lag intervals [1,15]. In addition, abrupt jumps in the frame size occur after a scene change. VBR video traces with non-uniform scene activity follow no specific probability distribution function (PDF) for the number of cells in scene change frames and for scene length [7]. VBR video traffic with moderate scene activity can also exhibit long-term correlation [12,14,16]. To model video traffic, methods that efficiently capture the switching between these levels of activity are required. In this paper a new model for MPEG-coded VBR video is examined and verified. Its usefulness in the characterization of full-motion video is discussed.

The discrete autoregressive (DAR) model has been used to model broadcast-video traces generated by a DPCM-based coding algorithm without motion compensation [7]. In the DAR(1) model, a finite-state Markov chain is used to generate sequence of states. These states are used to determine the frame sizes. This model requires only that the mean, variance, and autocorrelation coefficient of intra-scene frames be determined. The DAR(1) model was not found to be accurate for all video traces tested, however. For the video conference traces, the DAR(1) proved to be a good source model.

A time-varying AR process was applied to model full-motion video coded by a DPCM/DCT scheme [18]. The codec generates video stream using a three-level motion classification, i.e. high-, medium-, and low-motion activities. Short-range dependence (SRD) was modeled using sub-AR(1) processes. A finite state-discrete-time Markov chain was used to choose between the sub-AR(1) models. The number of frames used to generate the model was 500 frames and two arbitrary thresholds were selected from bit-rate histogram of the video trace. Recently an enhanced Markov chain based approach has been used with success to analyze the traffic from single and two layer MPEG-2 coders [15].

A scenic model [3] based on the DAR(1) modeling approach has been used to model VBR traffic. Scene changes were estimated using differences in the number of bits between consecutive frames rather than by using a Markov process. To discern the scene changes, the VBR video trace was first passed through a median-filter having a length of 0.5 seconds. Using the output, the short-time mean was calculated using the 5 frame average-filter. The short-time average value exhibits a significant change in the value at a scene boundary. Tests of the model using full-motion video traces showed that for large buffer sizes the scenic model estimated cell-loss probability more accurately than the DAR(1) model. A self-similar model was developed to estimate the long-range dependence (LRD) and was used in conjunction with SRD for VBR full-motion video traces [8]. In this work, self-similar traffic models were used to match the LRD, SRD, and probability density function (PDF). In queueing simulations the model underestimated cell-loss rates when compared to the actual trace. In [10] presence of a scene change in an MPEG video trace was determined using the difference in the frame bit count between two consecutive I-frames. Two thresholds were used as a measure of the scene activity. Intra-scene fluctuations for I-frames were estimated using an AR(2) process. Each frame type was fit to a lognormal distribution using the histogram. Composition of each frame type according to group of picture (GOP) format generated a video trace possessing the characteristics of VBR video traffic.

In this paper a nonlinear time-series modeling method is developed for VBR MPEG video traces. The TAR model [17] consists of a number of AR models. Each AR model has its own correlation structure and model order. Switching of the AR models is dependent on amplitude thresholds and the amplitude of a time-delayed sample. Therefore the conditions for transitions between the sub-regions is deterministic. The best TAR model is selected from evaluations of all possible TAR models realized from a given set of thresholds and time-delays. A new TAR model construction algorithm is examined which uses a minimum variance criterion for model selection. The algorithm significantly reduces computational cost and accurately finds the best TAR model.

1 Samsung Telecommunications Research, Samsung, Korea, bsjang@pharaoh.telecom.samsung.com.kr

2 Charles_Thompson2@uml.edu; 978-934-3360

The structure of the MPEG-coded full motion video trace is introduced in Section 2. The TAR model construction algorithm is presented in Section 3. Section 4 is devoted to the TAR modeling of a VBR MPEG-coded full-motion video trace. In section 5, bit-loss rates are estimated for different buffer sizes and transmission speeds using the leaky bucket algorithm. Finally Section 6 is devoted to conclusions.

2. VBR MPEG Video Trace

A bits per frame trace of an MPEG-1 coded video sequence is shown in Figure 1. For different amplitude ranges, piecewise-constant frame sizes are observed. These constant regions may extend several hundred frames. Large amplitude changes are also observed at intermittent intervals. Scene changes generate large jumps in the frame size. The bit rates in successive frames in a scene are strongly correlated. The figure shows the bit count for frames 31,000 through 33,000 for the full-motion video James Bond. Fluctuations in the trace are necessary to retain constant video quality [2].

In the case of the MPEG coding method [11], three types of frames are used to code contents of the video. These are the I, P and B frames. The I-frames use intra-frame coding based on the discrete transform and entropy coding. The P-frames include motion compensation from previous I- or P-frames. The B-frames include bi-directional motion compensation. Typically, I-frames require more bits than P-frames. The B-frames have the lowest bandwidth requirement. After encoding is completed, deterministic and periodic frame sequences are generated. This sequence is called group of pictures (GOP) format.

