

Video Source Modelling for ATM Networks

John Murphy and Jerry Teahan *[†]

[†]Electronic Engineering Department
Dublin City University
Ireland

There is a need to model telecommunications sources for network design and planning, such as the investigation of various congestion control mechanisms. Traditionally it was sufficient to model voice traffic for the telephone network and data traffic for computer networks. With the emergence of ATM networks there will be a mixture of voice, data, video and still image traffic on one network. The sources may take advantage of the ability of ATM to provide bandwidth on demand by supporting variable bit rate traffic. In particular we have investigated video coding schemes, especially the H.261 compatible ones, with both one and two layer codecs. The one layer codec presented in this paper is based on a discrete time Markov chain. The two layer model is modelled by its individual layers. Both are derived from an analysis of video conference data. The models are simulated using a commercial discrete event simulation tool. Evaluation of the characteristics of each source are presented, as are the capabilities and limitations of the models. Comparisons are drawn between actual data and the models. This is part of an overall objective to gather source models for most services that use ATM networks.

1. INTRODUCTION

The Asynchronous Transfer Mode (ATM) network is widely accepted as being the broadband network of the future [5] [8]. ATM allows the design of networks that give efficient and flexible allocation of resources by only allocating the amount of the resource that is needed. This ability of ATM networks to support variable bit rate traffic can be used to meet the dramatically changing traffic requirements of many applications such as video telephony or full motion video.

New important issues related to ATM networks, such as congestion control and cell loss probability, have to be investigated. In addition even though the ATM concept provides 'bandwidth on demand' there is a need to ascertain the network's efficiency or throughput. Performance evaluation of ATM broadband networks can be addressed by simulation studies using realistic and accurate models. This paper presents our contribution to the area of Video Traffic Characterization (VTC). We will present video source models which are based on analysis of real videophone type data and discuss the results of the simulations carried out using these models. While the results we produce are dependent

*All correspondence should be addressed to the first author.
Tel. +353 1 704-5444 Fax. +353 1 704-5508 email murphyj@eeng.dcu.ie

on the model of the video codec used, the method used to design the model is general and can be adapted to any codec design.

The video data, on which the models in this paper are based, were generated as in [2] using a H.261 one and two layer codec. This codec produced a data set consisting of bit rates for each video frame. The first task in producing accurate video source models was to characterize this data and this is addressed in Section 2. The design issues surrounding a video model in the context of the design of ATM networks and the investigation of the congestion control schemes is presented in Section 3. Simulation was carried out using a commercially available discrete event simulator called SES/workbench and is described in Section 4. This package provides a set of software tools for graphical animation and simulation of models. The accuracy and limitations of our model are discussed by considering that the number of ATM cells produced (on a per frame basis) during simulation statistically matches the number of cells per frame contained in the video data. This is presented in Section 5. Finally in Section 6 there is some discussion and plans for future work.

2. DATA CHARACTERIZATION

Video, like voice, is naturally bursty [1] [3] but can be manipulated to give a constant output bit rate. To gain the maximum efficiency from an ATM network only the real amount of information should be sent. This will change as the information content of the picture changes and so a variable bit rate output from the codec is expected. To gain maximum compression a single one layer codec could be used. Sudden movements in the picture or scene changes will produce large information transfers and so will produce high bit rates. Small changes from frame to frame will result in low bit rates.

In ATM networks it is sometimes hard to characterize the sources. Performance evaluation of the network becomes difficult and subsequent guarantees of quality of service (QOS) are hard to achieve [7]. However a QOS for a constant bit rate source might be easier to guarantee than that of a variable source. This has led to the development of two layered codecs [2]. One layer has a constant bit rate, called the base layer, and delivers a minimum QOS to the user. The second layer contains the enhancement to the base layer and is bursty in nature. This improves the quality of the picture to the user. The relevant quality of pictures of the two schemes is subject to the coding involved in the codecs, the bandwidth available and the probability of cell loss.

The empirical data on which the models presented in this paper are based were generated using algorithms implemented as in [2]. These are one and two layer implementations of the H.261 codec standard. This video compression standard can handle varying qualities of service from low quality video telephony (64 kb/s) to video conferencing (384 kb/s) and higher quality video (2 Mb/s) using a bit rate defined by $p \cdot 64 \text{ kb/s}$ where $p=1 \dots 31$.

