An Empirical Study of Client Interactions with a Continuous-Media Courseware Server [1]

Jitendra Padhye, Jim Kurose
Department of Computer Science
University of Massachusetts
Amherst MA 01003 USA
{jitu,kurose}@cs.umass.edu

Abstract

While considerable research has gone into investigating various networking and operating system mechanisms for supporting the transfer and playout of stored continuous media (e.g., audio and video), very little information is available about how users actually *use* such systems. Understanding how users interact with such a system – i.e., developing a user workload characterization – is crucial in designing and evaluating efficient continuous media (CM) resource allocation and access mechanisms. We have designed and built an interactive WWW-based, multimedia, client/server application, known as MANIC (Multimedia Asynchronous Networked Individualized Courseware), that streams synchronized CM (currently audio) and HTML documents to remote users. MANIC was used by more than 200 users during the Spring 1997 semester to listen to, and view, the stored audio lectures and lecture notes for a full-semester senior-level course at the University of Massachusetts. In this paper we provide empirical and analytic characterizations of observed user behavior in MANIC. We characterize both session-level behavior (e.g., the length of individual sessions) as well as *interactive* user behavior (e.g., the time between starting/stopping/pausing the audio within a session; and the "distance" of rewinding, fast forwarding and indexed jumping within the audio stream). We also examine how well our measured data can be fit analytically by various distribution functions. Finally, we consider the possible implications of our results for CM system architecture.

1 Introduction

Efficient delivery of continuous media (CM) such as audio and video, is one the important goals of current networking and operating systems research [6, 9, 16]. A large amount of

[1]This material was supported in part by the National Science Foundation under Grant No. CDA-9502639, NCR-9508274, and the University of Massachusetts through an Instructional Technology Grant. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the University of Massachusetts.

research has gone into investigating various networking and operating system mechanisms to make continuous media transfer more efficient. However, little information is available about how users actually *use* such systems. Understanding how users interact with such a system – i.e., developing a user workload characterization – is crucial in designing and evaluating efficient CM resource allocation and access mechanisms. For example, understanding how often users fast-forward, rewind, pause or make indexed-jumps within a continuous media stream (and how "far" they move within the media stream when they do so), can help in designing and evaluating call admission strategies [3], scalable compression techniques [17], and client/server buffer management strategies [4] for CM systems.

In this paper we provide an empirical characterization of observed user behavior with MANIC (Multimedia Asynchronous Networked Individualized Courseware) [7], a WWW-based multimedia client/server application, that streams synchronized CM (currently audio) and HTML documents to remote users. MANIC was used by more than 200 users during the Spring 1997 semester to listen to, and view, stored synchronized audio lectures and lecture notes for a full-semester senior/graduate-level course at the University of Massachusetts. Our particular focus here is on both session-level characteristics (e.g., the length of the individual sessions) as well as *interactive* user behavior (e.g., the time between starting/stopping/pausing the audio within a session; and the "distance" of rewinding, fast forwarding and indexed jumping within the audio stream).

Our measurements indicate a high degree of user interactivity. We find that once audio playout begins, in more than 45% of the cases it lasts less than 3 minutes before the user pauses or jumps elsewhere in the audio stream; the average length of time spent listening to an audio stream before a pause or jump, though, is slightly over 8 minutes. From a modeling perspective, we find that session lengths, and the length of time a user spends in continuous audio playout (before a stop, pause, or jump) or with playout halted (before resuming audio playout) has a "heavier" tail than the exponential distribution that is typically assumed in previous CM analytic modeling [3, 17]; both the lognormal and gamma distribution are shown to provide a closer fit to our observed data than the exponential distribution.

Surprisingly, we also find that forward jumps in the audio stream are more than seven times more likely than backwards jumps (we had expected to see a relatively larger number of short backward jumps, as a student reviewed a not-understood portion of the lecture). We also find that while there are a significant number jumps over relatively short periods of time (e.g., approximately a third of the jumps were less than 3 minutes forward in the audio stream), the average length of a jump was much larger (over two thousand seconds). Finally, at the session level, we find that the average session length is slightly less than half an hour, and that the vast majority of the sessions (more than 80%) are less than an hour in length.

We believe that while these measurements are themselves of independent interest, they are also of broader value. From a system architecture standpoint, the level of observed user interactivity suggests that mechanisms designed specifically to support the uninterrupted, long-term playout of CM (e.g., VoD servers for entertainment movies) may not be well-suited for the more interactive educational/training environments. From a modeling perspective, our results question the accuracy of some previously adopted modeling assumptions [3, 17] and provide specific "real-world" parameters values that can be employed in future studies. We stress, however, that our results are for a specific application, a specific set of multimedia course material, and a specific user population. Other populations and other servers with different CM applications and/or content may well exhibit significantly different user behavior. Nonetheless, the present work does provide an important first data point characterizing the interactive behavior of users in an operational CM server.

The remainder of this paper is organized as follows. Section 2 provides an overview of the MANIC architecture, describing the CM content and the user population. In Section 3, we describe how user interactions are captured and present empirical session-level characteristics. In section 4, we examine the finer-grained behavior of users within a session. Section 5 discusses implications of our measurements and Section 6 concludes the paper.

2 Overview of the MANIC Architecture and Content

In this section, we overview the MANIC architecture and discuss the structure of the multimedia course content that was accessed by students. For a detailed description of MANIC, the reader is referred to [7]. The CM course material that was accessed during this study can itself can be accessed via `http://gaia.cs.umass.edu/cs653-1997`.

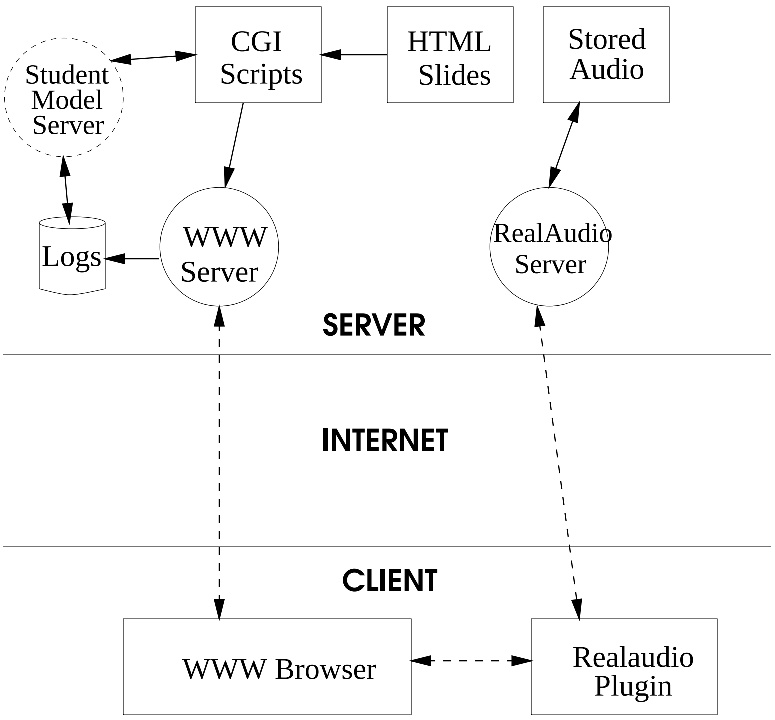


Figure 1: MANIC architecture

The architecture of the MANIC system is depicted in Figure 1. In its primary mode of operation, the system allows the user (client) to listen to pre-recored lectures with accompanying text. Pages of text ("slides") are displayed as HTML files in the browser window. Audio is delivered by the server using the RealAudio Server, and played out at the client side user the RealAudio browser plugin [13]. Slide display and audio playout are synchronized i.e., the displayed HTML contents in the WWW browser window change automatically as the audio progresses. This is achieved using the synchronized multimedia facility provided by the RealAudio server/player. Details about this synchronization mechanism, and how it is exploited in MANIC, can be found in [7]. In addition, MANIC can highlight (and then de-highlight) different portions of the text as the audio progresses, as discussed below.

