SAM: A Simplified Seasonal ARIMA Model for Mobile Video over Wireless Broadband Networks[1]

Abdel Karim Al Tamimi
Department of Computer Science and Engineering
Washington University in Saint Louis
aa7@cse.wustl.edu

Raj Jain
Department of Computer Science and Engineering
Washington University in Saint Louis
jain@cse.wustl.edu

Chakchai So-In
Department of Computer Science and Engineering
Washington University in Saint Louis
cs5@cse.wustl.edu

Abstract

Wireless broadband technologies like WiMAX[2] are spreading especially in areas where wired broadband is not expected to reach. Video streaming is continuously acquiring a larger share of Internet's traffic resulting in a need to have a reliable video traffic model. In this paper, we analyze several video streams compressed for mobile streaming and develop their optimal Seasonal ARIMA models. Although these optimal models are very different, we find that a simple model, which we call Simplified Seasonal ARIMA Model (SAM), represents all of the streams very well. This model is ideally suitable for video generation in mobile video simulation studies. We also present the parameter values suitable for such studies.

1. Introduction

Video streaming is one of the fastest growing applications on the web. Survey results have shown that 75 percent of the U.S. Internet users have watched an online video with the average person spending from 3 – 3.5 hours a month watching streaming videos, which represents a 29% increase from the last year. Paid video download revenues have reached 218 million dollars in the last year, and expected to reach 2.4 billion dollars in 2012. Revenues coming from the advertisements on streaming video and audio have reached an astronomical figure of 1.37 billion dollars [1, 2]. With the advent of wireless broadband services like WiMAX[2] the percentage of the broadband users are expected to rise especially in remote and rural areas. In addition to that, because of WiMAX's support of the mobility, more people are expected to use applications like Mobile TV as a source of information and entertainment on the go. These advances will result in a higher demand for video streaming.

An accurate video traffic model will facilitate a better understanding of the constraints of the network environment and its impact on video performance especially on time sensitive contents.

In this paper we present optimal models for a number of full-length movies and then observe how our simplified model, SAM, produces results that are very close to the optimal models. This one model is, therefore, ideal for workload generation in simulation studies.

The next section will describe the main features of MPEG video encoding. Section 3 describes the characteristics of the ARIMA model and the AIC index. Section 4 reviews some of the most important contributions and related works. Section 5 demonstrates and discusses the methodology of our analysis and the results that we have obtained. Section 6 concludes our results.

[1] This work was sponsored in part by a grant from the Application Working Group of WiMAX Forum.

[2] "WiMAX," "Mobile WiMAX," "Fixed WiMAX," "WiMAX Forum," "WiMAX Certified," "WiMAX Forum Certified," the WiMAX Forum logo and the WiMAX Forum Certified logo are trademarks of the WiMAX Forum.

978-0-7695-3454-1/08 $25.00 © 2008 IEEE
DOI 10.1109/ISM.2008.112

Proceedings of IEEE International Symposium on Multimedia (ISM 2008), Berkeley, CA, December 15-17, 2008.

2. MPEG Video Encoding

In this section, we explain the fundamentals of MPEG (Moving Picture Expert Group) encoding. This basic introduction will help justify our approach of using Seasonal ARIMA models.

MPEG is a collection of standards used to code both audio and video information. There are four main MPEG families and our emphasis is on MPEG-4 which is designed for low bit rate videos like web streaming media, and conventional videophone [3].

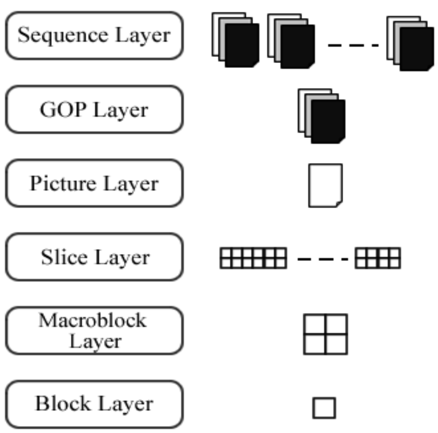

Figure 1. MPEG Video Hierarchy

MPEG layer hierarchy, as shown in Figure 1, consists of 6 different layers. A block is a group of pixels (8x8 pixels) that hold the visual information. To ease the computation, four blocks (16x16 pixels) are grouped together to form a macroblock. A single row of macroblocks in a video frame is called a slice. These slices are then grouped to form a video frame or a picture.

