Modeling Interactive Behaviour of a Video Based Multimedia System

Philip Branch, Greg Egan, Bruce Tonkin

Advanced Network Systems and Applications Group (ANSPAG)
Monash University, Clayton 3168, Melbourne
Australia

Abstract – **In this paper we report on statistics of user behaviour obtained during a semester of student usage of our video-on-demand based multimedia system. The statistics show that interactive behaviour is adequately modelled by exponential distributions, but that a better match is obtained from lognormal distributions. From the statistics we develop a model of interactive behaviour that consists of a Markov chain of video interactions overlaid by a Poisson process generating other interactions. We show that the model performs quite well, and provided caution is exercised, Markov processes are adequate models of interactive behaviour.**

I. INTRODUCTION

In this paper we attempt to characterise the interactive behaviour of users of a video based, multimedia system developed for education. We show that, while exponential distributions are adequate in describing interactive behaviour, lognormal distributions are better. We also construct a Markov model of interactive behaviour and compare its performance at predicting outcomes with observations.

Models of interactive behaviour are very important in designing networked multimedia systems [1, 2]. As users interact with the system, they introduce a new level of variability in its bandwidth demands. For video based systems, when in pause, little or no bandwidth will be required, but when in fast-forward or fast-rewind, substantially more bandwidth than the playback rate may be needed. Usually, behaviour is modelled with Markov processes. However, very little empirical data appears to have been gathered to support the validity of Markov modeling or even to find reasonable parameters for such models.

The McIVER system (Multi-campus Interactive Educational Resource) has been developed by our group as an experimental system to investigate multimedia usage and networking [3]. Film Studies, Art History, Languages and Communications students have used the system at Monash University for three semesters.

In our earlier video-on-demand trials we found that in educational applications, video was used in a very interactive manner. We found that, while VCR-like controls such as pause, fast-forward and rewind are heavily used, other functions are needed for effective use. We incorporated other media in the system, so that students could annotate the video, take notes, book-mark significant scenes, read lecturer's comments and extract stills for inclusion in essays. Consequently, our video-on-demand system evolved into a video-based multimedia system.

II. STATISTICS OF INTERACTIVE STREAMING MODES.

A. *Statistical Collection and Analysis.*

As part of the trials, user behaviour was recorded. Every user interaction generated an entry in a log file. From analysis of the log files, we can describe a hierarchy of interactive behaviour. At the highest level is a 'session'. A session is the interval between a student logging onto the system until they log off. During a session, students usually examine several videos. The examination of a single video is a 'viewing'. The duration of a session consists of the time spent viewing each video plus the time between viewings. During a viewing, the student can be in a number of modes including paused, normal playback or fast forward or rewind. Collectively these states are 'streaming modes'. Also, while in any of these states, a student might carry out non-streaming interactions that occur in negligible time and which occur in parallel with the other viewing modes.

It should be pointed out that the trials of this technology were not in any way a contrived exercise. The students who used the system were not technology-oriented students. Their primary interest in the system was to access video material for their studies.

B. *Descriptive Statistics.*

TABLE 1.
DESCRIPTIVE STATISTICS FOR FULL DATA SET

	N	Minimum	Maximum	Mean	Std. Deviation
Session	63	26	3664	1095	1000
Views per session	131	1	9	2.1	1.96
Viewing	131	19	3200	362	530
Pause	369	1	2110	62	194
Play	280	0	2141	93	215
Fast	76	0	196	18	30
Total Pause	131	2	2110	172	194
Total Play	131	0	3200	196	214
Total Fast	131	0	196	10	30

Table 1 summarises the statistics for the streaming interactions. N is the sample size of the statistic described. The other statistics are the minimum, maximum, mean and standard deviation for the function. These are all in units of seconds.

C. *Histograms of Interactions*

This section contains plots of histograms of interactive behaviour. The histograms show that all interactions exhibit positively skewed behaviour. For the streaming interactions this consists of many short interactions and far fewer long interactions, while for the non-streaming modes it consists of many short intervals and fewer long intervals between interactions.