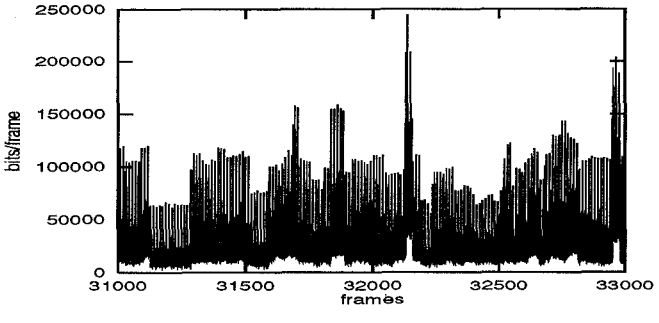

Figure 1: Representative sample path for a typical MPEG trace

The MPEG-1 trace data used in this paper is available to public for research purpose from [4]. All the traces obtained from [4] have rather active scene changes. The traces for news and sports have more frequent jump bit-rates than the movie sequences. The maximum number of frames in a trace is 40,000 which is equal to about 30 minute-long video sequence for a source frame generation rate of 30 frames per second. The GOP pattern consists of 12 frames as having a pattern IBBPBBPBBPBB. The traces are encoded color images having a maximum of 12-bits per pixel. The picture size is 384 X 288 pixels [5].

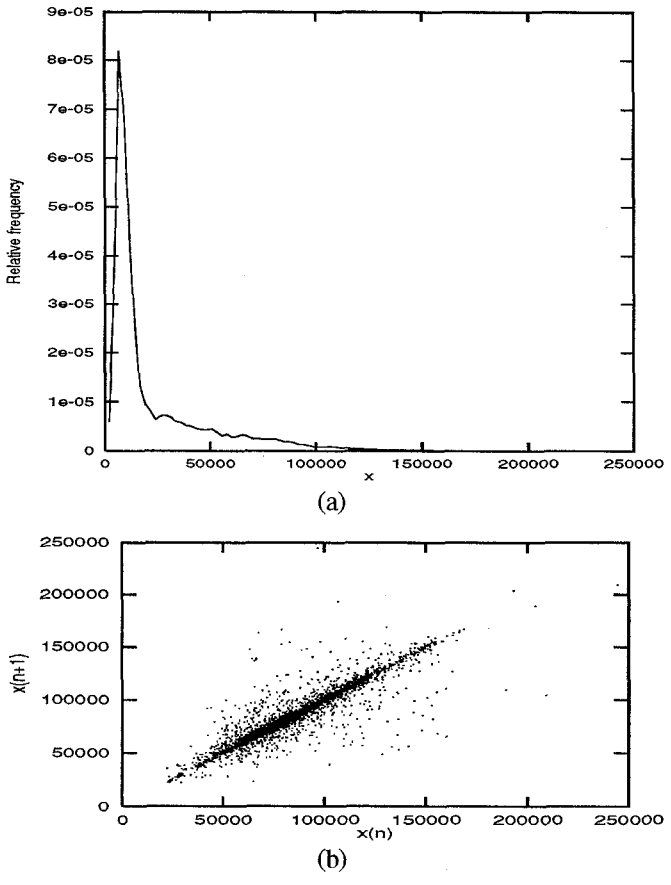

Figure 2: (a)PDF of MPEG video trace (b)Scatter diagram of I frames

The statistics of the MPEG traces are listed in Table 1. The peak to mean ratio is a rough measure of the degree of variability present in the trace. In all cases, the ratio is large. The PDF for the James Bond trace has a narrow peak near 24000 bits per frame. This peak primarily due to the B-frames, whereas the long tail includes contributions from I- and P-frames. The highest bit-rates come from I-frames.

Scene Type	Mean (bits/frame)	Variance	Peak to Mean
Movie(Star Wars)	9313	$1.7\ 10^8$	13
TV news	15358	$3.8\ 10^8$	12
Movie(James Bond)	24308	$6.6\ 10^8$	10
Soccer game	25110	$4.5\ 10^8$	8
Movie(Terminator)	10905	$1.0\ 10^8$	7

Table 1: VBR MPEG-coded video traces

The one-lag scatter diagram of the I-frames, based on the 12 frame period of the GOP format, is shown in Figure 2(b). A positive correlation with a single cluster is observed in the figure. The slope of the linear regression line through the cluster has a value of nearly unity. This implies that I-frames remain close in amplitude at one-lag. Table 1 also demonstrates that

different video sequences using a single video coder can generate statistically varying traces. This state-of-affairs may introduce problems in providing quality of services (QoS) and reliable network performance. Thus a model is needed to characterize efficiently and accurately the degree of activity in the video trace.

3. The TAR Model

The TAR model is a nonlinear time-series process comprised from a number of linear sub-models. Each amplitude switched AR process is constructed for specific amplitude range or sub-region. The AR model to be used at time n is determined by the amplitude $x(n-D)$ where D denotes a time-delay. Also the amplitude thresholds are used to determine which AR model is to be activated. The TAR model for sub-region m is defined as

$$x(n) = a_0^{(m)} + \sum_{i=1}^{p_m} a_i^{(m)}\, x(n-i) + u^{(m)}(n) \qquad (1)$$

if $R_m \le x(n-D) < R_{m+1}$. The variable x is the time-series observed, m is the index denoting the sub-region, $a_i^{(m)}$ are coefficients of the model for region m, $u^{(m)}(n)$ is a Gaussian distributed noise with $N(0: \sigma_{(m)}^2)$, R_m the denotes threshold amplitude, and p_m represents sub-AR model order in the region m.