For compression, successive video frames are considered together as a group of pictures (GOP) that represents an independent unit in the video scene. GOP sizes vary depending on the encoding options, increasing GOP size results in a lower bit rate and a better compression, but it also results in a less robust compression. Sequence layer is comprised of a sequence of GOPs. A sequence layer can be thought of as a video scene or shot.

There are three types of compressed video frames: Intra-coded Frames or I-frames, Predicted frames or P-frames, and Bidirectional predicted frames or B-frames. An I frame, or Intra-frame, represents a reference frame and it is compressed independently. No information from other frames is used in the compression. Therefore, this frame can be decompressed even if other frames in the GOP are lost. P (Predicted) frames result from encoding a video frame by its difference from the prediction based on previous I frame or P frame.

B (Bi-directionally predicted) frames result from encoding a video frame by its difference from prediction using both the previous I or P frame and next I or P frame. P and B frames belong to Inter-frames group and are considerably smaller than I frames. But they do require larger buffers to accommodate the backward and/or forward predictions [4].

In order to allow the receiver to decode the frames as intended, the sequence of P and B frames is altered on the sending side. Figure 2 shows the relationship of the different frame types and how the transmission order of frames differs from the encoding order.

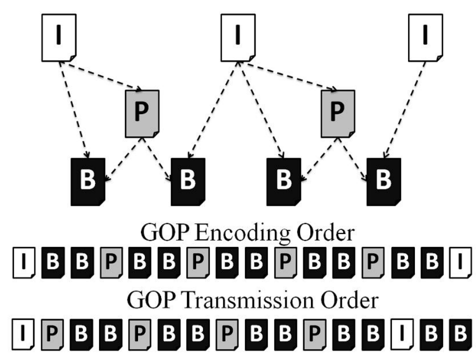

Figure 2. GOP Encoding and Transmission Orders

A GOP is usually presented as a sequence of frames starting with an I-frame then followed by a number of P and B-frames. One of the common patterns of GOP is G12B2, which means a GOP of size 12 and 2 B-frames between successive I and P frames. The encoding order for this sequence is: IBBPBBPBBPBB.

Because of the nature of MPEG encoding both long range dependence (LRD) and short range dependence (SRD) are present in frame size sequences. It is important for these relationships to be captured in any valid video traffic model. See [4, 18, 19] for more background information on MPEG compression.

3. ARIMA and Seasonal ARIMA Models

ARIMA (Auto Regressive Integrated Moving Average) is a general time series model that consists of three main components: an autoregressive component p that represents the dependency of the current data value to previous values, a moving average component q, which is also called a smoothing model and is useful in decreasing the local noise to allow better modeling and prediction, it depends on the previous error values, and lastly: a differencing part d that helps make the process stationary. Here, p, d, and q are non-negative integers.

Equation 1 shows how ARIMA(1,1,1) model is presented in terms of the current and previous data values.

$$y(t) = w(t) + y(t-1) + \phi.(y(t-1) - y(t-2)) - \theta.w(t-1) \quad (1)$$

Where ϕ is the coefficient of autoregressive (AR) part, θ is the coefficient of moving average (MA) part.

Seasonal ARIMA models are used for data series that exhibit periodic behavior. This seasonal behavior shows a continual repetition after a certain period. Seasonal ARIMA is described as ARIMA (p,d,q) × (P,D,Q)s. Here, P, D, and Q represent the order of seasonal AR model, seasonal differencing, and the order of seasonal MA model, respectively. “s” represents the period length of a season. For example, monthly data of a process that repeats yearly has a season period of 12 [5].