Some of the main elements of H.261 include interframe and intraframe prediction, motion compensation, discrete cosine transformation (DCT), quantization of DCT coefficients and run length coding of the DCT coefficients. Interframe prediction takes advantage of the fact that consecutive frames will be quite similar allowing just the difference between frames to be transmitted rather than the entire frame data. In contrast intraframe prediction occurs when consecutive frames are substantially different, for ex-

ample at the start of a video sequence or a sudden motion/scene change. The video sequences considered here are of just a single person talking with some zooming motion. This is more typical of a video phone type picture than video conferencing because there is little changing of scenes and only moderate motion. The spatial resolution of the five minute sequence was CIF (384, 288) and the temporal resolution was 25Hz. In order to capture the data set for the encoded video sequences, on a per frame basis, the outputs of the codecs were sampled every 0.04 seconds. This produced approximately 7500 samples to be considered. It can be seen from Figure 1 below that the bit rate of this one layer codec varies considerably over the time interval. This short sequence and the type of picture that was used are limiting factors for the reliability of the models produced.

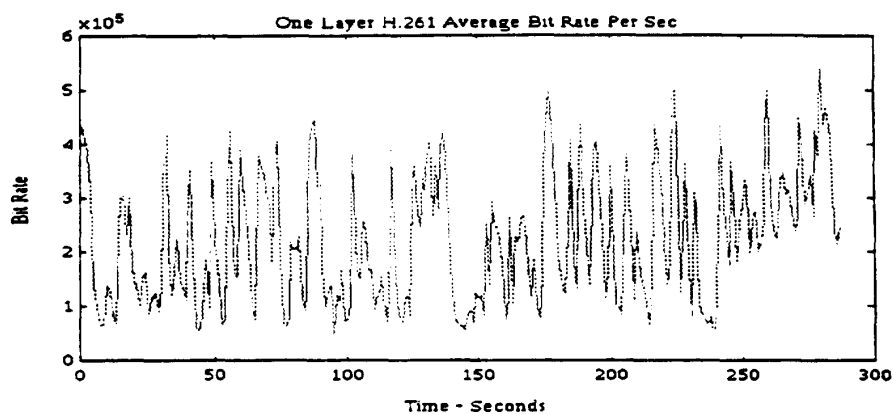


Figure 1. Bit Rate Of One Layer H.261 Codec

Even though the H.261 standard is complied with there is scope for different codecs. The two codecs discussed in this paper are the one layer 256 kb/s source and the two layer 384 kb/s source. The main characteristics of the sources are :

	One Layer	Two Layer Base Layer	Enhancement Layer
Mean Bit Rate (in kb/s)	222	369	39
Standard Deviation	Mean	0	High

3. MODELING THE VIDEO SOURCE

The one and two layered codecs were considered separately. For the one layer model a discrete state Markov chain was designed. This follows from work which verified that models based on multi-state Markov chains are sufficiently accurate for traffic studies [4] [9]. It was decided to take the frame as the basic unit of time for the chain and the ATM cell as the basic quantity. The objective is to capture the burstiness of the source by neither overestimating or underestimating the number of cells in each frame.

The Markov chain model was produced by carrying out the following analysis on the empirical data. The first issue was to examine the probability density function (PDF) of the size of frames produced. A comparable measure is the number of frames versus the frame size normalized to bits per second as is shown in Figure 2 below.

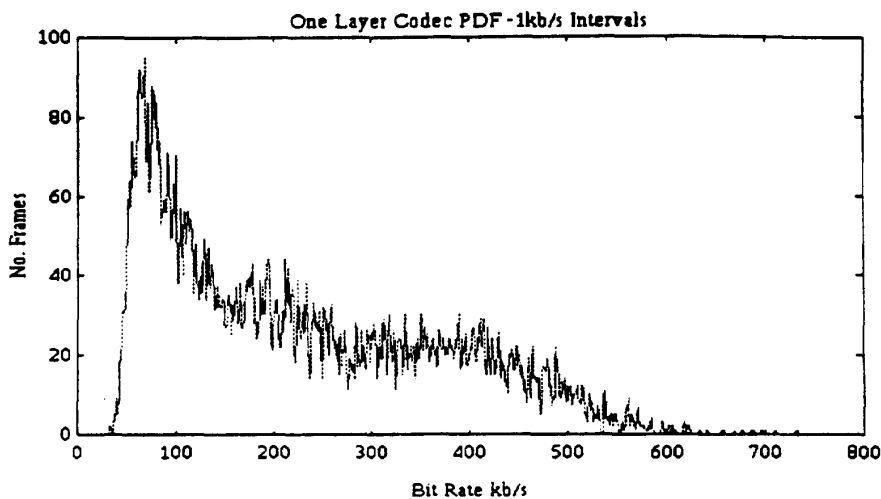


Figure 2. Shape Of One Layer Codec Probability Density Function (PDF)