The analysis will proceed by first splitting the video trace in I, B and P traces. Each will be modeled separately. Once the model is determined, the time-series can be constructed following the GOP format. In the next section, the sub-AR model is characterized by estimating $a_i^{(m)}$ and $u^{(m)}(n)$. A new algorithm for the TAR model construction is presented in Section 3.2. Criteria for the selection of R_m, p_m and D are also introduced in this section.

3.1 Characterization of the Sub-Region

Each sub-region is characterized in terms of the coefficients and residual variance of the sub-AR process. The sub-region is classified by amplitude thresholds and a time-delay amplitude. Figure 3 illustrates j^{th} classification in the sub-region m. Current point $x(n)$ is determined to be the j^{th} realization for the sub-region m since $x(n-D)$ shown as X in the figure is in the threshold range R_m and R_{m+1}. The $x(n)$ is a realization of random process having a dependence on p_m lagged time-series points. Hence we can denote the classified time-series points as

$$x_0^{(j)} = x(n),\ x_1^{(j)} = x(n-1),\ \cdots,\ x_{p_m}^{(j)} = x(n-p_m) \qquad (2)$$

where $x_0^{(j)}$ is a measured value from the p_m lagged past data points at the occurrence j.

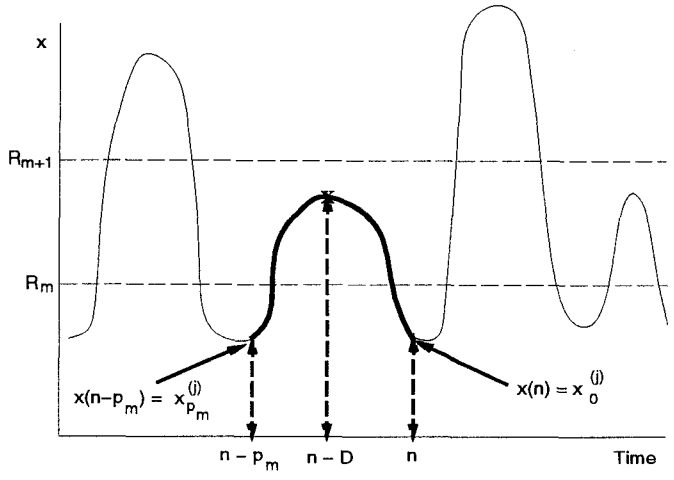

Figure 3: The j^{th} realization in the region m

Figure 4 shows realizations of a random process according to their order of occurrence for the region m. The total number of realizations is N_m and the model for each sequence uses p_m lagged values of x. In the bottom line of the figure, $\bar{x}_i^{(m)}$ denotes the ensemble average of the random process at each lag point i. The estimation of $a_i^{(m)}$ proceeds by minimizing squared error between the estimated and measured value of $x(n)$. This error J_m in the sub-region m can be written as

$$J_m = \sum_{j=1}^{N_m} \left\{ x_0^{(j)} - a_0^{(m)} - \sum_{i=1}^{p_m} a_i^{(m)} x_i^{(j)} \right\}^2. \qquad (3)$$

Setting the partial derivative of J_m in terms of $a_0^{(m)}$ equal to zero yields

$$a_0^{(m)} = \bar{x}_0^{(m)} - \sum_{i=1}^{p_m} a_i^{(m)} \bar{x}_i^{(m)} \qquad (4)$$

where $\bar{x}_i^{(m)}$ is the ensemble average. Substituting $a_0^{(m)}$ into Eq.(3) and differentiating the result with respect to $a_i^{(m)}$ yields the following correlation structure for the sub-region m as

$$\Phi^{(m)}(0,i) - \sum_{k=1}^{p_m} a_k^{(m)} \Phi^{(m)}(k,i) = 0 \qquad (5)$$

where $\Phi^{(m)}(k,i)$ is an ACF and written as

$$\Phi^{(m)}(k,i) = \sum_{j=1}^{N_m} \hat{x}_k^j \hat{x}_i^j / N_m \qquad (6)$$

where k and i denote different lags up to the order p_m, and $\hat{x}_i^{(j)} = x_i^{(j)} - \bar{x}_i^{(j)}$. The function $\Phi^{(m)}(k,i)$ represents correlation of the lags i and k resulting from averaging over the realizations.

