Packet Loss Correlation in the MBone Multicast Network: Experimental Measurements and Markov Chain Models

Maya YAJNIK, James F. KUROSE, and Don TOWSLEY

CMPSCI Technical Report 95-115

Packet Loss Correlation in the MBone Multicast Network: Experimental Measurements and Markov Chain Models[1]

Maya Yajnik, Jim Kurose, and Don Towsley
Department of Computer Science
University of Massachusetts at Amherst
Amherst MA 01003
{yajnik,kurose,towsley}@cs.umass.edu

UMASS CMPSCI Technical Report # 95-115

Abstract

The recent success of multicast applications such as Internet teleconferencing illustrates the tremendous potential of applications built upon wide-area multicast communication services. A critical issue for such multicast applications and the higher layer protocols that support them is the manner in which packet losses occur within the multicast network. In this paper we present and analyze packet loss data collected via experiments run on 14 multicast-capable hosts at 11 geographically distinct locations in Europe and the US and connected via the MBone. In this work we experimentally and quantitatively examine the *spatial and temporal correlation in packet loss* among participants in a multicast session. Our results show that there is a significant spatial correlation in loss among the multicast sites. We also find a fairly significant amount of of burst loss (consecutive losses) at a site. We also use these empirical measurements to derive a Markov chain characterization of temporal loss correlation.

[1]This work was supported in part by the National Science Foundation under grant NCR-911618

1 Introduction

The recent success of multicast applications such as Internet teleconferencing tools [6, 12] for audio [22, 10, 4], video [9, 3], and whiteboard [11], and distributed interactive simulation illustrates the tremendous potential of applications built upon wide-area multicast communication services. A critical issue for such multicast applications and the higher layer protocols that support them is the manner in which packet losses occur within the multicast network.

In this paper we present and analyze packet loss data collected via experiments run on 14 multicast-capable hosts at 11 geographically distinct locations in Europe and the US. These hosts are connected via the Multicast Backbone (MBone) network [6, 15]. The primary goal of this work is to examine the *spatial and temporal correlation in packet loss* among participants in a multicast session. (Informally, by "spatially" correlated loss, we mean the loss, i.e., lack of reception, of the same packet at many sites; by "temporally" correlated loss, we mean the loss of consecutive packets at a given receiver.) We also use these empirical measurements to derive a Markov chain characterization of temporal loss correlation. Our results show that:

- There is a significant correlation in loss among the multicast sites. For example, in one two-hour period, we find that almost 70% of transmitted packets were not received by one or more sites, and that nearly 30% of the packets transmitted were lost at two or more sites. This implies that over all multicast participants, a packet will frequently not be received at one, two, or more sites.

- There is a significant amount of burst loss (consecutive losses) at each site. For example, in this same two hour trace, we find that the length of the longest number of consecutive packet losses is more than 100 packets at each of the sites, and that the average number of consecutive losses varies from approximately 2 to 8 among the sites.

- Using "entropy" as a measure of the uncertainty in predicting whether the next packet will be lost at a receiver, given that the history of the last n packets is known, we find that an eight-state Markov chain (i.e., a value of $n = 3$) is appropriate for modeling the temporal correlation in loss.

The underlying packet loss process is of tremendous importance to error control protocols. This is particularly so with multicast communication, since many of the proposed error control protocols cited below recover from packet loss by having receivers interact with other receivers rather than with the data source itself. Thus, the spatial correlation of loss is of particular importance. Although there has been a considerable amount of research on multicast error control protocols [1, 2, 4, 5, 12, 18, 19, 20, 23, 24], these works have either not examined or considered the underlying loss process, or have assumed that packet losses are both spatially and temporally independent; the two exceptions are [1, 4]. The work by Bhagwat *et al.* [1] describes

a recursive analytic method for computing the probability that a packet is not received at one or more receivers given a specific multicast tree and known, independent loss probabilities on each link. The work by Bolot *et al.* is the work most closely related to our present work. In that work, packet loss measurements are presented from a 10,000-packet trace between MBone sites in France and England. With respect to temporally-correlated loss, they find that "losses appear to be isolated" – a result somewhat different from ours; they do not address the issue of spatially correlated losses.

The remainder of this paper is structured as follows. In the following section we describe the measurement tools we constructed and the experiments conducted. In section 3, we examine the spatial correlation of loss in the packet traces. In section 4, we examine the temporal correlation in loss at two specific sites. We also describe data analyses which indicate that a eight-state Markov chain is suitable for characterizing the temporal correlation in packet loss. Section 5 concludes this paper.

2 Experimental Background

Our experiments consisted of simultaneously monitoring and recording the received multicast packet transmissions of the "World Radio Network" at 14 hosts at 11 different Mbone sites. Table 1 lists the receiving hosts and their geographical locations, IP addresses, operating systems and machine types. At three of these sites (those in California, Massachusetts, and Germany), there was an additional multicast capable host (not shown in Table 1). However, unless otherwise noted, the measurements we report involve only the 11 hosts shown in Table 1.