A few criteria were used to select the states of the Markov chain. It was felt that at least five states would be needed to capture the shape of the PDF but the number could not be too large as the model would become complicated. It was felt that all states should be close to being equiprobable so that for short simulation runs they can all be reached. It was noticed that if there were a small number of states then a birth death process could be achieved. However if the number of states were larger, the assumption of a birth death process failed. The eventual number of states chosen was nine, which is a compromise between keeping a small number of states and also trying to have a birth death process. The birth death process greatly simplifies the modelling of the Markov chain as will be seen in Section 4.

Careful selection of the nine states resulted in the following transition matrix. There were small entries at a distance more than one from diagonal. The assumption made was to neglect these values. However the entries in the matrix must sum to unity and so a method of proportioning the excess to the other entries was found. This was done by starting at either end of the Markov chain and conditioning on only the allowed transitions to the nearest neighbors. By starting at either end there would be two slightly different matrices. The eventual matrix chosen is shown in Figure 3.

The probabilistic nature of transitions between these states is represented in matrix format by generating a count of all the transitions in and out of a particular state. This matrix is normalized in order to obtain the transition probabilities between states in the

	State0	State1	State2	State3	State4	State5	State6	State7	State8
State0	0.81	0.19							
State1	0.18	0.59	0.23						
State2		0.19	0.60	0.21					
State3			0.18	0.57	0.25				
State4				0.22	0.57	0.21			
State5					0.22	0.60	0.18		
State6						0.17	0.75	0.08	
State7							0.17	0.81	0.02
State8								0.55	0.45

Figure 3. Transition Probability Matrix For The One Layer Codec

Markov-chain model p_{ij} where ij are the row column elements of the normalized transition matrix.

To ensure that the states were almost equiprobable states were chosen so that the total frame count in that state did not exceed 15 percent of the total frame count. We also attempted to obtain a matrix that contained transitions of length at most one state. This simplification of the Markov-chain to a 'birth death process' greatly simplifies the complexity involved in simulating the model. Within each state there is an upperbound and lowerbound on the bit rate of the source. This can be converted to an allowable number of ATM cells by conversion to the frame size and then fragmenting that frame size into SAR-SDU's and by the AAL to SAR-PDU's. This SAR-PDU joins with a five byte header to form an ATM cell. Therefore for each state there are integer numbers of upperbound and lowerbound cells allowable in that frame time. The actual number picked is uniformly distributed across that interval.

The two layer codec was modelled differently as each layer was modelled separately. The PDF of the two layers is shown in Figure 4. The base layer is almost a constant bit rate and was modelled by just one state. Within that state the bit rate is uniformly distributed from 350 kb/s to 388 kb/s giving the correct mean rate of 369 kb/s.

The enhancement layer could have been modelled by a markov chain similar to the one layer codec model. However in this paper we investigate the possibility of modelling the layer by a hyper-exponential distribution. The hyper-exponential function produces a mean bit rate of 39 kb/s. To fit the distribution to the data a standard deviation of 124 kb/s was chosen.