Several CGI scripts running on the WWW server allow the user to perform various control/browsing actions by clicking on the control buttons provided within each slide. These commands include stopping and restarting the audio, moving to the previous or the next slide and jumping to arbitrary slides via an index. A detailed description of the implementation is provided in [7]. Here, we only describe what these commands do, and how they affect audio playout.

A screen shot of the MANIC user interface displaying a typical slide is shown in Figure 2. The slide contains two bullet points and a diagram. As noted above, the contents of a slide are partitioned into predefined groups that can be individually highlighted as the corresponding audio is played out. Highlighting is achieved rather crudely in the current version of MANIC by loading a new HTML page whenever a different highlight is to be displayed. All HTML pages associated with a given slide have the same text and graphics content, but different font color tags depending upon which portion is to be highlighted. In Figure 2, there are two such "highlights", each corresponding to a bullet point[2]. The audio corresponding to the first highlight was playing when the screen shot was taken. A highlight is the smallest unit of user interaction with MANIC system. That is, the smallest granularity with which a user can index or jump through the audio stream is by one highlight at a time.

Let us now consider the user commands in MANIC. As shown in Figure 2, the following commands are available:

- **Continuous Play:** This command starts the audio playout at the current highlight. The audio will continue to play uninterrupted until the end of the current lecture, unless stopped (as explained below) by the user. Slides and highlights change, as discussed above, as the audio progresses. Once the end of the lecture is reached, the audio stops playing [3].

- **Play Current Slide:** This command starts audio playout at the beginning of the current highlight. Audio playout stops once the end of the current slide is reached.

[2]This fact is not evident from the black and white screen shot. A highlight can contain any number of bullet points and/or arbitrary HTML code.

[3]Slides are grouped together into "lectures", corresponding to the actual lectures given live to in-class students. Each lecture is about 70 minutes in length. Each lecture was stored in a separate RealAudio file for ease of management. Since audio playout stops at the end of a lecture, the user must click on the "Next" button and then use one of the two "play" commands to restart the audio. Note that this limits the maximum length of continuous audio playout to the length of lecture.

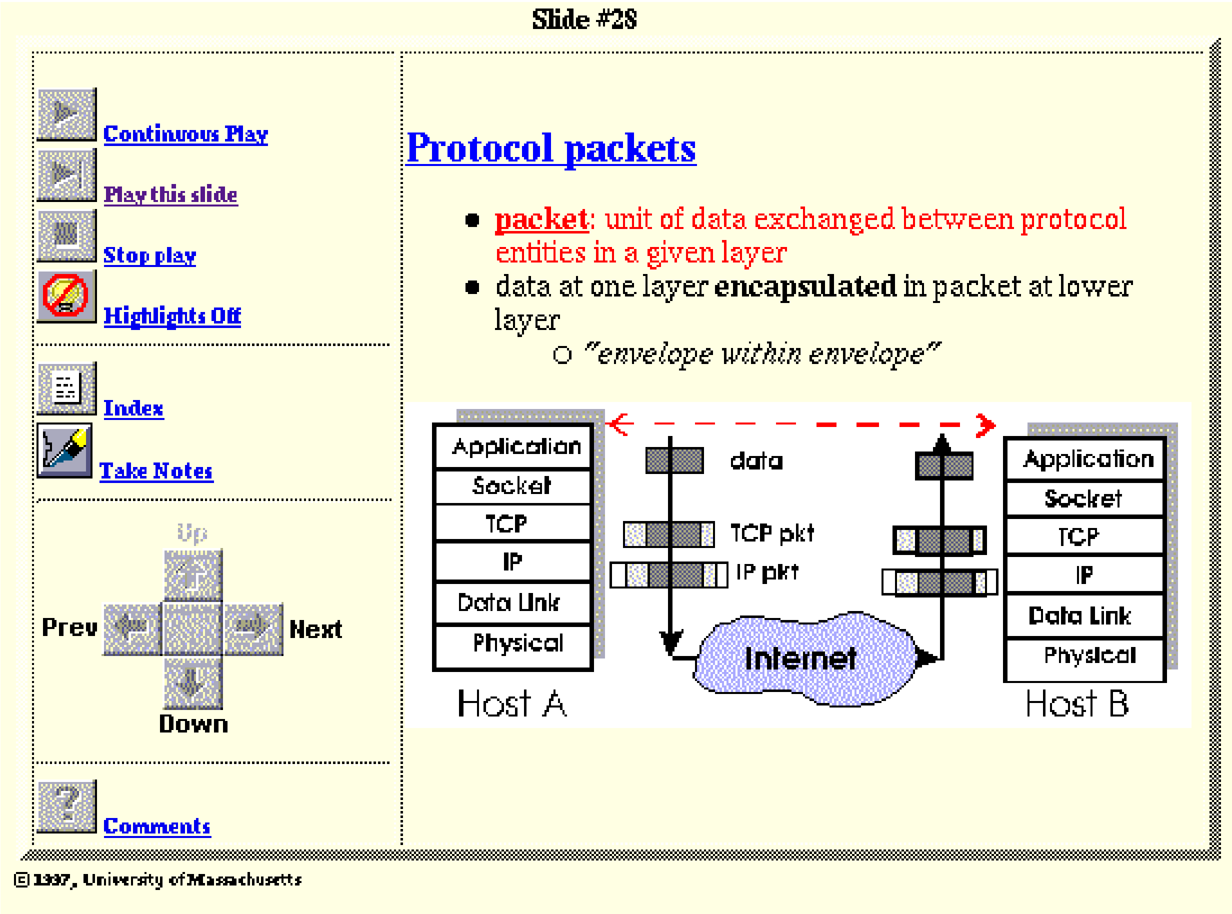


Figure 2: MANIC user interface

- **Next, Previous:** These commands stop the audio if it was playing, and display the next or previous slide in the lecture in the browser window, respectively.
- **Up, Down:** These commands stop the audio if it was playing, and move the highlight to next or previous highlight within the current slide, respectively. The Up and Down commands provide the means for "fast forwarding" or "rewinding" through material at the granularity of a highlight - the smallest possible unit of forwards or backwards jumping available within MANIC. [4]
- **Index:** This command brings up a window that displays an index of course material by lecture number, table of contents or keyword search. The audio continues to play while the index is displayed. However, if the user clicks on an index entry, the audio stops playing and the slide corresponding to the table entry is displayed in the browser window.
- **Stop Playing:** This command stops the audio playout. The slide displayed in the

[4]MANIC also provides a "non-highlight" mode [7] for users with slower network connections. In this mode, the user can FF/rewind only at the granularity of a single slide. Approximately 10% of the dowloaded slides were viewed in non-highlighted mode.

browser window remains unchanged.