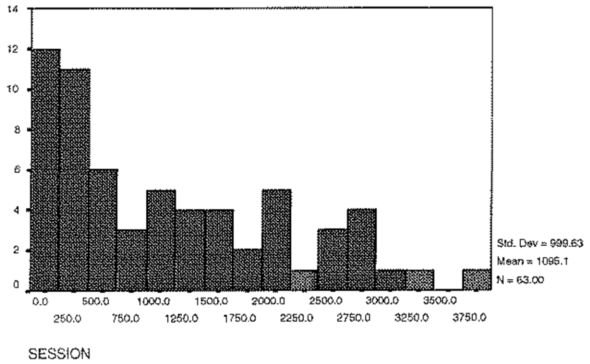

Figure 1. Session Duration

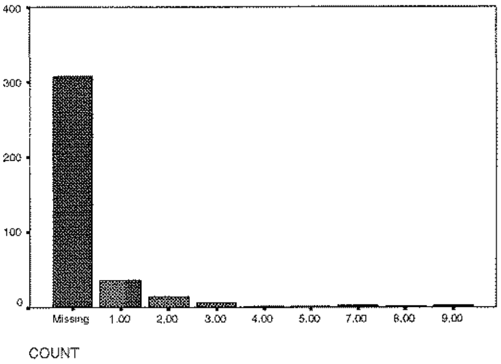

Figure 2. Number of Viewings per Session

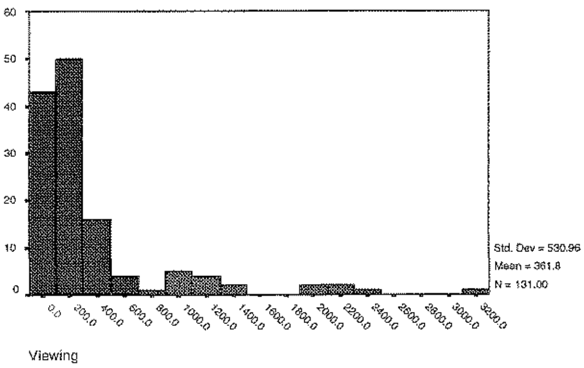

Figure 3. Viewing Duration

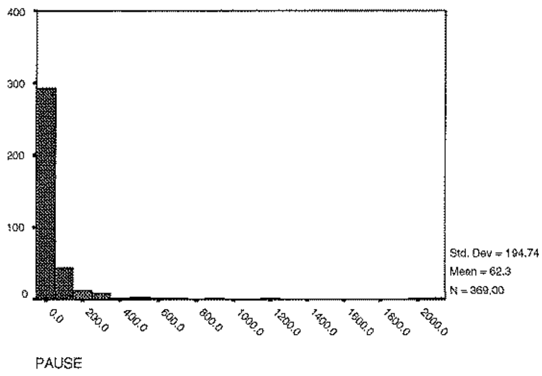

Figure 4. Pause Duration

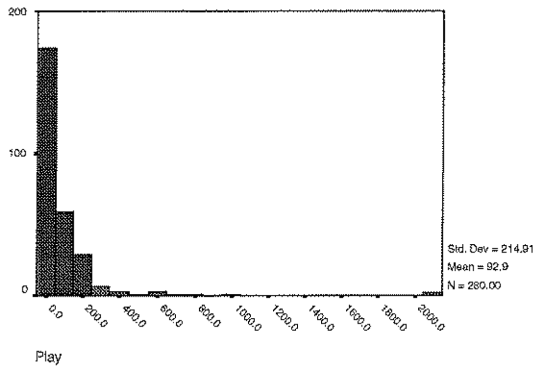

Figure 5. Play Duration

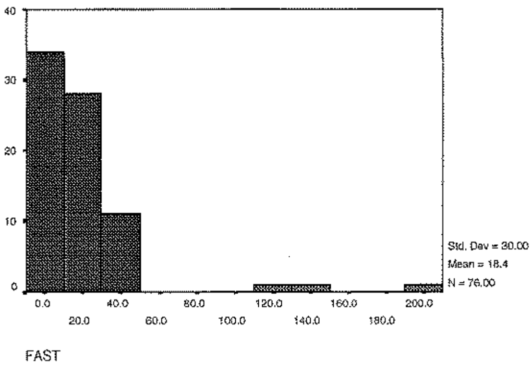

Figure 6. Fast Forward/Rewind Duration

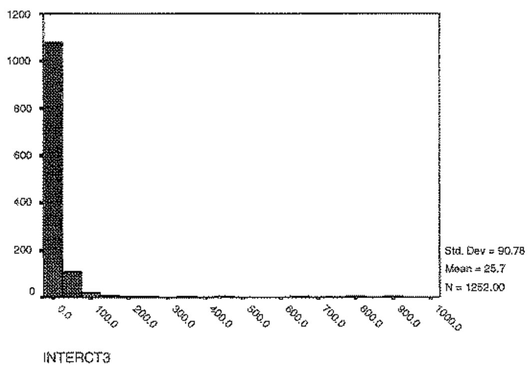

Figure 7. Non-streaming Interaction Interarrival Time

D. *Kolmogorov-Smirnov Test for Exponential and Lognormal distributions*

Of particular importance in understanding interactive behaviour is the distribution of time spent in each viewing mode. The Kolmogorov-Smirnov test is used to test the hypothesis that a set of sample points come from a particular distribution. Because of its success in modelling telephone call holding times, the exponential distribution is most commonly used in modelling interactivity. However, using the Kolmogorov-Smirnov test, we have found that the lognormal is a far better match than any other distribution. During the statistical analysis, other distributions, including the Pareto, were tested against the data. For every variable except the non-streaming interactions, the lognormal was an acceptable match to the data. For every variable, except the session duration, it was the best match to the data.

Table 2 summarises the Kolmogorov-Smirnov tests applied to our data [4]. To accept the hypothesis that the sample data comes from a continuous distribution whose parameters are estimated from the data, the maximum difference between the cumulative distribution function of the the samples and the continuous distribution must be less than a specific value for the hypothesis to be acceptable. Table 2 shows the data for each interaction.

TABLE 2.
K-S TESTS FOR STATISTICS OF INTERACTIONS.