The TAR model will possess a different $\Phi^{(m)}(k,i)$ for each sub-region. Evaluating Eq.(5) for i from 1 to p_m yields a matrix equation which can be used to estimate coefficients $a_i^{(m)}$ in the sub-region m as

$$\mathbf{a} = [\Phi]^{-1}\Phi(0,i) \tag{7}$$

where $\mathbf{a}$ is a coefficient vector, $[\Phi]^{-1}$ is the inverse of the ACF matrix, and $\Phi(0,i)$ is an ACF vector. At the lag $i=0$ in Eq.(5) the variance of the residual $u^{(m)}$ is calculated as

$$\sigma_u^{(m)^2} = \Phi^{(m)}(0,0) - \sum_{k=1}^{p_m} a_k^{(m)}\Phi^{(m)}(k,0) \tag{8}$$

Hence the sub-AR model is estimated using conditional correlation. The structure of the model varies in each sub-region m in accordance with the chosen thresholds and time-delay. In the next section, a method for selecting values for these parameters is discussed.

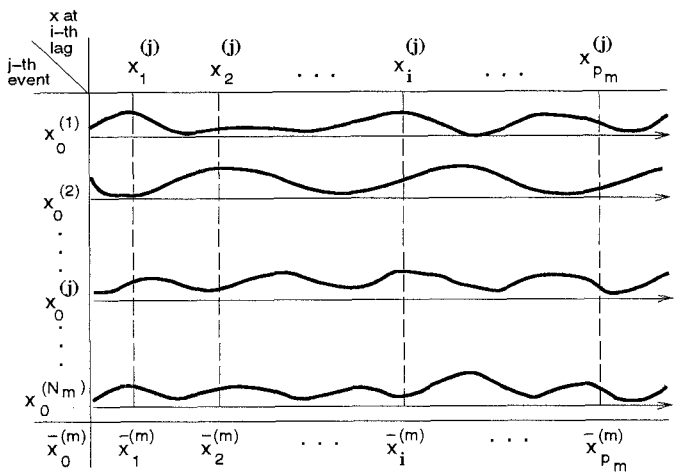

Figure 4: Total N_m realizations in the region m

3.2 Selection of the Best Model

A method to determine the optimal R_m, p_m and D is needed to select the best TAR model among all possible cases. The classical method used in time-series studies uses the Akaike information criteria (AIC) for the TAR model construction [17]. In addition to the AIC for modeling the full-motion video trace, a new model selection criterion is implemented to accurately evaluate the candidate TAR models. The best TAR model is selected by composing the parameters R_m,p_m and D that yield the closest fit to the data.

In the TAR model construction, effort is required to obtain the sub-AR models. The computational cost can be reduced by reducing the number of sub-regions evaluated. This can be done while retaining model accuracy. Given a set of $K+1$ threshold amplitudes, an AR model is required in the interval R_m and R_{m+h} with $R_m < R_{m+h}$ for $h = 1, 2, \cdots, K-m$. The number of AR model increases as the number of thresholds increase. Let the thresholds be defined as $R_0 < R_1 < \cdots < R_m < \cdots < R_{m+h} < \cdots < R_K$ where $R_0 = 0$ and $R_K = P$. The P is a value greater than the peak rate of the MPEG video trace. The number of TAR model to be evaluated for k threshold-points between R_0 and R_K is $\binom{K-1}{k}$ for $1 \le k \le K-1$. Hence the number of sub-AR model for the given k will be $\binom{K-1}{k}(k+1)$. By summing all cases, the number of sub-AR model will increase as $O(2^K)$. One can reduce the number of sub-AR models required by only evaluating representative models for the sub-regions once. Since the amplitude domains of some of sub-AR models are shared by the different TAR models, the number of sub-AR model will be reduced without repeated revaluation of sub-AR models in each TAR model. The sub-AR models to be evaluated are decreased to $O(K^2)$ [9] using representative sub-AR models.

The conventional criterion for determining the best TAR model is the AIC. In such a case, one must evaluate and compare the AIC from $O(2^K)$ candidates. Typically one would select the model having the minimum AIC. However the AIC is of limited use in choosing the best model for video traffic. Due to the large and intermittent fluctuations in the video traces, the variance of the residual used in the AIC is not a reliable discriminator. We will introduce sum of weighted variance as the criterion for selecting the best TAR model. Referring to the Gaussian noise term in Eq.(1), the residual of the TAR model can be represented as a mixture of Gaussian distributed random variables having a stationary marginal distribution [19]. The weighting factor of the mixture is defined as

$$\omega_m = (R_{m+h} - R_m)/R_K \tag{9}$$

for the sub-region m. The variance of region m is shown in Eq.(8). The weighted variance for a TAR model is obtained as

$$T_D([0,m][m,m+h]\cdots[m+k,K]) = \tag{10}$$
$$\omega_0\Phi^{(0)}(0,0) + \omega_m\Phi^{(m)}(0,0) + \cdots\omega_{m+k}\Phi^{(m+k)}(0,0)$$

for the delay value D. The TAR model having the minimum $T_D(\cdot)$ is selected for the best model. The AIC is used for order selection in the sub-AR model. For the sub-region m, the AIC is defined as follows

$$AIC(m,p) = (N_m \log(\sigma_u^{(m)^2}2) + 2p)/N_m \tag{11}$$

where p is the AR model order which is $1 \le p \le p_{\max}$. The $p_{\max}$ denotes the maximum sub-model lag size. Selected model order is p_m for region m if the $AIC(m,p_m)$ has the minimum AIC among the $p_{\max}$ cases. Thus p_m coefficients are used to calculate the roots of the polynomial as

$$1 - \sum_{i=1}^{p_m} a_i^{(m)} z^{-i} = 0 \tag{12}$$

where z is transform variable. If roots of Eq.(12) satisfy

$|\alpha_i^{(m)}| < 1$ then the AR model is stable. Therefore a suitable TAR model uses only stable AR processes.