In each case when the audio stops, the user must explicitly restart playout by using one of the two "play" commands. We note that while MANIC provides the capability to start/stop and index through audio (i.e., by using the commands described above), a student can also manipulate audio via a slider bar in the Real Audio pop-up window (although we explicitly encouraged students not to do so); we discuss the difficulties posed by this second way to control audio shortly.

The brief description above has discussed MANIC's functionality as a CM-on-demand server. The client can start and stop the audio playout. Fast-forward, rewind and indexed-jumps are affected by jumping to predefined, fixed points in the audio stream, using the Next, Previous, Up, Down, and Index controls.

2.1 System Parameters and User Population

The current version of MANIC is designed around a 25-connection RealAudio server and a NCSA WWW server. Both the audio and the WWW server run on a Silicon Graphics *Challenge* workstation with two CPUs, 256 MB of RAM and 64 Gigabytes of hard disk space. Users can access course material from a variety of PC and workstation platforms using Netscape with a RealAudio plugin. While the audio server is capable of handling 25 simultaneous connections, we have seldom seen more than 5-7 users active at the same time.

The system was used to deliver a 3-credit, senior/grad level course on Computer Networks taught at the University of Massachusetts at Amherst, during the Spring 97 semester. The course itself was offered live to approximately 120 registered on-campus students and (by satellite) to 40 off-campus students. The audio and video of the live presentation of the course (instructor lectures, questions and answers, discussion) were professionally recorded by the UMass Video Instructional Program and the audio encoded in RealAudio format for 14.4 Kbps access (in a few cases, lectures were encoded for 28.8 Kbps access). Each lecture was made available on-line within 1 or 2 days after it was delivered in the classroom. Overall, the course consisted of 25 lectures, each of which was approximately 70 minutes in length. There were approximately 400 HTML slides.

It is important to note that the course was offered live, at the same time it was being made available to students via MANIC. It is likely, therefore, that a student using MANIC to view part of a lecture may well have already attended that lecture live. We note that the behavior of student viewing material for the first time may well be different from a student reviewing material. We have anecdotal evidence, however, that MANIC was also used to make up for missed classes, and several students admitted (after the class was over) to having "attended" more classes via MANIC than in person. Since the course was made available on the Internet free of charge, there were also additional users who had not officially registered for the class.

Before accessing MANIC for the first time, each user was required to select a unique (potentially anonymous) username. The users were expected to use the same username each time they subsequently used the system. During the period of the Spring semester, we recorded 220 unique usernames. For each username, a separate activity log was maintained. While it is possible that some users may have used multiple aliases and changed their identity from one session to the next, we believe that the effect of any aliasing is minimal, particularly because we characterize aggregate behavior, rather than the behavior of a particular user.

3 Measurements and Session-level Characteristics

The activity of each MANIC user (for the particular course described above) was logged by the MANIC server throughout the semester. These per-user logs were then analyzed off line to produce the user characterizations described in this section and the next. The 215 user logs that were analyzed to produce this report contained more than 33,000 such log entries[5]. In addition to the logs kept by the MANIC server, the Real Audio server also kept its own logs of audio only activity. As discussed in the Appendix, a users occasionally directly manipulated the audio using the Real Audio pop up window (rather than the MANIC audio controls). In such cases, we thus had to coordinate the Real Audio logs with our MANIC logs.

Let us begin by describing the format and the contents of the MANIC activity logs. Recall that each user is identified by a unique username. This username is transmitted in all HTTP requests and maintained at the server through the use of *cookies*. This allows the WWW server to maintain logs of actions performed by individual users. Each log entry specifies either a user command (such as stop, play, previous/next/up/down) or the slide and highlight sent by the server to the user's browser window. The log entries are timestamped with a granularity of 1 second.

The log fragment shown in Figure 3 is a typical example of the logs generated by the MANIC system. Lines 1 and 2 indicate that the user started on the first highlight of slide 1. The audio played continuously (line 3) until partly into the second highlight of slide 2. After that, the user used the index (line 7) to jump to slide 10 (line 8). Note that the audio stops playing at this point. The user then moved forward two slides by using the "next" button (lines 9-12). The audio was restarted from the first highlight of slide 12 (lines 12-13), and it was stopped shortly thereafter (line 15). The last line (line 16) simply indicates that when the user hit the "stop" button, slide 12 was reloaded. This is an artifact of the way we update highlights and maintain log entries.

3.1 Session Length

Perhaps the most important session-level user characterization is the length of each session itself, i.e., the length of time the student is sitting at the computer using MANIC before he/she leaves. This value is easy to measure if there are explicit log-in and log-off procedures. Although we instrumented MANIC with a log-on procedure for this study, this

[5]This data is available on-line. See ftp://gaia.cs.umass/edu/pub/jitu/MANIClogs/README

```
 1 Sat Apr 26 16:00:42 1997- -start-1
 2 Sat Apr 26 16:00:42 1997-slide=1&buttons=N&highlight=1-1
 3 Sat Apr 26 16:00:47 1997- -continuous_play-1
 4 Sat Apr 26 16:00:48 1997-slide=1&buttons=N&highlight=1-1
 5 Sat Apr 26 16:02:14 1997-slide=2&buttons=DPN&highlight=1-1
 6 Sat Apr 26 16:04:34 1997-slide=2&buttons=UDPN&highlight=2-1
 7 Sat Apr 26 16:04:42 1997- -index-1
 8 Sat Apr 26 16:04:43 1997-slide=10&buttons=PN&highlight=1-1
 9 Sat Apr 26 16:04:47 1997- -next-1
10 Sat Apr 26 16:04:47 1997-slide=11&buttons=DPN&highlight=1-1
11 Sat Apr 26 16:04:49 1997- -next-1
12 Sat Apr 26 16:04:49 1997-slide=12&buttons=DPN&highlight=1-1
13 Sat Apr 26 16:04:52 1997- -continuous_play-1
14 Sat Apr 26 16:04:59 1997-slide=12&buttons=DPN&highlight=1-1
15 Sat Apr 26 16:05:12 1997- -stop-1
16 Sat Apr 26 16:05:13 1997-slide=12&buttons=DPN&highlight=1-1
```

Figure 3: Example of MANIC user logs

explicit login procedure could be bypassed by the savvy user. More importantly, we chose *not* to have an explicit log-off procedure in MANIC, primarily because we could not ensure that students would indeed explicitly log-off when done. (They could simply walk away from, or turn off, their PC or workstation when their session was "over.")