Statistic	N	Exponential	Lognormal	Max for 5% Sig.
Sessions	63	0.099	0.129	0.154
Viewings	131	0.157	0.055	0.107
Pause	228	0.498	0.058	0.081
Play	178	0.218	0.041	0.091
Fast	53	0.207	0.122	0.168
Non-stream	1252	0.348	0.112	0.035

V. A MARKOV CHAIN MODEL OF INTERACTIVE STREAMING MODES.

A. *A Markov chain Model of Video States.*

Although our statistics show that the lognormal distribution is a better model of interactivity than the exponential, this is no reason to discard Markov methods entirely. Markov models are easy to understand and manipulate and if they can provide an approximate model of performance are well worth developing.

Consequently, we now model interactive behaviour with a finite Markov chain [5]. A finite Markov chain is a discrete time Markov process consisting of a finite number of states between which the process may switch. From the statistics, we calculate the probability of a transition to every other state each unit of time. In this model, we are only interested in the transitions during a viewing, and not with the commencement rate of viewings.

We segment time into seconds. At the end of each second, a client goes from state i to state j with probability of p_{ij} . The possible states of a finite Markov chain can be classified into sets of transient and ergodic states. Once a set of transient states is departed, it is never re-entered, and once a set of ergodic states is entered it is never departed. If there is only one state in the ergodic set, then it is an absorbing state, and the transition probability p_{ii} is 1. A Finite Markov chain, all of whose non transient states are absorbing, is called an absorbing chain. In any finite state Markov chain, the probability that after n steps it will be in an ergodic state, tends to 1 as n tends to infinity. Where the chain is an absorbing chain, the time until the process enters an ergodic state is the time until absorption.