The "World Radio Network" is an application run by the Internet Multicasting Service from Washington DC. It transmits multicast packets over the MBone at 80ms intervals, each of which contains approximately 5Kbits worth of audio data within a vat [10] audio packet. These vat audio packets, in turn, are sent as UDP datagrams which are then encapsulated in IP multicast packets. By listening to the WRN multicast address at each of the 11 sites, it is possible to determine which packets arrive and which are lost. (Note that while these packets contain audio data, our results are not tied to this specific application. We ignore the actual contents of these packets, essentially considering them as periodic test packets that are sent into the multicast network).

At each of the 11 receivers in Table 1, a process was run that listened to the WRN multicast address and recorded and timestamped the vat headers of the arriving WRN packets. The packet header contained a sequence number which uniquely identified each multicast packet sent by the WRN. These data collection daemons were remotely controlled by commands sent from a central control program to start, stop, and otherwise control the data collection daemons. Once the data was collected, the control program instructed the daemons to send the trace files

Workstation Name	Location	IP Address	Operating System	Machine Type
alps	Georgia	130.207.8.16	SunOS 4.1.3	sun4m
anhur	Sweden	192.16.123.94	SunOS 4.1.3	sun4m
cedar	Texas	128.83.141.15	SunOS 5.3	sun4m
collage	California	192.100.58.17	SunOS 4.1.3	sun4c
erlang	Massachusetts	128.119.40.203	SunOS 4.1.2	sun4c
float	Virginia	128.143.71.21	SunOS 5.3	sun4m
law	California	128.32.33.106	IRIX	IP20
tove	Maryland	128.8.128.42	SunOS 4.1.3	sun4m
ursa	Germany	192.35.149.160	SunOS 4.1.3	sun4m
willow	Arizona	192.12.69.86	SunOS 4.1.1	sun4c
zen	Missouri	128.252.169.30	SunOS 4.1.3	sun4m

Table 1: The eleven receiving hosts

via ftp to our centralized site. The traces described in this paper were collected at two times on July 13, 1995. The first trace is for a 2-hour period starting from 13:45 EDT, corresponding to a total of 90,000 packets. The second trace began at 22:55 PM of the same day and lasted for one hour (45,000 packets). We will subsequently refer to these traces as the "day" and "night" traces respectively. These times were chosen roughly to correspond to times when one might expect the MBone to be rather busy and rather unused. These traces, as well as others that we have obtained, can be obtained from our ftp site ftp://gaia.cs.umass.edu/pub/yajnik.

While the spatial and temporal correlation of loss can be measured and characterized without knowing the specific MBone topology connecting the 11 multicast sites in our experiments, we will find this information useful when interpreting the spatial correlation of the loss measurements. To understand this topology, it should first be noted that the MBone itself is a virtual network built on top of the Internet. The nodes in the MBone are multicast-capable routers logically connected to each other via IP routes known as "tunnels." Thus, a logical point-to-point connection between two multicast routers will typically contains many IP routers.

Figure 1 provides a *logical* view of the multicast connectivity among the 11 sites in our experiments. Each of the eleven hosts are shown, as are *selected* intermediate MBone routers between the hosts and the WRN root. Each selected MBone router shown is a nearest common ancestor of all downstream end hosts on the multicast tree. The multicast tree itself was constructed by joining together the multicast paths from each of the end hosts to WRN. These multicast paths were determined using the multicast `ping` program with the record-route option (which returns a full path along the MBone from source to destination) and by querying intermediate multicast routers using the `mrinfo` program (which returns information about a