Without a specific logout procedure, it is difficult to measure a session's length, or indeed even define what is meant by a "session." We thus adopted the following approach to *estimate* session length. In order to define a session, we first define a Session Gap Threshold (SGT). The SGT is an amount of time, such that if a particular user is inactive for more than this amount of time (e.g., no audio in played out, and no MANIC commands are entered), the user's session is considered to have terminated. Clearly, the larger the value of the SGT, the smaller the number of the perceived sessions in our logs. In the limit, an SGT value of infinity will result in a number of sessions equal to the number of distinct users of the system. On the other hand, an SGT value of zero will result in a number of sessions equal to the total number of log entries.

To determine a reasonable value for the SGT, we plot the number of "sessions" that would be present in the logs for various values of SGT, as shown in Figure 4. From Figure 4, it can be seen that a reasonable value of the SGT would be approximately 30 minutes, as the slope of the graph becomes very close to zero after that. This implies that a user that is inactive for more than 30 minutes is likely to remain inactive for a much longer period of time. Thus, for all further measurements, we assume that an SGT of 30 minutes is used to define our notion of "session."

Once we have selected a value for SGT, we can measure the lengths of individual sessions. The distribution of session lengths indicates the pattern of user "holding times" i.e., the amount of time each user spends in the system. In Figure 5 we plot the distribution

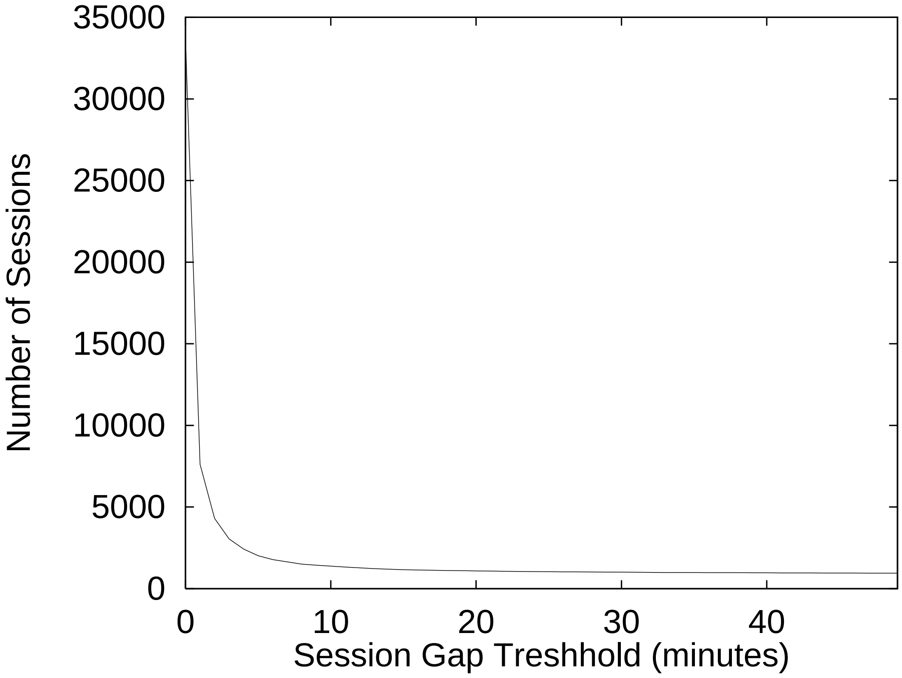


Figure 4: Number of sessions as a function of SGT

of session lengths with an SGT of 30 minutes. The leftmost graph in Figure 5 shows a histogram of the observed session lengths. We have normalized the heights of the histogram bars so that the area under the histogram is one. The average session length is 1731 seconds (approximately a half hour); the standard deviation is 2662.

The leftmost graph in Figure 5 also plots the three fitted distributions: an exponential distribution of the form

$$f(x) = \lambda e^{-\lambda x},$$

a gamma distribution of the form

$$f(x) = \frac{\lambda_g e^{-\lambda_g x} (\lambda_g x)^{\alpha - 1}}{\Gamma \alpha}$$

and a lognormal distribution of the form

$$f(x) = \frac{1}{x(2\pi\sigma^2)^{0.5}} e^{\frac{-(ln(x)-\mu)^2}{2\sigma^2}}.$$

The distribution parameters $\lambda, \lambda_g, \alpha, \sigma^2, \mu$ were obtained by the method of moments - matching the mean (in the case of the exponential distribution) or the mean and variance (in the case of the gamma and lognormal distributions) of the fitted distributions to their empirically observed values. By informal inspection, we note that gamma distribution most closely fits the measured data. We also note that the exponential distribution tends to underestimate the probability mass for both short and long sessions.

The right graph in Figure 5 plots $1 - F(x)$ on a logarithmic y-axis to highlight the tail behavior. Here we see that both the gamma and lognormal distributions tend to match

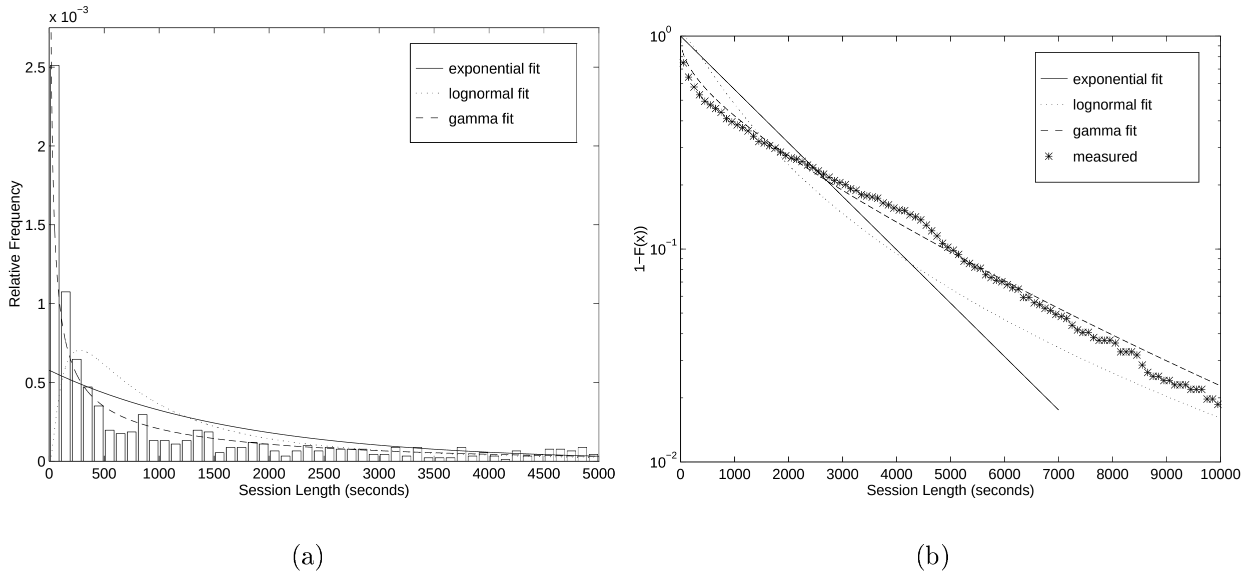


Figure 5: Session length characteristics: empirical measurements and fit data

the tail well, while the exponential distributions tends to underestimate the tail of the distribution.