One of the most used evaluation methods for models is Akaike’ Information Criterion (AIC), which was developed by Hirotsugo Akaike to calculate the goodness of fit of a statistical model. AIC is a measure of the goodness of a model. It represents a tradeoff between the complexity of the model and how closely it fits the data [6]. AIC depends on (p,q,P,Q) values, changing these value results in different AIC values.

4. Previous and Related Work

Several models to represent VBR (Variable Bit Rate) MPEG traffic have been proposed in the last decade. Some of the models proposed are based on Markov chain models [7, 8, 9], which are known for their inefficiency of representing LRD characteristics of MPEG traffic. Though improvements have been made to get better results [8].

Others were based on the fact that MPEG traffic is a self-similar process [10]. More sophisticated approaches to model the traffic using wavelets have been proposed [11]. Due to the high influence of LRD on the MPEG traffic, multiplicative processes like Fractional ARIMA (FARIMA) have been considered, which have been shown to be better than wavelets and Gamma-Beta AR (GBAR) models [12].

Seasonal ARIMA also has been used before to model network traffic and especially the GSM traffic [13]. Most of the models proposed before were based on video traces coded with obsolete codecs; hence new models that can capture the new characteristics of the new codecs are required.

Most of the proposed models except a few [8, 13, 14], are based on short video traces. That might lead to the skepticism about the validity of the model since the level of texture and motion varies between different scenes, affecting the statistical characteristics of the model.

The most important drawback of the previous models is that they offer a specific model for each movie or a movie scene without providing a general approach to tackle the video modeling problem. They also suffer from a high complexity level of analysis and implementation due to a large number of parameters in these models [15].

We conducted a study on different video traces and found that a composite model consisting of separate models for I, P, and B streams is the optimal approach to model video scene traffic. During that study, we also found that a single model is also capable of achieving comparable results with less complexity and analysis time. As a result of our analysis a general model using Seasonal ARIMA was introduced that is capable of fitting all the tested video traces. In this paper we test the feasibility of a general model code named SAM that is capable of capturing all the characteristics of MPEG video traffic based on full length video traces.

5. Analysis and Results

Our goal is to have a general model that is capable of capturing the main aspects and characteristics of video traffic. To achieve this goal we considered 6 full movie traces available from the Video Traces Research Group [16]. The selected movies are: Lord of The Rings Trilogy and Matrix Trilogy.

We focused on the video traces compressed with the specifications that are more likely to be played on a mobile device. Hence, we chose the following specifications:

- MPEG-4 Part 2: This encoding standard, as we have mentioned before, supports low bit rate encoding and it is the most common standard used for web streams.
- Advanced Simple Profile (ASP): this profile provides the best coding performance, supports wide range of bit rates, and supports B-Frames for better video compression [17].
- CIF size (352 × 288): one of the most common screen sizes used in mobile devices.
- Frames rate 25fps (as provided by the Video Traces Research Group). The analysis is valid for 30 fps video also.

The total number of frames in these movies is around 188 thousand frames for Matrix trilogy, and around 266 thousand frames for LOTR trilogy. The frame size statistics in the analyzed traces, as shown in Table 1, indicates a high level of motion (high standard deviation), a high level of texture (high mean) in these frames. There is also a good range of variation between the different movies.

Our first step in video traces analysis is to find all the statistical information that helps reveal the characteristics of the video traces. Hurst index indicates the existence of LRD since it exceeds 0.5.

Table 1 Statistical Analysis of Video Traces

Movie	Standard Deviation	Mean	Variance	Hurst Index
LOTR 1	9594.778	9342.26	92059757	0.9158
LOTR 2	11178.38	11481.00	124956269	0.9158
LOTR 3	10794.25	11145.63	116515800	0.9233
Matrix 1	7946.338	7348.922	63144295	0.9011
Matrix 2	10687.00	9508.467	114212020	0.9147
Matrix 3	12701.56	10522.08	161329728	0.9253

A closer look at the trace files shows a repeated pattern as shown in Figure 3. This pattern reveals the seasonal characteristic of MPEG encoded videos. As can be noticed from Figure 3 there is a general pattern continually repeated with a period equal to the GOP size. In our case, it is 12 since we used G12B2 coding pattern. The seasonal part can also be computed using a simple mathematical algorithm to compare frames sizes to identify I-frames.