Figure 5 illustrates the TAR model construction algorithm. Since the optimal thresholds are initially unknown, the histogram and scatter diagram of MPEG trace are used to select the initial thresholds. High activity regions in the histogram or dense clusters in the scatter diagram can provide useful information for the initial threshold points. More elaborate search technique for determining the thresholds based on K-mean clustering and dynamic invariance have been used to analyze VBR trace[15]. We found that the initial threshold selection using equally spaced amplitude intervals is reliable for the MPEG traces used here.

The maximum D and p_{max} are set to 5 throughout this paper. Index counters used in the algorithm are p, t and D. The p and D index respectively the sub-AR model order and time-delay. Initial value of D is set to 1 and it increases until D reaches to 5. The combination of sub-AR models for a T_D requires recursive function calls in the simulation program to efficiently retrieve the representative AR models [9]. The $L(K)$ in the figure represents the total number of representative sub-AR models in terms of the maximum threshold point K, and t is index counter for $L(K)$. Thus $L(K) = O(K^2)$. At each p iteration step $x(n)$ is classified and $a_i^{(m)}$ extimated in the given threshold interval, R_m and R_{m+h}. This portion of the algorithm takes the most time.

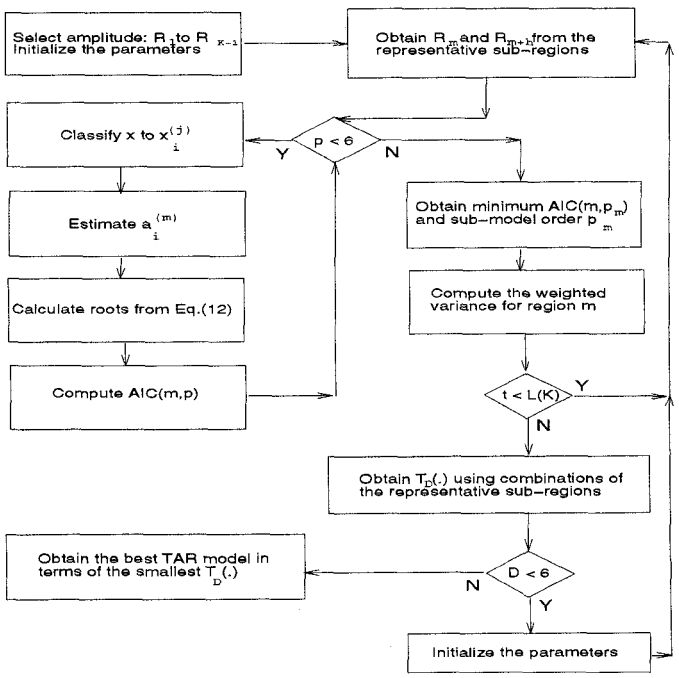

Figure 5: The TAR model construction algorithm

4. Model for a Full-Motion Video Trace

In this section, the TAR modeling procedure is applied to VBR MPEG video data shown in Figure 1. TAR models are constructed for each frame types of I, P and B. Composition process for the trace using the TAR models is implemented using the GOP pattern. The number of data points for I-, P- and B-frames are 3334, 10000 and 26666 respectively. The initial threshold points are determined by equally spaced amplitude ranges. Thresholds for I-, P- and B-frame types are respectively selected as $\{0, 50K, 60K, \cdots, 200K, P_I\}$, $\{0, 20K, 30K, \cdots, 120K, P_P\}$ and $\{0, 6K, 8K, \cdots, 40K, P_B\}$ where P_I, P_P and P_B are values greater than the peak rates for each frame type. Using the algorithm developed in the previous section, the optimal TAR models were retained for each frame type and shown in Table 2.

$T_D^{(I)}$	D_I	k_I	$T_D^{(P)}$	D_P	k_P	$T_D^{(B)}$	D_B	k_B
1.00	2	5	1.00	2	8	1.00	2	6
1.01	4	6	1.05	2	6	1.00	2	7
1.05	4	7	1.10	4	9	1.01	2	7
1.09	2	6	1.10	3	9	1.03	3	5
1.12	4	8	1.10	2	5	1.04	4	5