Our measurements indicated that that slightly under half of sessions did not play any audio at all. Figure 6 presents the session lengths from this perspective. Here, we divide the sessions into two groups, depending on whether or not a session played any audio. Figure 6 shows that sessions with audio tend to last much longer session than those without audio. The mean session length for those with with audio is 2827.22 seconds, while the mean session length for those without audio is 304.52 seconds.

4 User Activity Within a Session

Having examined the length of a user session, we next characterize user behavior within a session.

Perhaps the coarsest characterization of user behavior with respect to the CM content is the fraction of time within a session that the user actually utilizes (listens to) the CM content. Recall that within MANIC, a user need not always be receiving streamed audio. For example, a user may choose to turn audio off while browsing through the HTML lecture slides (or other WWW documents). For example, the log fragment shown in Figure 3 indicates that the user had the audio on between 16:00:48 and 16:04:42 (lines 4-7), off (as the result of the indexed jump to a new slide in line 8), and then on again between 16:04:59 and 16:05:12 (lines 13-15).

We define a session's "audio utilization" to be the ratio of the amount of time the audio is actually being played during a session to the session length. The histogram in Figure 7 plots the fraction of sessions (y-axis) that were observed to have a given audio utilization

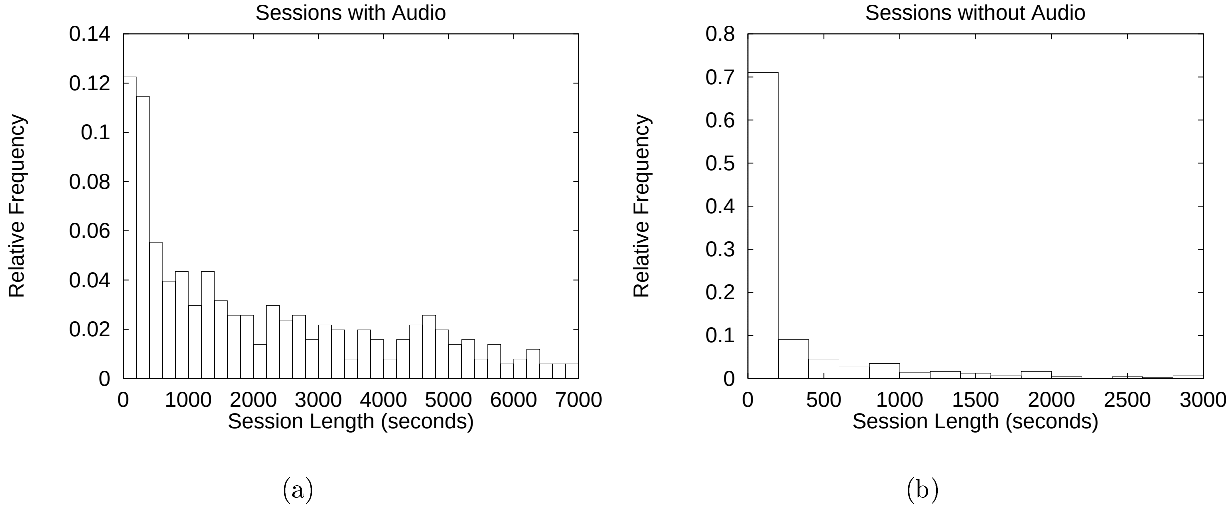


Figure 6: Sessions with and without audio playout

(x-axis). The average audio utilization is 67.59% and the standard deviation is 27.50. These numbers are calculated only for those sessions that include at least some audio playout. Recall that almost 50% of the sessions contain no audio playout at all. The average audio utilization of *all* sessions is 37.50% and the standard deviation is 39.34. We note that the relatively low average utilization indicates that a CM server that allocates resources on the basis of worst case demand (e.g., assumes that each session will always be playing audio) will severely over-reserve resources.

As discussed above, an individual user's CM activity within a session will consist of alternating intervals of having the audio on and off. Indeed, a number of analytic models of users in interactive CM systems explicitly model this on/off behavior [3, 17]. Figure 8 plots the distribution of the length of the individual audio ON periods. Once again, the heights of the histogram bars have been normalized so that the area under the histogram is one. Interestingly, more than 45% of the audio ON periods are less than 180 seconds in length. The average length of an audio ON period is 509.50 seconds and the standard deviation is 815.92. The mass at 4200 seconds corresponds to approximately 70 minutes of continuous audio playout, the approximate length of an individual lecture.

Figure 8 also plots fitted exponential, gamma, and lognormal distributions. We see again that the exponential distribution appears to provide the poorest fit, while both the lognormal and gamma distributions match the distribution of the observed data over a wide interval.

Figure 9 plots the distribution of the time spent between periods of audio playout – time when no CM is being streamed to the user. During this time users may be browsing through class slides (as discussed below), browsing other WWW documents, thinking, or sleeping. The average length of silence between successive audio subsessions is 171.91 seconds; the standard deviation is 392.24. The fitted distributions again indicate that the gamma distribution is a particularly good fit to the observed data (except at very large