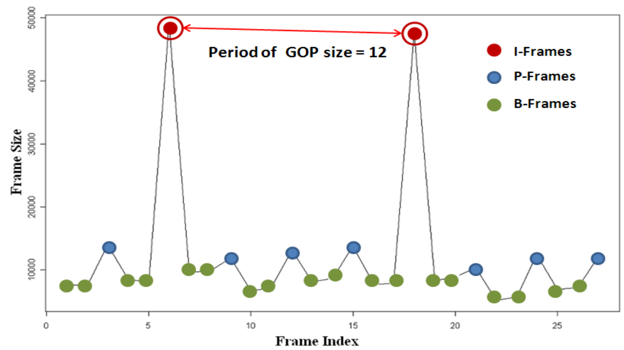

Figure 3. Seasonality in Video Traces

These results encouraged us to use Seasonal ARIMA, which considers both SRD represented by the relationships between the different frame types in one GOP, and LRD represented by the relationship between successive GOPs.

We found that Seasonal ARIMA models provide a very good fit to the measured traces. Although each trace had a very different optimal Seasonal ARIMA model, what surprised us was that a single simplified model was also very close to the optimal models in all these cases. This simplified model is a $(1, 0, 1)\times(1, 1, 1)_{12}$ Seasonal ARIMA model which was chosen after several tests.

We have conducted several statistical tests over the chosen full video traces. We had to make sure that the general model or SAM is a valid model and as close as possible to the "optimal model" for the full movie traces. The optimal models were Seasonal ARIMA models that were obtained using extensive analysis of video traces as described in [5].

As mentioned before, AIC index is a measure of the goodness of a model. We, therefore, compared the AIC values for the general model, i.e. SAM, to the AIC values for the optimal model for each of the full traces. Table 2 shows results that confirm that the SAM is within 1% of the optimal models in all cases.

Table 2 AIC Comparison between SAM and Optimal Models

Movie	AIC (Optimal)	AIC (SAM)	Difference% (|S-O|/O)
LOTR 1	15209108	15214697	0.036%
LOTR 2	18195617	18220707	0.137%
LOTR 3	16495282	16515722	0.123%
Matrix 1	11222747	11227109	0.038%
Matrix 2	20321203	20361456	0.198%
Matrix 3	34489730	34764677	0.797%

Table 3 Statistical Comparison between SAM and Optimal Models

Optimal Model				
Movie	**MAE**	**MARE**	**SNR^{-1}**	**NMSE**
LOTR 1	1850.149	0.3256206	0.0848033	0.1652013
LOTR 2	2038.680	0.2806260	0.0708604	0.1456091
LOTR 3	1940.064	0.2889833	0.0685161	0.1415653
Matrix 1	1553.833	0.3700388	0.0957917	0.177721
Matrix 2	2126.052	0.3839772	0.0993043	0.1779137
Matrix 3	2830.622	0.3941804	0.1267721	0.2137702
SAM				
Movie	**MAE**	**MARE**	**SNR^{-1}**	**NMSE**
LOTR 1	1851.281	0.3240269	0.0848344	0.1652620
LOTR 2	2043.132	0.2799332	0.0709581	0.1458099
LOTR 3	1944.378	0.2888479	0.0686010	0.1417407
Matrix 1	1553.584	0.3694246	0.095829	0.1777901
Matrix 2	2132.762	0.3864979	0.0995010	0.1782661
Matrix 3	2845.982	0.3957961	0.1277827	0.2154743

We then conducted extensive statistical analyses to further compare SAM with the optimal models. Table 3 shows that for each of the statistical measures considered, SAM is very close to the optimal models for all movies. The statistical measures shown in Table 3 are: Mean Absolute Error (MAE), Mean Absolute Relative Error (MARE), inverse of Signal to Noise Ratio (SNR^{-1}), and Normalized Mean Square Error (NMSE). Other statistical measures (such as Peak to Signal Noise Ratio [PSNR]) have been considered and have yielded similar results as shown in Table 3.