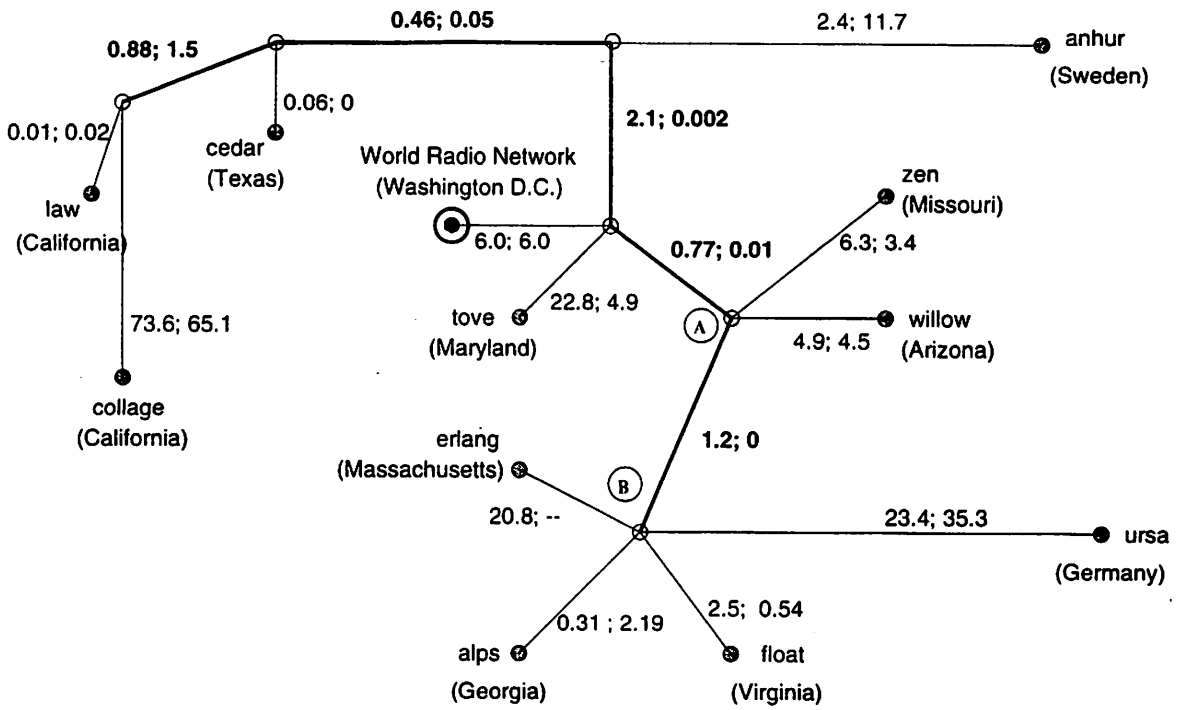

Figure 1: Transmission Tree with Losses

multicast capable router's connectivity to other multicast capable routers). We ran the `ping` and `mrinfo` programs before and after gathering data, and found that the multicast routes had not changed. The numbers labeling the virtual links in Figure 1 will be discussed later, in the following section.

3 Spatial Correlation of Loss

In this section, we examine the spatial correlation in packet loss among the receiving hosts.

Figures 2 and 3 provide histograms of the percentage of packets lost at exactly one of the receivers, by exactly two receivers and so on, for the day and night traces. The workstation, `collage`, was not included in the statistics either trace because of its extraordinarily high loss rate. The workstation `erlang` was not used to take measurements in the night trace.

A number of interesting observations can be made from the histogram results. First, note that even excluding `collage`, 67% of transmitted packets for the day trace and 58% of packets for the night traces were lost by at least one receiver. This implies that in a sender-initiated reliable transport protocol (e.g., a protocol in which receiving hosts request a retransmission