Table 2: The selected optimal TAR models for each frame type

The five best TAR models for each frame type are shown in the table in terms of the normalized T_D. All T_D's are normalized by the minimum value of its class. For the best cases in each frame type, D was found equal to 2. The k represents the number of threshold between R_0 and R_K for each frame type. The thresholds for the best TAR models are obtained as

$$R_m^{(I)} = (0, 60K, 70K, \cdots, 100K, P_I) \quad (14)$$

$$R_m^{(P)} = (0, 20K, 30K, \cdots, 90K, P_P) \quad (15)$$

and

$$R_m^{(B)} = (0, 6K, 8K, \cdots, 16K, P_B) \quad (16)$$

Some sub-region selections could not be evaluated due to the lack of a sufficient number of data points. The sub-AR models for the best TAR models in each frame type have sub-AR model orders as (1,4,1,4,1,4), (1,2,5,1,3,4,1,4,1), (1,1,4,1,5,1,5) for I, P and B respectively. From left to right, the sub-AR model orders are for intervals of lower thresholds to higher thresholds. In total 22 linear models are used for the MPEG full-motion video trace.

A MPEG trace was generated using the best TAR models of each frame type. All sub-AR models in each frame type have residuals of Gaussian random variables with zero mean and standard deviation $\sigma_u^{(m)}$. The residual series in Eq.(1) is obtained as

$$u^{(m)}(n) = \sigma_u^{(m)} g(n) \quad (17)$$

where $g(n)$ is Gaussian random process with N(0:1) and $\sigma_u^{(m)}$ is calculated from Eq.(8). Note that different residual series are obtained for each frame type and for each sub-region. Finally, a synthesis video trace $x(n)$ was attained after the composition process of frame types generated by the best TAR models.

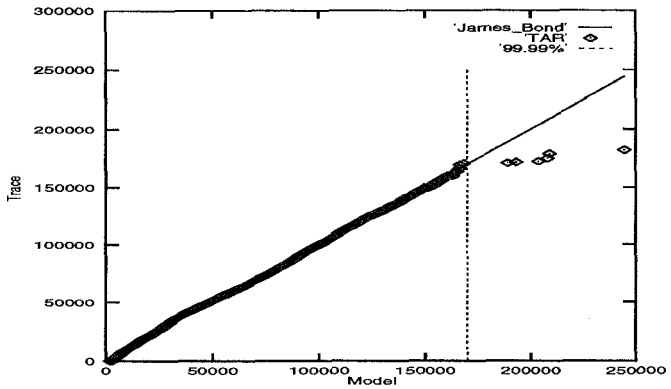

Figure 6: Q-Q plot for actual vs. TAR traces

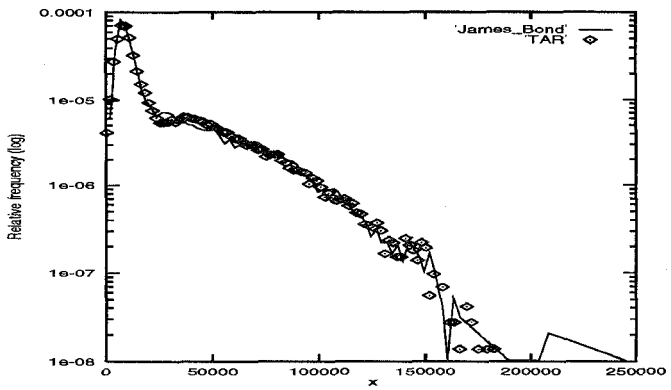

Figure 7: PDF for actual vs. TAR traces

The Q-Q plot and PDF are used to compare the synthesized and actual traces. In the Q-Q plot comparison shown in Figure (6), deviation from the line is observed in the amplitude range greater than 99.99% of the highest amplitude in the actual trace. This range is corresponds to the I-frames occurring at peak bit-rates. PDF comparison shown in Figure 7 demonstrates accurate TAR modeling below a frame size of 150,000 bits.

The TAR model captures the SRD in the local amplitude regions using low order AR models. Transition between the local regions depends on the thresholds and a time-delay amplitude. The $x(n)$ composed using the TAR model captures the LRD of actual video trace. The deterministic composition of each frame class can allow the TAR model to capture the LRD of the MPEG stream. The ordering between frame classes (I, P, B) has an impact on the LRD of the TAR model. Reference [10] has shown that LRD in the MPEG trace is captured by the composite of each frame type that is modeled using independent random processes. Figure 8 shows the results of ACF comparison.

5. Bit-Loss Rate Comparisons

The leaky bucket algorithm [12,15] will be used to estimate bit-loss rates. The number of bits $x(n)$ will be the input to the queue and the clock will be the frame counter. The maximum number of bits serviced during a frame period will be denoted by the constant drain rate. The drain-rate S can be retrieved dividing by the time duration of each frame. In our case this will be 1/30 of a second.