To further validate the SAM model, we conducted several visual comparisons of the data generated using SAM model and actual traces. This includes comparing the frame size traces, autocorrelation function graphs, and cumulative distribution functions. Each of these graphs showed that SAM is a very good and valid model. Figures 5.a, 5.b and 5.c show the visual validation results for Matrix 1 movie. Due to the limitation on the paper size we are not able to present the graphs for the other movies but they are very similar to the ones shown here.

Graphs in Figure 4 show that SAM is capable of capturing both statistical and behavioral characteristics of the modeled data. It is capable of capturing the complex relationships between encoded MPEG video frames.

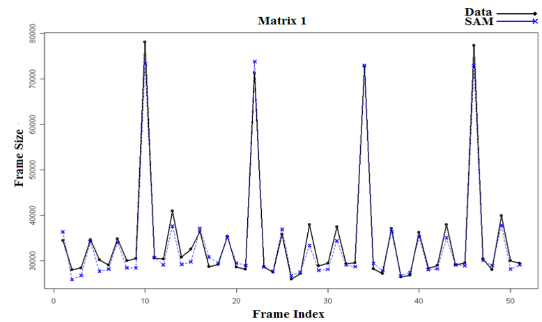

Figure 4.a Video Trace Comparison

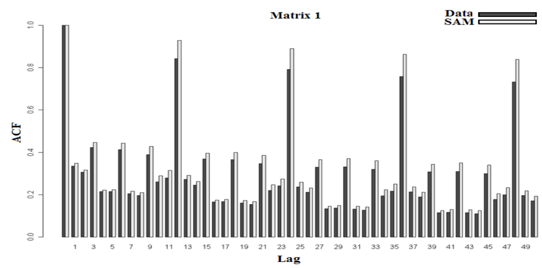

Figure 4.b ACF Comparison

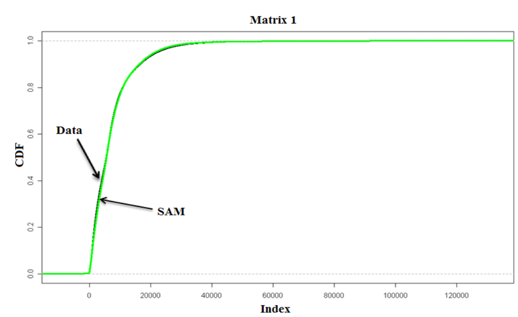

Figure 4.c CDF Comparison

The SAM model has four parameters: autoregressive (AR) parameter, moving average (MA) parameter, seasonal autoregressive (SAR) parameter, and seasonal moving average (SMA) parameter. The variations in these parameter values are presented in Table 4 by showing the min-max range of each parameter. It is interesting to note that there is little difference among the parameters for various movies. A single set of parameter values (shown by the mean) is a good set of parameters to use for video traffic generation.

Table 4 General Model Parameters for all Video Traces

	AR	MA	SAR	SMA
Mean	**0.93**	**-0.67**	**0.21**	**-0.86**
[Min, Max]	[0.924, 0.938]	[-0.691, -0.637]	[0.1, 0.271]	[-0.895, -0.805]

The next step is to use our conclusions and findings in this paper to create a video traffic generator that is able to mimic the behavior of an MPEG video stream. Our aim is to develop a video traffic generator for WiMAX NS2 package, which will help eventually to ease and facilitate a better understanding of WiMAX network environment and QoS support challenges.

6. Conclusion

In this paper we have presented a general model called SAM for full-length video traces using Seasonal ARIMA with the emphasis on video traffic expected on wireless broadband (like WiMAX) enabled mobile devices. We have shown through our extensive analysis both visually and statistically that SAM is a valid and accurate model for all video traces that we analyzed. We have also found that SAM parameters for these video traces, in spite of their variation in texture and motion levels, are close to each other. These results support the idea of a one general model with common parameters set that can represent most of video traffic. These results are a big motivation for our next step to create a video traffic generator for a WiMAX NS2 simulator.