We model the behaviour of client viewing by an absorbing chain. The time until absorption is the time until the user enters exit. This is the duration of the viewing. Our chain has three transient states (pause, play, and fast) and one absorbing state (exit). The viewing time is the sum of the total time spent in pause, play and fast.

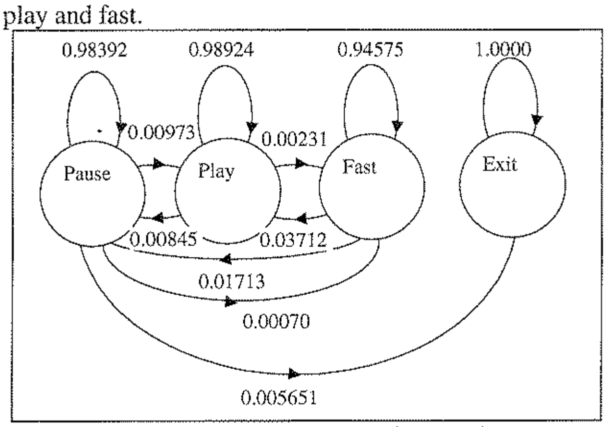

Figure 8. Markov chain of Interactive Behaviour

Figure 8 is a Markov chain of the behaviour. In the McIVER system, exit is only available from the pause state. From our measurements of student usage, we construct the following transition matrix.

$$P = \begin{array}{c} \\ \text{pause} \\ \text{play} \\ \text{fast} \\ \text{exit} \end{array} \begin{array}{c} \begin{array}{cccc} \text{pause} & \text{play} & \text{fast} & \text{exit} \end{array} \\ \begin{bmatrix} 0.98392 & 0.00974 & 0.00070 & 0.00565 \\ 0.00845 & 0.98924 & 0.00231 & 0.00000 \\ 0.01713 & 0.03712 & 0.94575 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix} \end{array}$$

The first statistics that can be derived from P are the mean times spent in each state once the process enters that state (the sojourn times). For state i:

$$t_i = \frac{1}{1 - p_{ii}} \qquad (1)$$

Similarly, we can find the variance of each sojourn time in state i:

$$v_i = \frac{p_{ii}}{(1 - p_{ii})^2} \qquad (2)$$

In deriving further values for an absorbing chain, a key matrix is the Fundamental matrix N defined by:

$$N = I + Q + Q^2 + Q^3 + Q^4 + ... \\ = (I - Q)^{-1}. \qquad (3)$$

If $\mathbf{n_j}$ is the mean number of times the process is in state j, and $\mathbf{M_i[n_j]}$ is the mean number of times the process is in state j if it started from state i, then

$$N = \{\mathbf{M_i[n_j]}\} \qquad (4)$$

If the process starts in state i then the total time spent in state j until absorption is N_{ij}. The variance of the time spent in state j is N_2 where

$$N_2 = N(2\,N_{dg} - I) - N_{sq} \qquad (5)$$

where N_{dg} is the matrix N with off diagonal elements equal to 0 and N_{sq} is N with all elements squared.

If $\mathbf{e}$ is the unit vector then

$$\mathbf{t} = N\mathbf{e} \qquad (6)$$

is a vector whose elements consist of the mean time to absorption from each initial state. If $\mathbf{t_{sq}}$ is the vector whose elements are those of $\mathbf{t}$ squared, then the vector of variance of the time to absorption is

$$t_2 = (2N\text{-}I)\mathbf{t} - \mathbf{t_{sq}}\,. \qquad (7)$$

In the next section we apply these results to our matrix of interactive behaviour

B. Values Derived from the Markov chain Model.

For pause, play and fast, the mean sojourn time calculated using equation (1) are 62, 92 and 18 seconds respectively. Using equation (2), the variance of the sojourn time in each state. For pause, play and fast the predicted variances are 3805, 8544 and 321 corresponding to standard deviations of 62, 92 and 18.