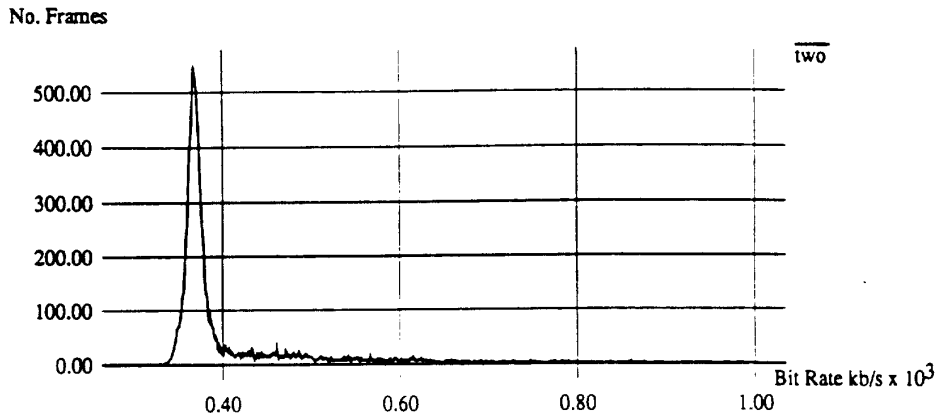


Figure 4. PDF For The Two Layer Codec

4. SIMULATION MODEL

The source models were simulated using SES/workbench [10], a discrete-event simulator that allows hardware and software simulation. Models are built using a top down hierarchical approach. At the top level variables and other parameters used in the model are declared. At the bottom or sub-model level a graphical representation of the models structure may be built using a library of standard nodes, and directed graphs. The models in this paper were mainly created by use of the graphical user interface. SES/workbench compiles the graphical code to C and creates an executable. The simulation execution platform was a cluster of Sparc-10 workstations. The runtime for the simulations was about 10 percent of real time, that is one hour of a source would only take 6 minutes execution time.

Simulating the one layer model is summarized in Figure 5 below. The basic time unit is a frame time or 0.04 seconds. The birth death process makes the model much simpler because instead of having 64 arcs connecting all the states to each other there are only 16 arcs.

Executing the model produces a number of ATM cells based on the range of bit rates that a particular state represents. As shown in Figure 5 the model can enter a new state at a rate of 25 per second. this emulates a codec running at 25 frames per second. The probability of a transition between states is determined by the normalized transition matrix. On entering a state the model produces a number of ATM cells. The number produced depends on the state being entered and is uniformly distributed across the range of attainable values in that state. Since the empirical data set for the coded video sequence presented the data on a bit rate per second format for each frame it was necessary to convert this data to ATM cells. The AAL 3/4 was used to show a possible implementation

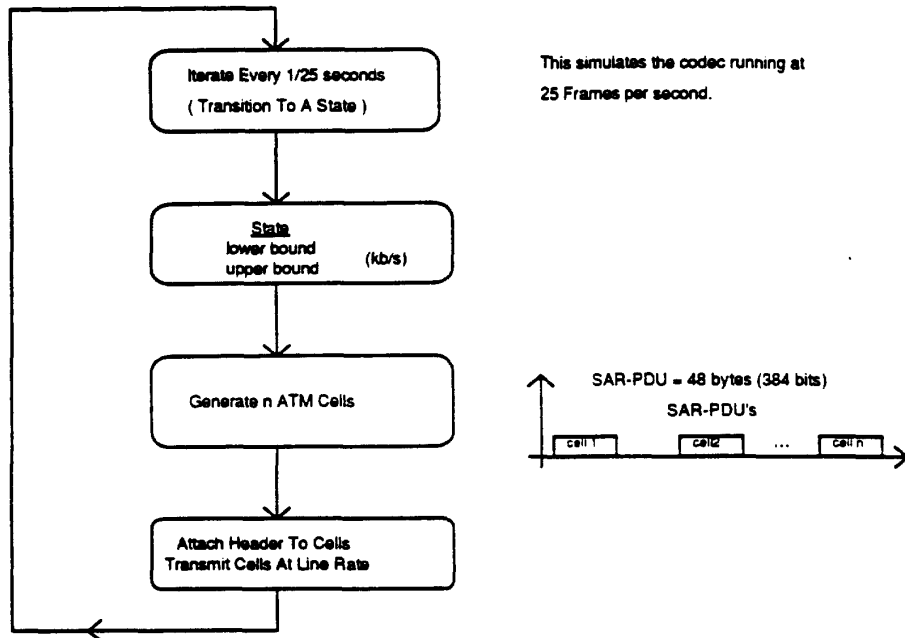


Figure 5. Process For Simulating The One Layer Codec

with 44 bytes of information per cell. A certain amount of shaping is also used for the generation of the ATM cells in that they are generated equally spaced in time over the frame duration.

Implementing the two layer model was quite similar to the one layer model. In this case however there is not a state diagram but rather a single state for the base layer and a state for the enhancement layer. A problem occurred again due to the discrete nature of the ATM cells generated. This resulted in a smaller mean being produced than expected. This was because the rounding to an integer number was rounding down. For the enhancement layer it was necessary to increase the mean of the hyper-exponential to 43 kb/s to produce a real mean of 39 kb/s. Both the enhancement and base layers run at the same time combining to form the two layer codec output.