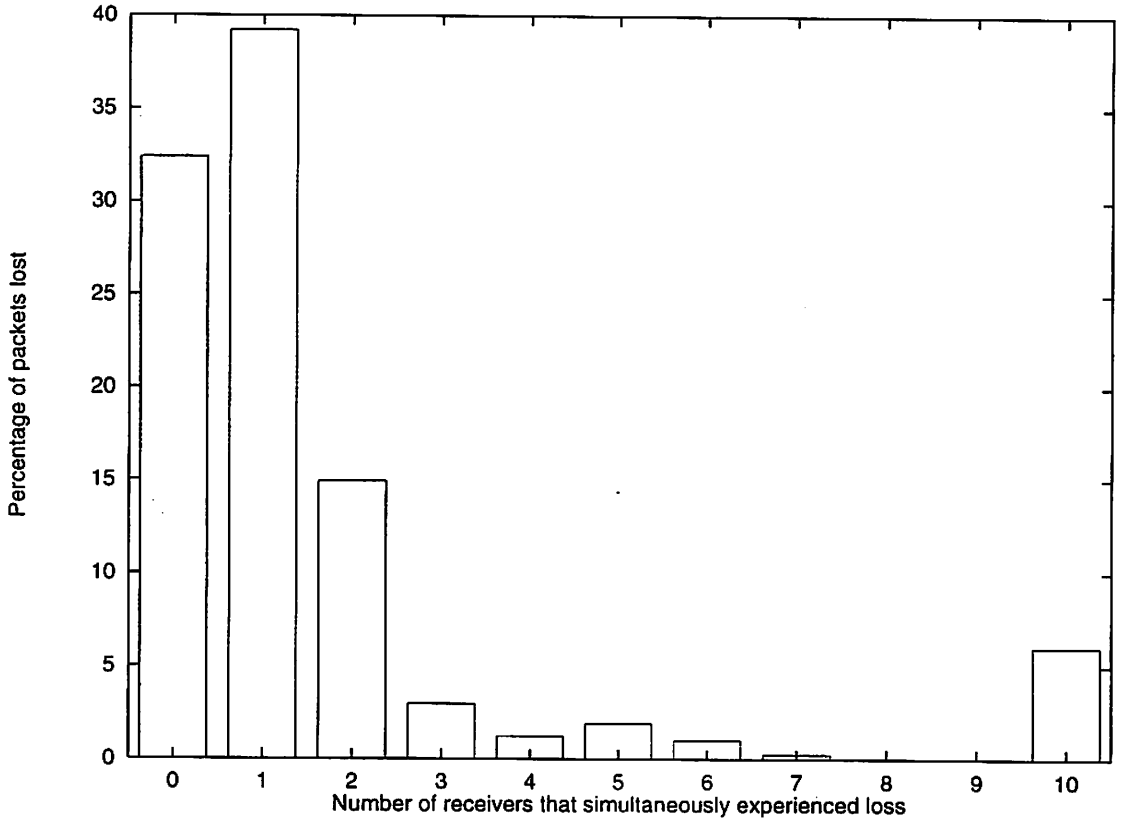

Figure 2: Simultaneous Losses: day trace

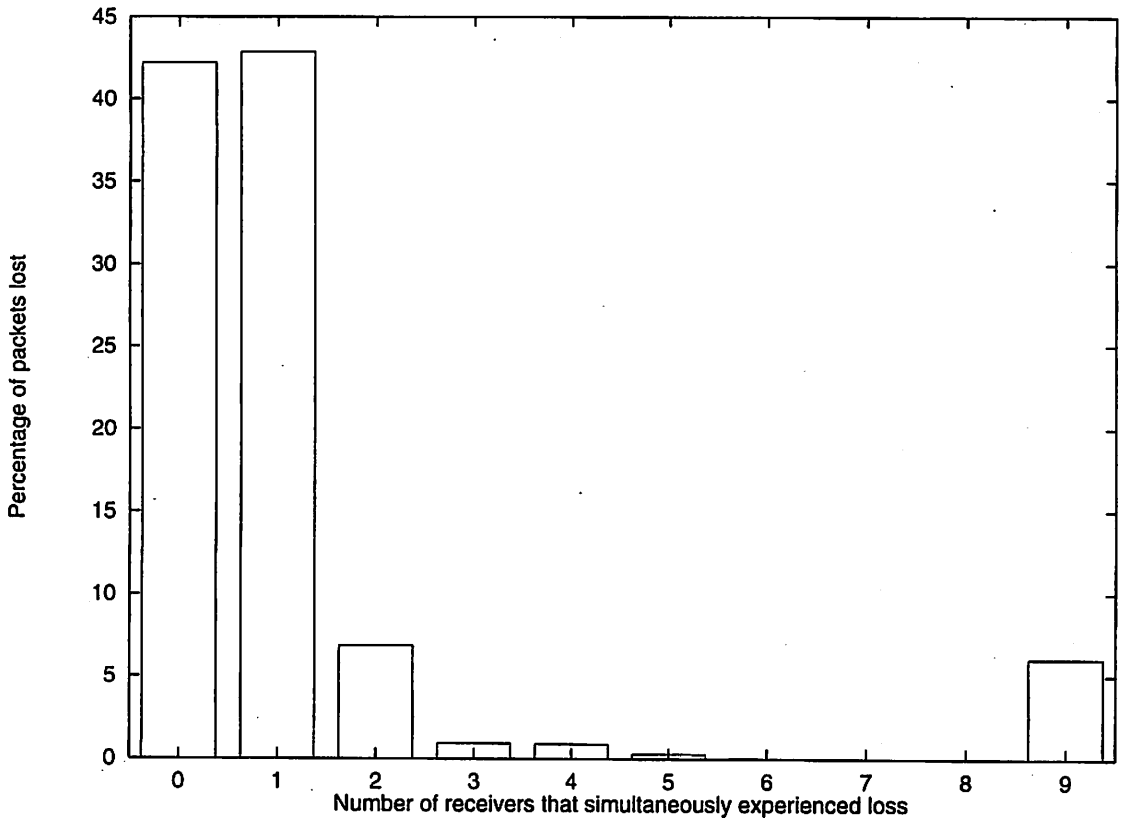

Figure 3: Simultaneous Losses: night trace

directly from the sender), the sender might need to transmit a packet several times, since the probability that a transmitted multicast packet is received at all sites (and hence would not require retransmission) is only .33 and .43 respectively. Second, note that 28% of the packets in the day trace and 15% percent of packets in the night trace were simultaneously lost at more than one receiver. By this measure, there is a significant amount of correlation in the loss process. It is also interesting to note, however, that 80% of the packet losses in the day trace and 86% of the packets in the night trace were lost at only one or two receivers. This would indicate that it in the context of receiver-initiated error recovery protocols, a receiver that lost a packet may be able to recover that packet from another receiver that is closer to it than the source. Finally, we note that 6% of the packets were simultaneously lost by all ten receivers in both traces – suggesting that this packet loss occurred close to the WRN source.

For a given packet, examining which receivers received the packet and which did not can provide a valuable clue as to *where* in the multicast tree the packet was lost. For example, in the multicast tree in Figure 1, if a packet is lost at `erlang`, `alps`, `float` and `ursa`, but received correctly at `willow`, it is likely that the packet was dropped between the multicast routers A and B. (It should be noted, however, that this need not be the case, as the packet could have been simultaneously and independently lost on all four downstream paths from B, although we consider this latter scenario to be much less likely.) Using this reasoning, we can determine the approximate percentage of packets lost on each of the links in Figure 1. The numbers shown for each link in Figure 1 indicate the approximate percentage of packets lost on that link in the day and night traces, respectively.