For additional values, we need the fundamental matrix N. We obtain this using equation (5).

$$N = \begin{bmatrix} 177 & 197 & 11 \\ 177 & 305 & 15 \\ 177 & 271 & 32 \end{bmatrix}$$

Using equation (6) the vector of total times in pause, play and fast is [177, 197, 11]. Using equation (7) the variance of the total time spent in each state during a viewing is

$$N_2 = \begin{bmatrix} 31152 & 81164 & 572 \\ 31152 & 92720 & 720 \\ 31152 & 91598 & 992 \end{bmatrix}.$$

Since users always start in pause, the vector of variances of total time spent in pause, play and fast is [31152, 81164, 572] and the standard deviations are [176, 284, 23]

Using equation (8) and (9) the vectors of mean and variance of viewing duration are respectively [385, 497, 480] and [195437, 208835, 205984]. Since clients always start their viewing in pause state, the mean time viewing duration is 385 seconds with variance of 195437 which gives a standard deviation of 442 seconds.

C. Comparison of Derived and Observed Values.

In this section we compare the results derived from the Markov chain with observed values.

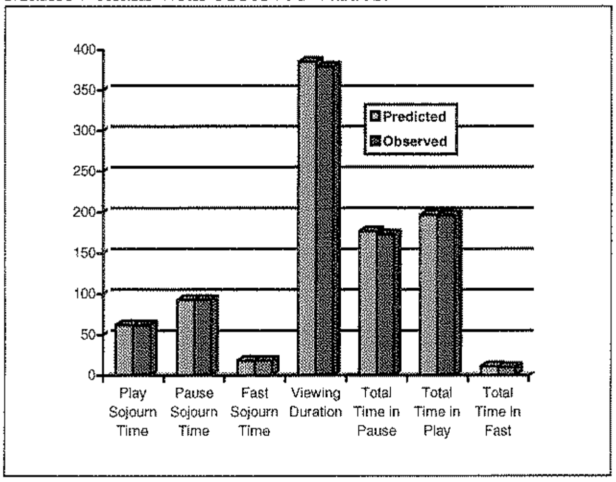

Figure 8. Comparison of Mean of Predicted and Observed Statistics.

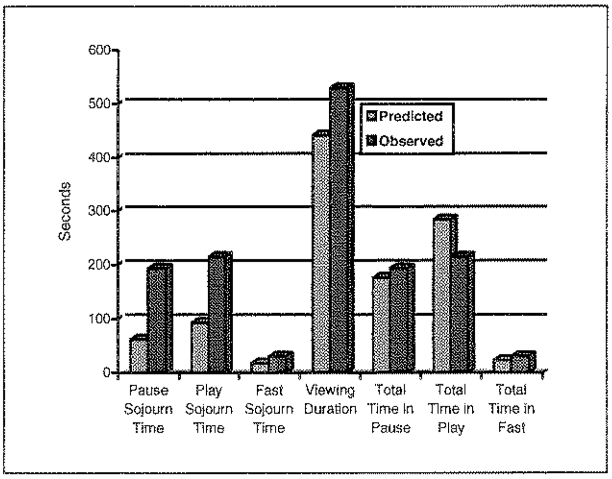

Figure 9. Standard Deviation of Predicted and Observed Statistics.

Figure 8 shows predicted and observed means. There is a very impressive agreement between predicted and observed means. However, some caution needs to be exercised in assuming that this shows the strength of the Markov model since the mean values do not depend on the data having an exponential distribution.

However, the difference between the predicted and observed standard deviations do depend on whether the data follows a negative exponential distribution.

Figure 9 shows the predicted and observed standard deviations for the statistics. Amongst the sojourn times, agreement is weakest for pause and strongest for fast. The predicted and observed standard deviations for the viewing duration and the total times in each state agree quite well.

From the comparison of the predictions and observations, we see that the Markov chain model performs quite well in predicting means and variances. We conclude that Markov techniques are suitable for modelling interactive streaming mode behaviour provided the outcomes are used cautiously.