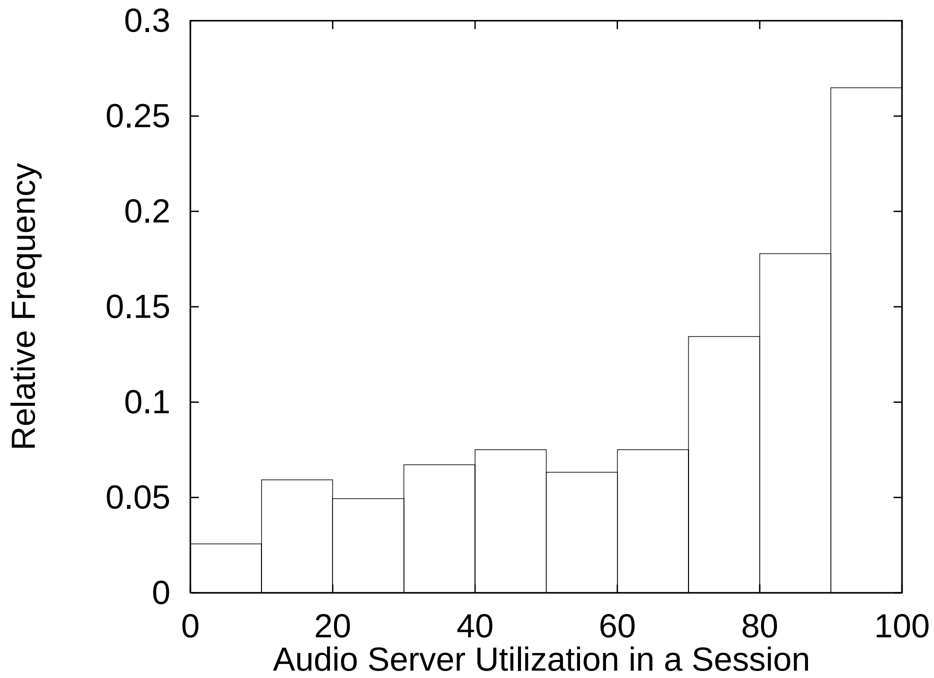


Figure 7: Audio utilization in a session

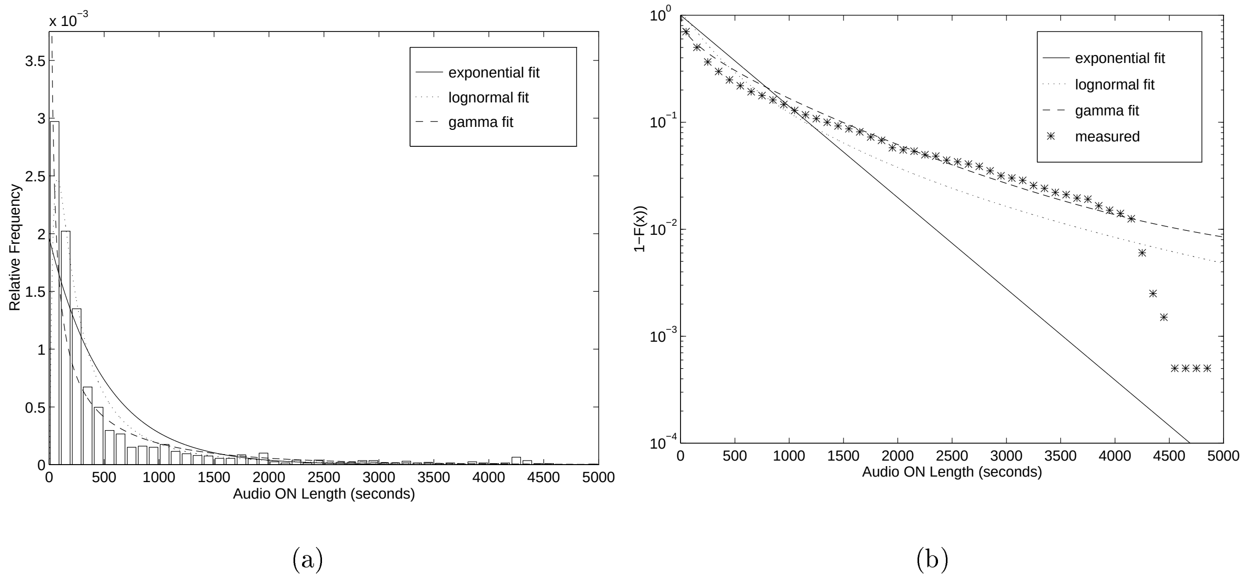
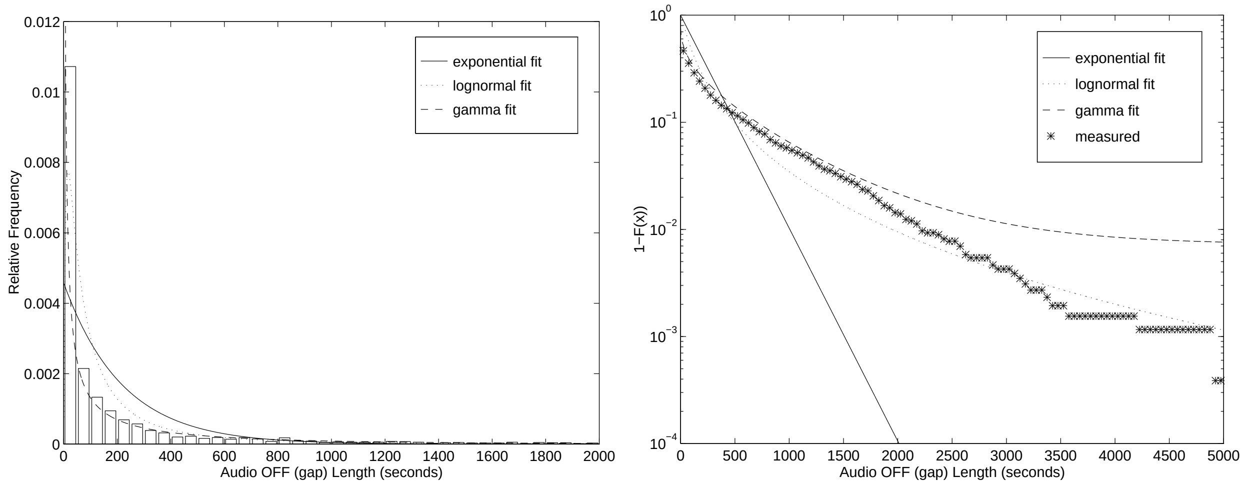


Figure 8: Audio ON lengths: empirical measurements and fit data



(a) (b)

Figure 9: Audio OFF gap lengths: empirical measurements and fit data

tail values), as is the lognormal distribution. Once again, we see that the exponential distribution is a poorer fit for both small and large values of the audio OFF period.

As discussed above, after a user ends an audio ON period, another audio ON period may eventually be initiated (following the audio OFF period). When re-initiating audio playout, a user may continue the audio where he/she had left off, or may choose to jump to a new point in the audio stream (perhaps, for example, via the use of the index). As an example, consider the log fragment shown in Figure 3. The audio-off period begins at clock time 16:04:42 (line 7), ends at clock time 16:04:52 (line 13) with a forward jump in the audio stream.

We now define the length of a "jump" in the audio stream more precisely. Consider the log fragment in Figure 3. In that example, the user has jumped from slide 2, highlight 2 (line 6) to slide 12, highlight 1 (line 12), between two audio subsessions. To measure the length of this jump, we construct a timeline as follows: we consider slide 1, highlight 1 to begin at time 0, and measure the time continuously, so that the last highlight of the last slide of the last lecture (highlight 1, slide 393, lecture 25) is considered to end at time 101685.9. On this timeline, slide 2, highlight 2 begins at 219.9 seconds[6] and slide 12, highlight 1 begins at 2795.9 seconds. Thus we consider the length of jump to be the difference between these two values, i.e. 2576 seconds. Note that we ignore the fact that the user has played out part of slide 2, highlight 2 before making the jump. Thus, we measure the length of jumps only at one highlight granularity. This scheme of measurement reflects the fact the MANIC user interface provides fast forward/rewind capabilities only at one highlight granularity – i.e. using MANIC controls, a user cannot jump into the middle of

[6]Note that the value 219.9 (not shown in line 6 in Figure 3) is not at all related to the wall time of 16:04:34. This latter time is simply the time of day the user happened to listen to this particular audio segment.