An interesting observation is that the loss on the "backbone" (i.e., between any two routers) is generally much smaller than between a router and an end host. This would suggest that loss is more likely to occur near the "edges" of the multicast network. (We add a caution here, however, that there may be multiple MBone routers not shown in Figure 1 that connect an end host to one of the MBone routers shown in the figure). That is, once a packet enters the network backbone , it is likely to be multicast successfully throughout the backbone and only be lost at the networks edges. Thus, if a packet loss occurs at the network edge, the receiver losing the packet may still be able to recover the packet from a "nearby" receiver that had a different route into the backbone network. For example, if a packet is lost on the path from B to `float` in Figure 1, may be able to retrieve the lost packet from `alps`, `erlang`, or `float` since a loss on the B-to-`float` path does not imply a loss at the other receivers. In the context of reliable multicast, this would thus suggest that local recovery from another receiver in times of loss would often be possible.

The final set of measurements we made regarding the spatial locality of loss was to determine whether any packets were being dropped at the receiving hosts themselves. To do so, we monitored the multicast at two different workstations on the same end local area network. We did this at three sites: `collage` (in California), `erlang` (in Massachusetts) and `ursa` (in Germany). Surprisingly, the end-host loss was found to be negligible. It was zero at `collage`

and ursa and 0.001% at erlang.

4 Temporal Correlation of Loss at a Single Receiver

This section describes the distribution of the lengths of these bursts of losses.

Tables 2 and 3 summarize the loss rates experienced by each of the receivers during the day and during the night. They also show the number of lossy bursts seen by each workstation, the average length of the bursts and the coefficient of variation of the burst length. The Coefficient of variation is defined as

$$c = \frac{\sqrt{E[(b - \bar{b})^2]}}{\bar{b}} \tag{1}$$

where b is the burst length or the number of consecutive losses and $\bar{b}$ is the mean burst length.

In general, the coefficient of variation was seen to be very high ranging from 1.640 at alps (Georgia) to 11.620 at erlang (Massachusetts), for the day trace. Long bursts of losses were observed. For example, for the day trace, alps saw long periods of loss upto 122 packets which is effectively 9.76s and anhur (Sweden) saw loss periods upto 2203 packets, equivalent to 176.24s.

Figures 4 and 5 show the distribution of burst lengths for two receivers, law and zen. It is clear that, although most of the bursts are of size less than 100, there are occasional long loss periods. Figures 6 and 7 show the percentage of total loss experienced in the various burst lengths. In the case of law, a loss period of length 1867 contributed to 22% of the loss.

Figures 8 and 9 show the burst length distributions for the night trace and Figures 10 and 11 show the corresponding variation in percentage of loss with burst length. The longest bursts seen in the night trace were not as long as the longest bursts seen by day. However, in the case of the receiver law, a single burst of 358 consecutive losses contributed to 11% of the total loss.

The cumulative distribution functions for this data are shown in Figures 12 and 13. These graphs show, for all burst lengths, b, the number of bursts of length b or smaller, and the corresponding percentage of overall loss that occurred in bursts length b or smaller. Again it is apparent that loss periods more than 100 packets long contribute heavily to the total loss, despite their small numbers.

Workstation Name	Loss Rate	Number of Lossy Bursts	Average Burst Length	Coefficient of Variation	Length of the Longest Burst
alps (Georgia)	8.27%	2194	3.392	1.640	122
anhur (Sweden)	10.46%	1150	8.186	8.249	2203
cedar (Texas)	8.58%	1337	5.778	8.874	1867
collage (California)	83.01%	510	146.484	9.492	17690
erlang (Massachusetts)	28.76%	2234	11.59	11.620	2993
float (Virginia)	10.48%	2450	3.851	8.840	1665
law (California)	9.42%	1670	5.076	9.123	1867
tove (Maryland)	28.77%	10046	2.578	12.409	2222
ursa (Germany)	31.37%	8480	3.330	6.746	567
willow (Arizona)	11.67%	3957	2.655	1.871	118
zen (Missouri)	13.08%	5423	2.170	4.118	437

Table 2: Burstiness of Loss: day trace

Workstation Name	Loss Rate	Number of Lossy Bursts	Average Burst Length	Coefficient of Variation	Length of the Longest Burst
alps (Georgia)	8.24%	1091	3.401	2.310	96
anhur (Sweden)	17.75%	618	12.927	4.005	522
cedar (Texas)	6.10%	427	6.424	1.845	96
collage (California)	72.68%	181	180.691	6.093	9011
float (Virginia)	6.60%	567	5.236	2.055	96
law (California)	7.57%	488	6.980	3.047	358
tove (Maryland)	10.96%	1975	2.496	2.479	96
ursa (Germany)	41.33%	1797	10.349	4.920	762
willow (Arizona)	10.58%	1640	2.904	2.323	96
zen (Missouri)	9.46%	1637	2.599	2.526	96