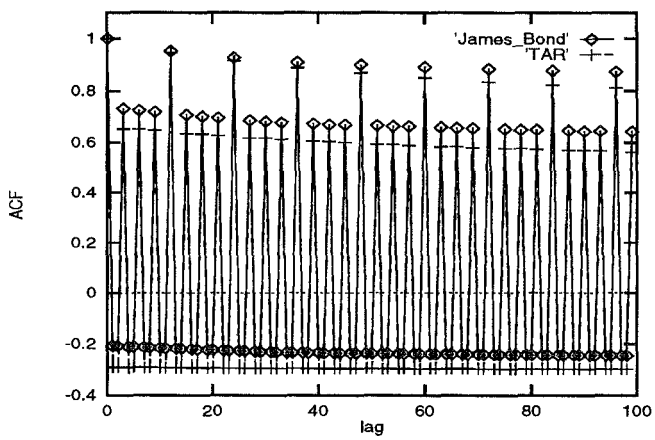

Figure 8: ACF for actual and TAR realizations

The maximum capacity of the buffer is M bits. The number of bits in the buffer at time n is equal to $B(n)$ where the auxiliary function $B(n)$ has an infinite capacity

$$B(n) = B(n-1) + x(n) - \min(B(n-1), S). \tag{18}$$

The bracket expression represents the number of bits serviced using a first-in-first-out (FIFO) discipline in the interval between $n-1$ and n. When $B(n)$ exceeds the buffer capacity M the number of bits exceeding the threshold are lost. The cumulative bit-loss $L(n)$ at time n is given by the expression

$$L(n) = L(n-1) + \max(0, B(n) - M). \tag{19}$$

The fraction of bits that are lost in N frames is equal to

$$\varepsilon = L(N) / \sum_{n=1}^{N} x(n). \tag{20}$$

Two parameters will be used in the presentation of the results. The first will be the bucket size given in seconds which are equal to the $M/S\delta t$ where δt is equal to 1/30. The second parameter is the ratio of the drain rate $S/\delta t$ and the average rate $\bar{S}/\delta t$. The average rate is a fixed parameter using a long-term average in the video sequence.

Figure 9 shows the bit-loss rates as a function of the drain rate $S/\bar{S}$ and the buffer size $M/S\delta t$ in seconds. As the buffer size increases, the difference between the actual and synthesized traces increased. Figure 9 illustrates the results of the simulation for bucket sizes equal to 0.01, 0.5 and 2.0 seconds. To obtain these results 40,000 frame samples were used. The figure compares the loss rate in a leaky bucket queue for the actual and TAR traces. The horizontal axis reflects the values of $S/\bar{S}$ and the vertical axis denotes the bit-loss rates. For bit-loss rates less than 10^{-3} the TAR model underestimates the loss rates.

The reason for disagreement is the improper modeling of outliers which correspond to frames having large bit-counts. As we have seen in the Q-Q plot comparison between actual and TAR models, less than or equal to 0.01% of the jump arrivals in I frames was not properly modeled.

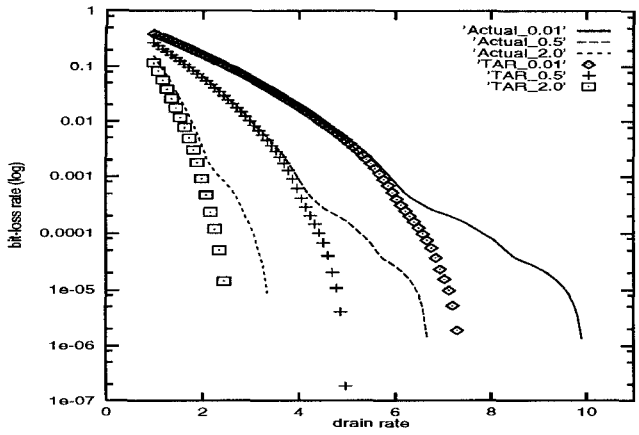

Figure 9: Bit-loss rate comparison

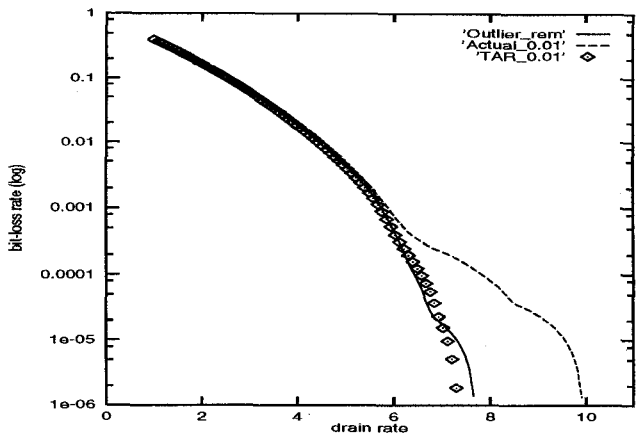

Figure 10:Bit-loss rate comparison with/without outliers

We tested the outliers effect on the estimation of the bit-loss rates using the TAR model. Only 5 frames whose bits are greater than 189,232 in the actual trace belong to the range of greater than 99.99%. To test this outlier effect, we removed the 5 frames whose bits are greater than 189,232 in the actual trace and measured the bit-loss rates using the actual and TAR traces with the bucket size 0.01 seconds. The result is presented in Figure 10. The bit-loss rates for the actual trace, the trace with the 5 frames removed, and the TAR trace are compared in the figure. A good agreement of bit-loss estimation is observed between the trace with the 5 frames removed and TAR trace.