D. *A Poisson Model of Non-Streaming Interactions.*

We now model the events that don't affect sustained video bandwidth. In our system, these include bookmarking of scenes of interest, using the slider bar to jump from one scene to another, taking notes and extracting stills for inclusion in student notes. These events occur in a negligible time, and don't alter the video state (pause, play or fast), but may cause a large bandwidth 'spike' to occur as buffers are refreshed.

We use the Poisson distribution to model the arrival of these events. The Poisson distribution specifies a discrete random variable, the number of times we can expect an event to occur during a specific interval.

If N is the random variable 'number of occurrences', and λ the mean number of occurrences during the interval, then the probability of n occurrences during the interval has probability mass function of:

$$\Pr(N = n) = \frac{e^{-\lambda}\lambda^{n}}{n!} \qquad n = 0,1,2,...$$

From the statistics, if the interval is ten seconds then λ is 0.68 events per ten seconds.

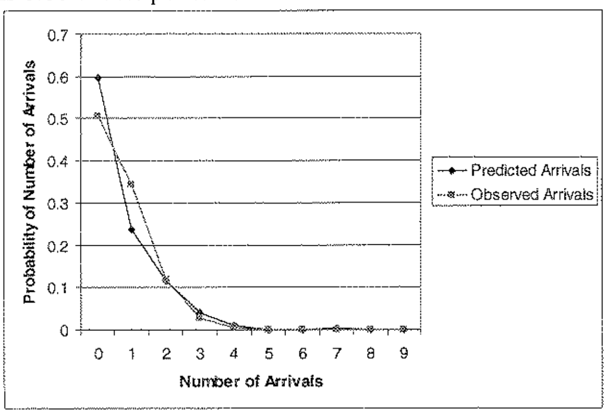

Figure 10. Expected and Observed Probability of Number of Arrivals in 10 Second Intervals

Figure 10 shows the probability of zero to 9 events occurring during any ten second interval. The two lines show the predicted and observed probability. Once again we see that Poisson arrivals are adequate at predicting the arrival rate of non-video interactions.

V. CONCLUSION.

This paper fills a significant gap in current multimedia knowledge. It identifies the importance of interactivity in video based multimedia and shows how interactive behaviour can be modelled. We show that exponential distributions are adequate for modelling of interactive behaviour, but where high levels of accuracy are required, lognormal distributions may be preferable.

We use a Markov chain to model interactive video behaviour, overlaid by a Poisson arrival process to model non-video interactions. Even though the time spent in each state is not negatively exponentially distributed, this Markov based model is quite effective.

The statistics obtained for this paper were from a homogeneous group using the system over a short period of time. Further research will include collating statistics over a longer period of time with heterogeneous groups, and in seeing whether the observations made for this system can be generalised to other multimedia systems.

ACKNOWLEDGEMENTS

We acknowledge the support of Monash University, Telstra Australia, Siemens Pty Ltd and the Australian Government's Cooperative Research Centre scheme.

REFERENCES

[1] V. O. K. Li, W. Liao, X. Qiu, and E. Wong, "Performance Model of Interactive Video-on-Demand Systems," IEEE Journal on Selected Areas in Communications, vol. 14, no. 6, 1996, pp. 1099-1109.

[2] J. Dey-Sircar, J. Salehi, J. Kurose, and D. Towsley, "Providing VCR Capabilities in Large-Scale Video Servers," presented at ACM International Conference on Multimedia, San Francisco, 1994, pp. 25-32.

[3] P. Branch and B. Tonkin, "Multicampus Video-on-Demand Trials at Monash University," Australian Journal of Educational Technology (AJET), vol. 13, no. 2, 1997, pp. 85-97.

[4] P. Sprent, Applied Nonparametric Statistical Methods, 2 ed. London: Chapman and Hall, 1989.

[5] J. Kemeney, J. Snell, and A. Knapp, Denumerable Markov Chains. Princeton, New Jersey: Van Nostrand, 1976.