Table 3: Burstiness of Loss: night trace

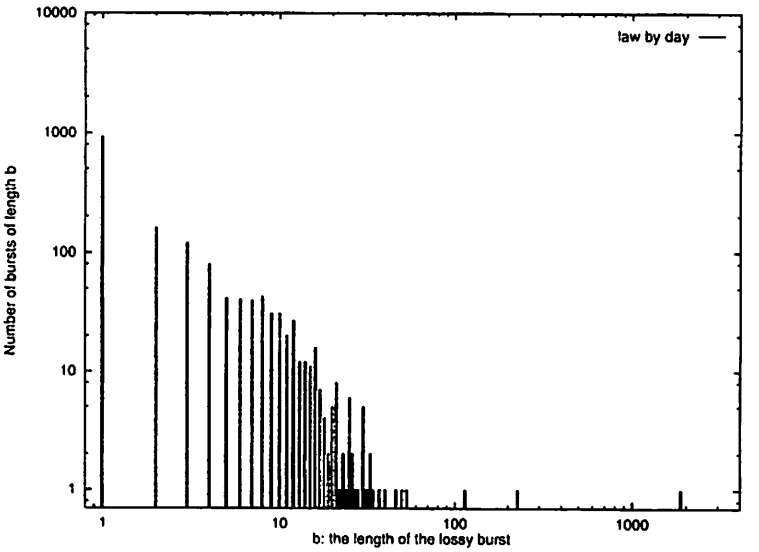

Figure 4: Distribution of Burst Length for "law": day trace

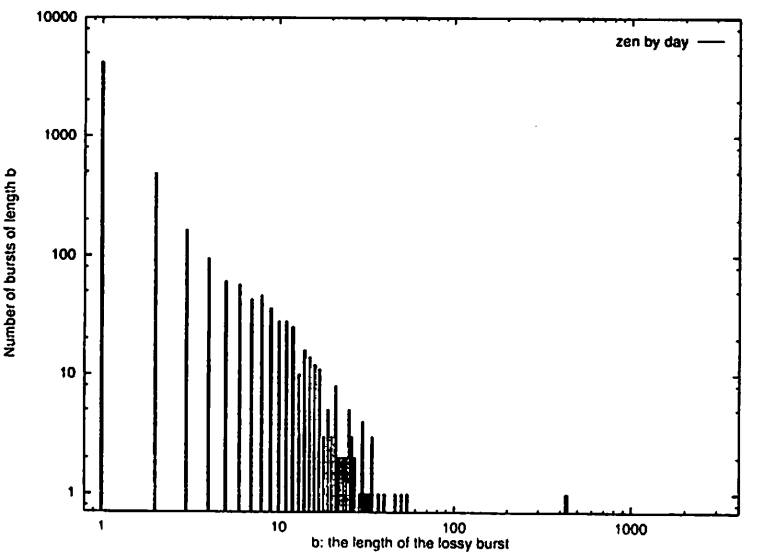

Figure 5: Distribution of Burst Length for "zen": day trace

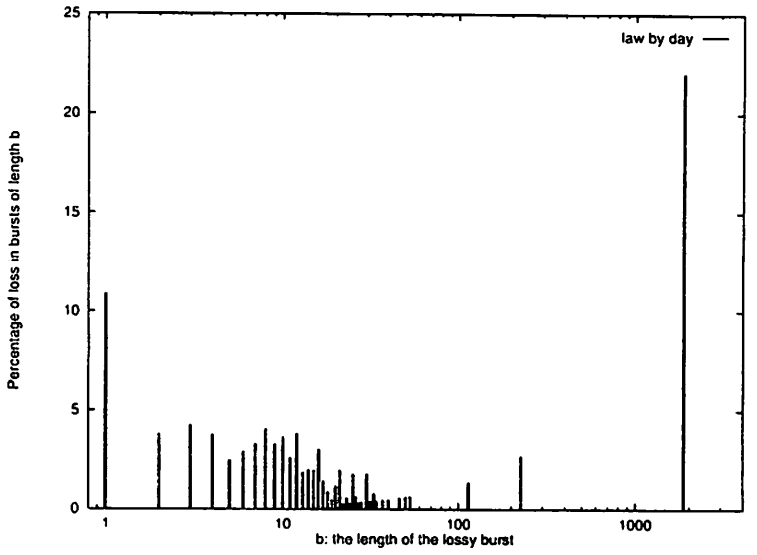

Figure 6: Percentage of Packet loss for "law": day trace

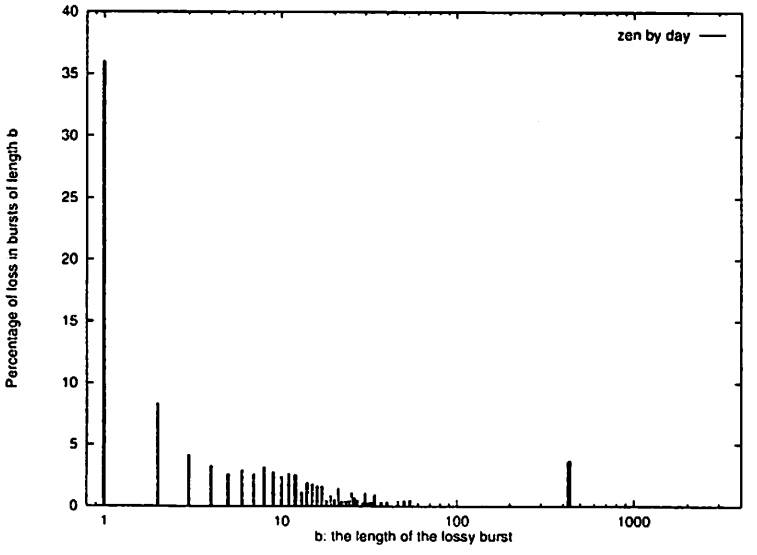

Figure 7: Percentage of Packet loss for "zen": day trace

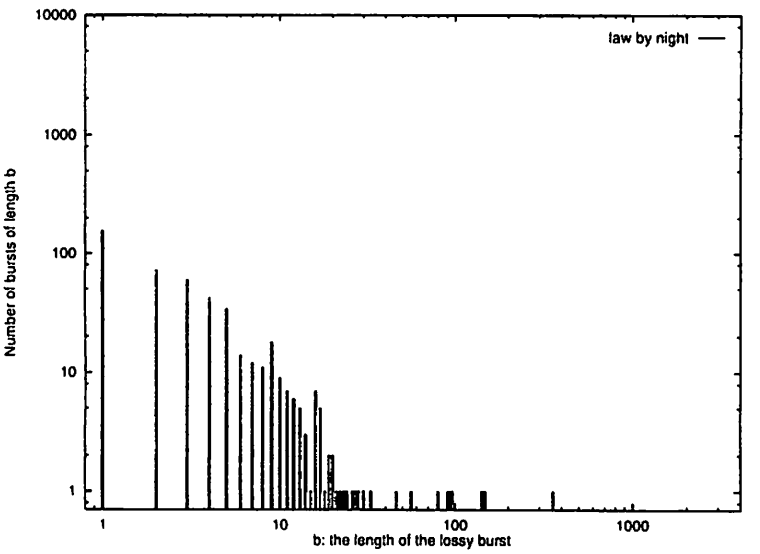

Figure 8: Distribution of Burst Length for "law": night trace

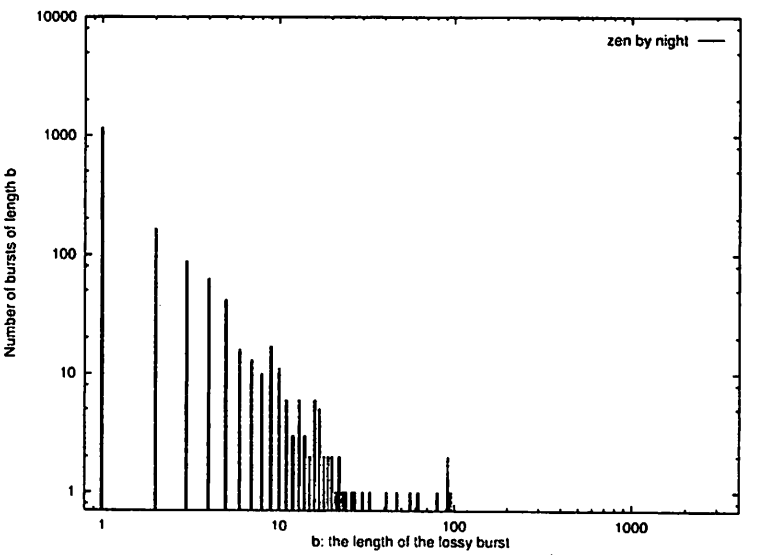

Figure 9: Distribution of Burst Length for "zen": night trace

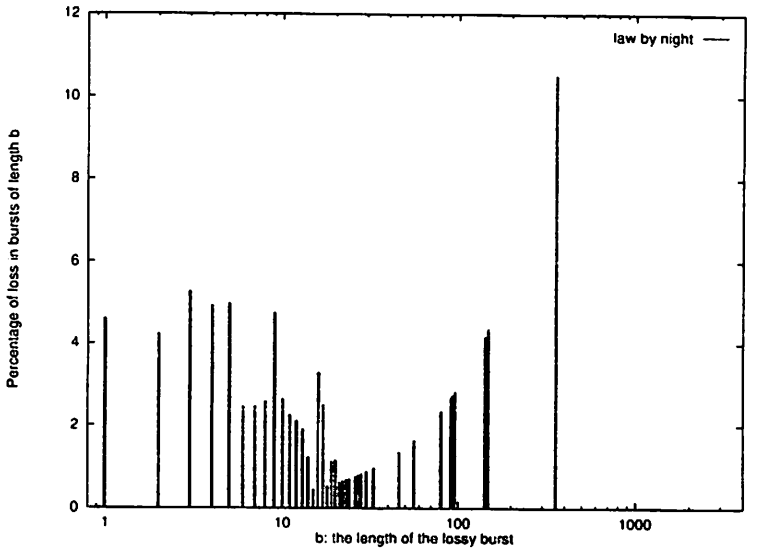

Figure 10: Percentage of Packet loss for "law": night trace

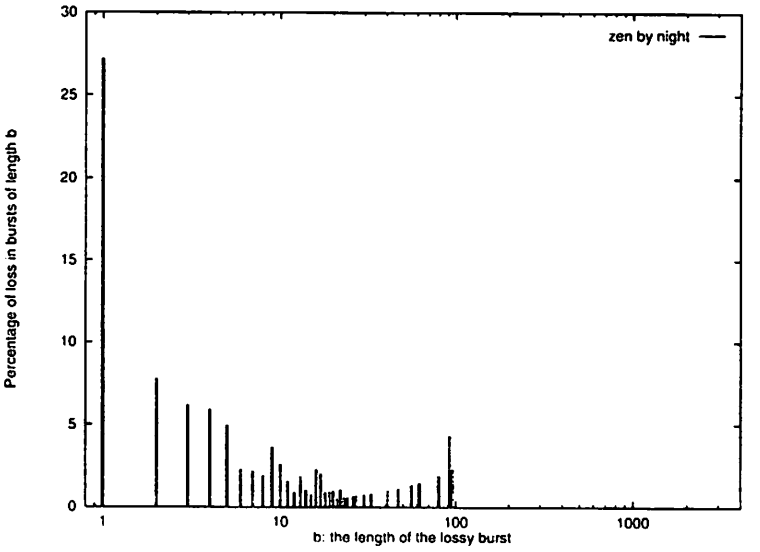

Figure 11: Percentage of Packet loss for "zen": night trace