7. References

[1] ComScore Press Center, Press Release, URL=[http://www.comscore.com/press/release.asp?press=2002]

[2] Liz Cannes, "Need-to-Know Web Video Stats: Traffic, Rentals, Revenues, UGC," URL=[http://newteevee.com/2008/01/17/need-to-know-web-video-stats-traffic-rentals-revenues-ugc/]

[3] Wikipedia, "MPEG Standards," URL=[http://nostalgia.wikipedia.org/wiki/MPEG]

[4] Seeling, Patrick, Fitzek, Frank H.P., and Reisslein, Martin , Video Traces for Network Performance Evaluation, Springer 2007.

[5] Duke University, Decision 411 Course, URL=[http://www.duke.edu/~rnau/411home.htm]

[6] Wikipedia, Akaike information criterion. Wikipedia page,URL=[http://en.wikipedia.org/wiki/Akaike_information_criterion]

[7] Dawood, A.M., and Ghanbari, M. "Content-Based MPEG Video Traffic Modeling," IEEE Transactions on Multimedia, Volume 1, Issue 1, Mar 1999, Page(s):77–87.

[8] Yang Sun, and Daigle, J.N, "A Source Model of Video Traffic Based on Full-Length VBR MPEG4 Video Traces," IEEE Global Telecommunications Conference, 2005. GLOBECOM'05, Volume 2, 28 Nov.-2 Dec. 2005, 5 pp.

[9] Lazaro, O., Girma, D., and Dunlop, J, "H.263 Video Traffic Modeling for Low Bit Rate Wireless Communication," 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2004. PIMRC 2004. Volume 3, Issue, 5-8 Sept. 2004 Page(s): 2124 – 2128.

[10] De la Cruz, L.J., Pallares, E., Alins, J.J., and Mata, J. "Self-Similar Traffic Generation Using A Fractional ARIMA Model. Application to the VBR MPEG Video Traffic," Proceedings of SBT/IEEE International Telecommunications Symposium, 1998 (ITS'98), Volume 1, 9-13 Aug 1998, Page(s):102 - 107.

[11] Lazaro, O., Girma, D., and Dunlop, J., "Real Time Generation Of Synthetic MPEG-4 Video Traffic Using Wavelets," Proceedings of IEEE VTS 54th Vehicular Technology Conference, Fall 2001 (VTC 2001), Volume 1, Page(s): 418 - 422.

[12] Lazaro, O., Girma, D., and Dunlop, J. "A Study of Video Source Modeling for 3G Mobile Communication Systems," Proceedings of First International Conference on 3G Mobile Communication Technologies, 2000, Conf. Publ. No. 471, Page(s):461 – 465.

[13] Yantai Shu, Minfang Yu, Jiakun Liu, and Yang, O.W.W. "Wireless Traffic Modeling and Prediction Using Seasonal ARIMA Models," Proceedings of ICC'03, Volume 3, 11-15 May 2003 Page(s): 1675 – 1679.

[14] Hai Liu, Ansari, N., and Shi, Y.Q. "A Simple Model For MPEG Video Traffic," Proceedings of ICME 2000, Volume 1, Page(s):553 - 556.

[15] Xiao-dong Huang, Yuan-hua Zhou and Rong-fu Zhang, "A Multiscale Model for MPEG-4 Varied Bit Rate Video Traffic," IEEE transactions on broadcasting. Volume:50, Issue: 3, page(s): 323- 334.

[16] Video Traces Research Group. URL=[http://trace.eas.asu.edu/]

[17] MPEG-4 Industry Forum. MP4 SP/ASP. URL=[www.mpegif.org/public/documents/vault/m4-out-30037.pdf]

[18] Patrick Seeling, Martin Reisslein, and Beshan Kulapala, "Network Performance Evaluation Using Frame Size And Quality Traces of Single and Two Layer Video: A Tutorial," IEEE Communications Surveys and Tutorials, Third Quarter 2004, Vol. 6, No. 2, Pages 58-78.

[19] Geert Van der Auwera, Prasanth T. David, and Martin Reisslein, "Traffic Characteristics of H.264/AVC Variable Bit Rate Video," IEEE Communications Magazine In print, 2008.