In summary the TAR trace shows very accurate estimation for bit-loss rates for the entire range of drain rates and for large buffer sizes compared to the actual bit-loss rates without outliers. The outliers in the actual trace hinders accurate modeling of the bit-loss rates below the 10^{-3} range. These high jump bit rates occur very rarely in the actual trace (only 5 frames out of 40,000 frames). It is realized as a difficult task to include these very rare events in any modeling approach.

6. Conclusion

The results presented in this paper demonstrated that the developed TAR modeling process is effective in modeling VBR video traffic. The Q-Q plot and PDF were well matched to the actual VBR MPEG video trace. Bit-loss rates were accurately predicted using the trace composed from the TAR models of each frame type with the exception of the amplitude range belonging to outliers. Unmodeled high rate frames, which were only 5 frames among 40,000 frames were shown to greatly impact the bit-loss at high drain rates. Due to these frames, errors in bit-loss estimates were observed at drain rates which are greater than six times average bit rate at a bucket size of 2 seconds.

The method developed in this paper estimates accurate modeling parameters such as thresholds, a time-delay, and sub-AR model orders. The optimal thresholds and a time-delay point provide accurate models for each frame type. The model orders for stationary sub-AR models capture the SRD presented in the amplitude ranges. The switching between sub-regions is deterministic due to the thresholds and the time-delay point.

Acknowledgement

The authors wish to thank Prof. Kavitha Chandra of UMass Lowell for her thoughtful comments on this work.

References

1. Beran, J., R. Sherman, M.S. Taqqu, and W. Willinger, "Long-Range Dependence in Variable-Bit-Rate Video Traffic", IEEE Trans. Comm, vol. 43, no. 2, pp. 1566-1579, February 1995
2. De Prycker, M., Asynchronous Transfer Mode Solution for Broadband ISDN, 2nd edition, pp. 222-237, Ellis Horwood, 1993
3. Frater, M.R., J.F. Arnold, and P. Tan, "A New Statistical Model for Traffic Generated by VBR Coders for Television on The Broadband ISDN", IEEE Trans on Circuits and Systems for Video Technology, vol. 4, no.6, pp.521-526, December 1994
4. ftp:ftp-info3.informatik.uni-wuerzburg.de /pub/MPEG
5. ftp:tenet.berkeley.edu /pub/dbind/traces/README
6. Heyman D., A. Tabatabai, and T.V. Lakshman, "Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks", IEEE Trans. on Circuit and System for Video Technologies, vol. 2, pp. 49-59, March 1992
7. Heyman, D.P., and T.V. Lakshman, "Source Models for VBR Broadcast-Video Traffic", IEEE/ACM Transactions on Networking, vol. 4, no.1, pp.40-48, February 1996
8. Hwang, C., 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 ACM SIGCOMM '95, Cambridge, MA, pp. 114-125, August 1995
9. Jang, B.S., "Threshold Autoregressive Models With Application To Video Traffic", Doctoral Thesis, Computer Science Department, U. of Massachusetts Lowell, Low-

ell, MA, pp. 69-72, April 1997

10. Krunz, M. and S.K. Tripathi, "Scene-Based Characterization of VBR MPEG-Compressed Video Traffic", Technical Report, Department of Computer Science and Institute for Advanced Computer Studies, U. of Maryland, June 1996

11. LeGall, D., "MPEG: A Video Compression Standard for Multimedia Applications", Communication of the ACM, 34(4), pp. 46-58, April 1991

12. Lucantoni, D.M., M.F. Neuts, and A.R. Reibman, "Methods for Performance Evaluation of VBR Video Traffic Models", IEEE/ACM Trans.on Networking, vol. 2, no. 2, pp. 176-180, April 1994

13. Maglaris, B., D. Anastassiou, P. Sen, G. Karlsson, and J. Robinson, "Performance Models of Statistical Multiplexing in Packet Video Communications", IEEE Trans. Comm., vol. 36, no. 7, pp. 834-844, July 1988

14. Onvural, R.O., Asynchronous Transfer Mode Networks Performance Issues, pp. 122-129, 2nd edition, Artech House, 1995

15. Chandra, K. and A.R. Reibman, "Modeling two-layer MPEG-2 video traffic", H.L. Bertoni, Y. Wang and S. Panwar Eds., pp. 87-94, Plenum Press, NY 1996

16. Schwartz, M., Broadband Integrated Networks, pp. 58-67, Prentice Hall, 1996

17. Tong, H., Non-linear Time Series A Dynamic System Approach, Oxford Science Publications, 1993

18. Yegenoglu, F., B. Jabbari, and Y-Q. Zhang, "Motion-Classified Autoregressive Modeling of Variable Bit Rate Video", IEEE Trans on Circuits and Systems for Video Technology, vol. 3, no.1, February 1993

19. Zhao, Y., X. Zhuang, and S.-J. Ting, "Gaussian Mixture Density Modeling of Non-Gaussian Source for Autoregressive Process", IEEE Trans. on Signal Processing, vol. 43, no.4, pp. 894-903, April 1995