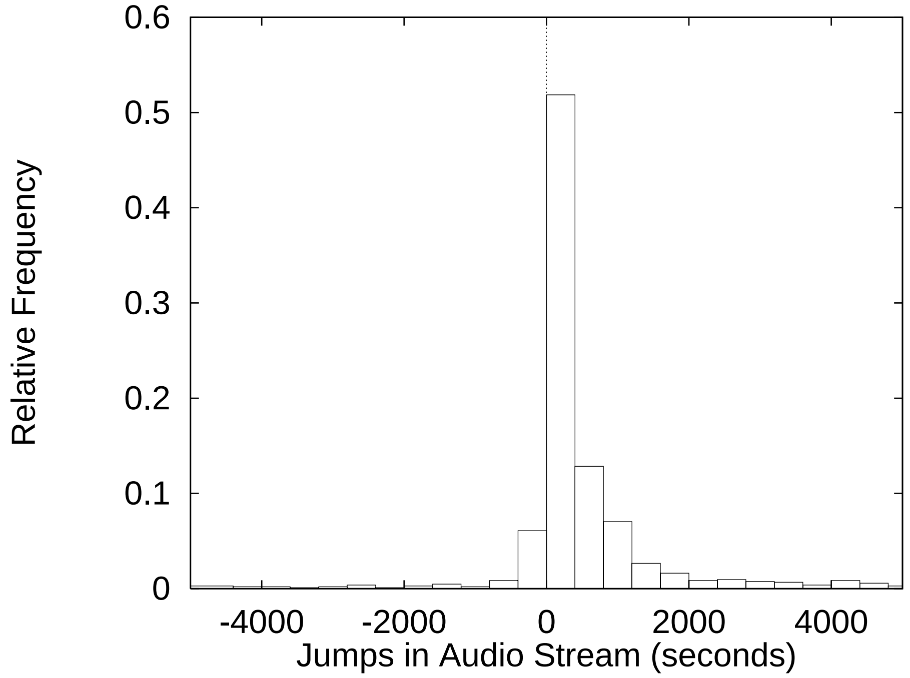


Figure 10: Distribution of lengths of jumps in the audio

a highlight. This results in slight overestimation of jump lengths. However, the error is no worse than length of the audio segment associated with one highlight.

We are interested in the length of these jumps for two reasons. First, if backward jumps are short, it might be possible to begin playout after a short backward jump without reloading new audio data from the server; playout could begin immediately using the already downloaded (and previously played-out and stored) audio data. Second, if forward jumps are short, playout can also begin immediately if the audio data has already been preloaded into the buffer as a result of server workahead [15, 12, 5].

In Figure 10, we plot the distribution of length of forward and backward jumps. Interestingly, there are very few backward jumps, and forward jumps are much more frequent. We had expected the students to make frequent backward jumps to consult previously viewed material. (This unexpected result may be due to the fact that many students might have been using MANIC for review purposes, rather than first-time viewing of the material.) The average length of a jump is 2137.62 seconds in case of forward jumps and 3394.84 seconds in case of backward jumps. Approximately 34.37% of the forward jumps and 33.33% of the backwards jumps were less than 3 minutes in length. Quite interestingly, Figure 10 indicates only a limited amount of "locality" in the jumps.

5 Consequences for CM Architectures and Modeling

Given the empirical and analytic results of the previous sections, we now consider the implications of these results for continuous media architecture and modeling.

Perhaps the most striking observed behavior is the high degree of user interactivity. Recall, for example, that once audio playout begins, in 45% of the cases it lasts less than 3 minutes before the user pauses or jumps elsewhere in the audio stream. This observed user behavior differs significantly from that of a "couch potato" user who initiates a CM playout that continues uninterrupted for a long period of time. From a resource allocation standpoint, a high level of interactivity would argue against prefetching significant amounts of CM material before its playout. While some amount of workahead is needed to remove network jitter [14] and reduce burstiness in the CM stream [15], downloading CM data too far in advance of its playout time can waste network and server resources by downloading data that is never actually played. From a network protocol standpoint, this would also argue against a rate control protocol (e.g., TCP) that simply downloads (prefetches) data as fast as the network will allow.

Two surprising observations in our data were the relatively small number of backward jumps ("rewinds") and the relatively large distance covered by jumps in the CM stream. The relatively small number of backward jumps suggests that relatively few interactive requests (fast forward/rewind/index) could possibly be satisfied by replaying previously played (but still buffered) CM. The latter observation suggests that it might also be difficult to satisfy a forward jump using prefetched (but not yet played out) CM. For example, recall that only a third of the observed forward jumps were less than three minutes forward. Even with a low bandwidth CM application such as 14.4K audio, three minutes of prefetched data is often non-trivial amount of data (more than a Megabit of data in this example), particularly for low bandwidth connections. Clearly, the advantage of potentially masking the latency of a forward jump by prefetching data must be weighed against the cost of prefetching data that is never played. This is an interesting question for future research. We also note that while the population as whole exhibits relatively little locality in their jumps, certain individuals may exhibit high locality in their jumping. In such cases, a model-based prefetching strategy that considers past user-specific actions [18] or makes content-specific prefetching decisions [10] may be able to successfully hide jump latency by buffering old data and/or prefetching future CM data. This too remains an issue for future research.

We noted that almost half of the MANIC sessions do not use audio and among those sessions that use audio, it is used only approximately two thirds of the time. Several observations can be made here. First, a CM server that allocates resources on the basis of worst case demand (e.g., assumes that each session will always be playing audio) will clearly over-reserve resources. A more reasonable basis for resource allocation would be the so called on/off model [3, 17] in which a user moves between "on" periods (where audio is streaming to the user) and "off" periods (where no audio is being streamed). A call admission policy based on such a model could then take advantage of the statistical multiplexing gains inherent in any population of users that are not always "on." The empirical data presented in this paper provides parameters for such models. The fact that half of the sessions use no audio at all also makes a compelling argument for handling the components of the CM stream (text, audio, video) as separate elements, allowing the heterogeneous client end systems to request and process/synchronize these streams according to their own desires.

From a modeling standpoint, our empirical data can be used in future simulation and analytic studies of CM systems at the session level, as well as at the detailed level of the activity within a session. We observed that the exponential distribution was a poorer fit to the empirical data than either the lognormal or gamma distributions. We also noted that the performance measures we have analytically characterized (the length of a session, the length of time spent in the audio "on" and "off" states) have a heavier tail than that of the exponential distribution. The behavior of heavy-tailed traffic sources has been the focus of much recent research in the networking community and the problems of modeling heavy-tailed phenomena with exponential models noted [11]. However, the consequence of such behavior for a CM system remains an interesting open question.

6 Conclusions

In this paper we have provided both empirical and analytic characterizations of observed user behavior with MANIC (Multimedia Asynchronous Networked Individualized Courseware), a WWW-based multimedia client/server application, that streams synchronized CM and HTML documents to remote users. Our goal was to characterize how people actually *use* such a CM system. Understanding how users interact with such a system – i.e., developing a user workload characterization – is crucial in designing and evaluating efficient CM resource allocation and access mechanisms. We examined both session-level characteristics (e.g., the length of the individual sessions) as well as *interactive* user behavior (e.g., the time between starting/stopping/pausing the audio within a session; and the "distance" of rewinding, fast forwarding and indexed jumping within the audio stream). We also considered the consequences of our measurements for CM architectures and their models.