5. MODEL VALIDATION

The objective here was to check that the models presented in this paper were producing data which was statistically comparable to the empirical data on which the models are based. A first approach was taken by visually comparing the PDF's of the empirical data with that of the models. However due to the cell structure of the mod-

els there is a discrete effect on the PDF, and so visual comparison was difficult. It was noted however that some edge effects were taking place at the boundaries of the states where a certain cell size was being favorably treated by being in two states. This was duly eliminated from the model. The basic measurements that were observed were the moments of the models and these are compared below for the one layer model.

Parameter	Empirical Data	Model
Mean (in b/s)	216 E+3	223 E+3
Second Moment	2.08 E+10	2.22 E+10
Root of Second Moment	1.44 E+3	1.49 E+3
Third Moment	1.9 E+15	2.2 E+15
Cubic Root of Third Moment	1.24 E+3	1.3 E+3

These results show that statistically the model is producing similar moments to the empirical data. To validate the model further, the correlation functions should be taken and compared for the model and the data. For the two layer model the mean bit rate produced is chosen to be identical to the empirical data and the deviation is similarly chosen.

6. DISCUSSION AND CONCLUSIONS

The models presented in this paper adhere to the H.261 video coding standard and have been implemented by a discrete event simulator. This will allow the sources to be used on many ATM network issues such as congestion control strategies [6]. The accuracy of the models in representing the real codecs is shown for the one layer codec by comparing the moments of the bit rates and these are seen to be compatible. The accuracy of the models is only as good as the accuracy of the codec algorithms that were implemented and the length of time the data was collected for. To improve the accuracy a longer time span should be taken. It is also likely that the algorithms will change over time and so it is not so much the exact model but the method of deriving the model which is important.

As the models will be used to investigate various ATM network issues rather than investigate the video sources themselves the exact algorithm used for the codec is not critical. It would be useful to design more video models and to test congestion control schemes so that they are not dependent on the type of source model being used to simulate a network.

Acknowledgments

We are grateful for the help and advice given by Sean Murphy on the codec design and data output. We acknowledge Prof. Charles McCorkell, Dublin City University, for his continuing encouragement and support for this work.

REFERENCES

1. P. T. Brady, 'A Statistical Analysis of On-Off Patterns in 16 Conversations', *Bell Syst. Tech. J.*, pp 73-91, 1968.
2. COST, 'Redundancy Reduction Techniques for Coding of Video Signals in Multi-Media Services', COST 211ter, Compendium ATM, Coding Procedures, Simulation Subgroup, pp 1-83, May 1993.
3. J. G. Gruber, 'A Comparison of Measured and Calculated Speech Temporal Parameters Relevant to Speech Activity Detection', *IEEE Trans. Commun.*, Vol. COM-30, No. 4, pp 728-738, 1982.
4. D. P. Heyman and T. V. Lakshman, 'Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks', *IEEE Trans. Circuits and Systems For Video Tech.*, Vol. 2, No. 1, pp. 49-59, March 1992.
5. I. M. Leslie, D. R. McAuley and D. L. Tennenhouse, 'ATM Everywhere ?', *IEEE Network*, Vol. 7, No. 2, pp 40-46, March 1993.
6. J. Murphy and L. Murphy, 'Bandwidth Allocation By Pricing In ATM Networks', *Broadband Communications '94, IFIP TC-6, Telecom Paris*, Paris, France. 2-4 March 1994.
7. I. Nikolaidis and R. O. Onvural, 'A Bibliography on Performance Issues in ATM Networks' *Computer Communication Review*, Vol. 22, No. 5, pp 8-23, Oct. 1992.
8. M. de Prycker, *Asynchronous Transfer Mode : Solution for Broadband ISDN*, Ellis Horwood, 1991.
9. R. M. Rodriguez-Dagnino, M. R. K. Khansari and A. Leon-Garcia, 'Prediction of Bit Rate Sequences of Encoded Video Signals', *IEEE J. Selected Areas Commun.*, Vol. 9, No. 3, pp 305-313, April 1991.
10. Scientific and Engineering Software, *SES/workbench Reference Manual*, Release 2.1, February 1992.