A number of avenues for future work remain. We are currently investigating network-level congestion control protocols for CM, and the effect of client interactivity on the amount of CM data that should be transmitted in advance to the client. We are also currently integrating video into the MANIC system. More rigorous goodness of fits studies [2] can be performed to quantify the benefits of one analytic distribution over another. Finally, even though one distribution may more accurately model the workload than another, the overall effect of using different workload models in various performance studies is another topic for future research.

Acknowledgments

The MANIC system is the collective effort of many people, including the authors. Hu Im Lee, Jesse Steinberg, Mia Stern, and Gary Wallace have been instrumental in the design and implementation of MANIC; without their work these measurements would not have been possible.

References

[1] Anderson D. P., Osawa Y., Govindan R., "A File system for continuous media," *ACM Transactions on Computer Systems 10,* 4 (Nov. 1992), 311-337

[2] R. D'Agostino, Goodness of Fit Techniques, Marcel Decker, 1980.

[3] J. Dey Sircar, J Salehi, J.F. Kurose, D. Towsley, "Providing VCR-like Capabilities in Large-Scale Video Servers", in *Proc. ACM Multimedia 1994,* Oct 1994.

[4] J. Dey, S. Sen, J. Kurose, D. Towsley, J. Salehi, "Playback Restart in Interactive Streaming Video Applications," *Proc. IEEE Conference on Multimedia Computer and Systems," (June 1997 Ottawa Canada).*

[5] W. Feng, F. Jahanian, S. Secrest, "An Optimal Bandwidth Allocation Strategy for the Delivery of Compressed, Prerecorded Video," *Multimedia Systems,* Sept. 1997, pp. 297-309.

[6] J. Gemmel, H. Vin, D. Kandlur, P.V. Rangan, "Storage Servers: A Tutorial and Survey", *IEEE Computer,* May 1995.

[7] J.F. Kurose, H.I. Lee, J. Padhye, J. Steinberg, M. Stern, 'MANIC: Multimedia Asynchronous Networked Individualized Courseware," Technical Report 96-72, Department of Computer Science, University of Massachusetts, Amherst MA 01003. Oct. 1996. `URL: ftp://gaia.cs.umass.edu/pub/Kuro96:MANIC.ps` *The MANIC homepage is* `http://gaia.cs.umass.edu/nmis/manic.html`.

[8] T. Kwon, Y. Choi, S. Lee, "Disk Placement for Arbitrary Rate Playback in an Interactive Video Server," *Multimedia Systems,* July 1997, pp. 271-281.

[9] K. Nahrstedt, R. Steinmetz, "Resource Management in Networked Multimedia Operating Systems," *IEEE Computer,* May 1995, pp 40–49.

[10] V. Padmanabhan and J. Mogul, "Using Predictive Prefetching to Improve World Wide Web Latency", *Proceedings of ACM SIGComm*, 1996, pp. 22-36.

[11] V. Paxson and S. Floyd, "Wide Area Traffic: the Failure of Poisson Modeling,"*IEEE/ACM Transactions on Networking,* Vol. 3, No. 3, (June 1995), pp. 226 - 244.

[12] M. Reisslein and K. Ross, "Join the Shortest Queue Prefetching for VBR Encoded Video on Demand," *Proc.19997 Int. Conference on Networking Protocols,* (Oct. 1997, Atlanta GA).

[13] Real Audio homepage: http://www.realaudio.com.

[14] K. Rothermel, T. Helbig, "An Adaptive Protocol for synchronizing Multimedia Systems," *Multimedia Systems,* Sept. 1997, pp. 324-336.

[15] J. Salehi, Z. Zhang, J. Kurose, D. Towsley, "Supporting Stored Video: Reducing Rate Variability and End-to-end Resource Requirements through Optimal Smoothing," *Proc. 1996 ACM Sigmetrics Conference,* (May 1996, Philadelphia PA).

[16] P. Shenoy, P. Goyal, H. Vin, "Issues in Multimedia Server Design," *ACM Computing Surveys,* December 1995.

[17] P. Shenoy, H. Vin, "Efficient Support for Scan Operations in Video Servers," *Proc. ACM Multimedia 95,* (Nov. 5-9, San Francisco).

[18] Mia K. Stern, Beverly Park Woolf, and James F. Kurose, "Intelligence on the Web?", *Proceedings of the 8th World Conference of the AIED Society,* August 18-22, 1997, Kobe, Japan.

Appendix

Factors affecting measurement accuracy

Our measurement accuracy is affected by several factors. In this appendix, we describe these factors and discuss their impact on our work.

- Use of Netscape and RealAudio plugin controls

  Recall that as discussed in Section 2, MANIC provides controls to both navigate through the class notes and control the audio. The CGI scripts that maintain the MANIC logs were based on a (naive, in hindsight) assumption that users would refrain from using browser and Real Audio plugin controls to move through the course materials[7] since this functionality is explicitly provided by MANIC. There were, however, users that used the browser/plugin controls (e.g,. forward and backward buttons on the browser, bookmark lists, history index, "pause" buttons on the Real Audio plugin etc.) – actions that would not appear in the MANIC logs but (for audio-related actions) would appear in the Real Audio logs. In order to account for this in the audio ON and OFF length results presented in section 4, we explicitly correlated the MANIC logs with the separate logs maintained by the RealAudio server to account for such actions.

- Measurement of forward and backward jumps

  As noted above, users could potentially make small forward and/or backward jumps in the audio stream using the controls on the RealAudio plugin. The RealAudio logs do not contain sufficient information to determine the lengths of these jumps. It is also difficult to determine lengths of such jumps from the logs maintained by our CGI scripts. Thus, in our analysis of jump lengths, we have only considered the jumps made using MANIC-provided controls.

- Network and processing delays

  Because of the network and processing delays, RealAudio log entries can be written later than when a corresponding event is noted in the MANIC logs. For example, while the MANIC log records a start audio command at time t, the RealAudio server may not begin actually playing the audio until time $t + \delta$. In most cases, δ is small (e.g., less that 15 seconds).In those few cases where δ was larger than two minutes, we considered the resulting silence at the client (waiting for audio playout to begin) to be part of an audio OFF period.

- Small breaks in the audio session can be hard to attribute

  Such breaks can arise due to several factors, including temporary network congestion and multiple clicks by the user on the "Play" buttons. It is hard to distinguish between the various factors that may cause of such breaks, especially since they are all of the same magnitude as the network delay. For example, if the RealAudio

[7] This problem gave rise to almost all the issues discussed here.

logs show that the playout started, say 10 seconds, after the time indicated by the timestamp of the corresponding MANIC log entry, we can not conclude if this was due to some or all of the following factors: network delay, processing overhead at the RealAudio server, user clicking on “pause” button just after he clicked on “play” etc. This results in our slightly overestimating the length of audio subsessions. We estimate that the magnitude of the overestimate is less than 4%.

- Use of multiple IP addressed and proxies by students

  If a user used multiple IP addresses simultaneously, or used a proxy to receive RealAudio content, it became difficult to match the MANIC and RealAudio logs. We had to exclude logs of 5 users from our consideration, as their logs contained such discrepancies on an extensive scale.