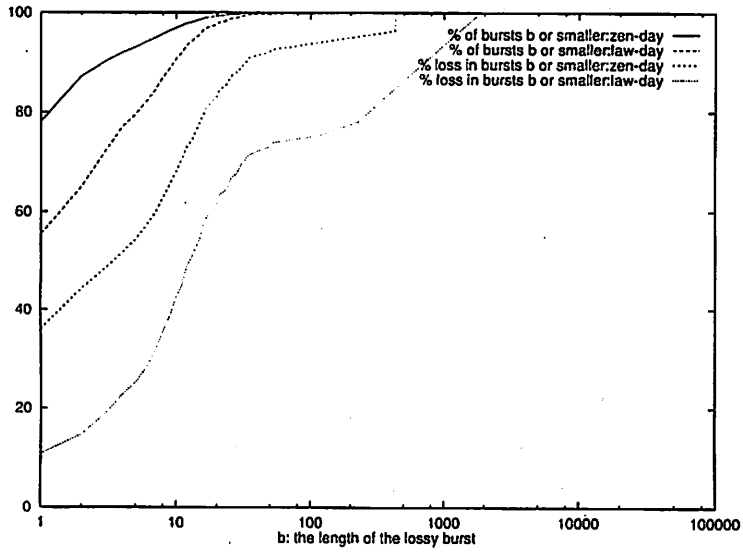

Figure 12: Cumulative Distribution Function: day trace

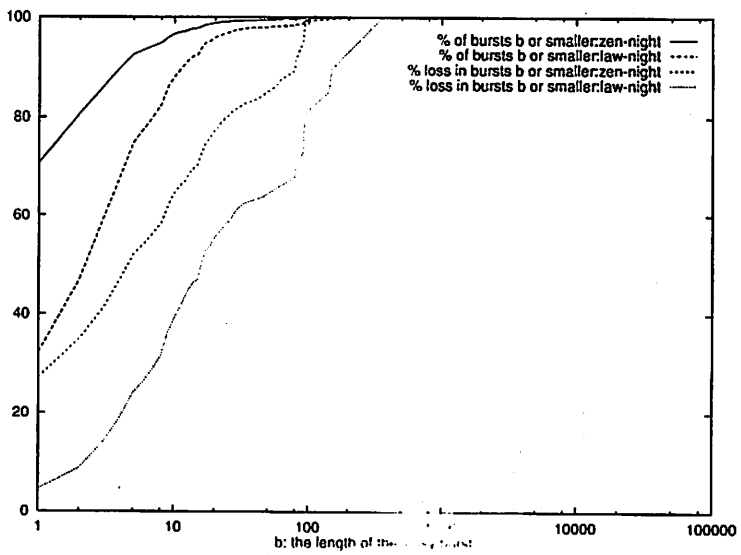

Figure 13: Cumulative Distribution Function: night trace

5 A Markov Model

One way to examine the temporal correlation in loss is to look at the packet loss events seen by a receiver as a series of random variables which take on the value 0 or 1. These indicate the status of successive packets. If a packet is lost, the value of the corresponding variable is 1 and if it arrives at the destination then the value is 0. Packets are sent out by the source regularly at 80ms intervals. Thus this can be considered to be a discrete-time, binary-valued time series.

A class of random processes that is rich enough to capture a large variety of temporal dependencies and yet structured enough to lend itself to some meaningful analysis is the class of all Markov processes [14] [13]. We shall therefore model the packet loss process at a (fixed) receiver as finite-state Markov chain, the state of the process being determined by the history, *i.e.*, the previous values of the process.

Let $X_i, i = 1, 2, 3..$ be a stationary and ergodic discrete time Markov process taking values in the set $\mathcal{X} = \{0, 1\}$. Such a process is characterized by its order k, and by a conditional probability matrix P_k whose rows may be interpreted as probability mass functions on $\mathcal{X}$ according to which the next random variable X_i is generated when the processes is in a state $x_{i-k}x_{i-k+1}\cdots x_{i-1}$.

$$Prob(X_i = x_i | X_{i-1} = x_{i-1}, ..., X_{i-k} = x_{i-k}) = P_k(x_i | x_{i-1}, ..., x_{i-k}) \tag{2}$$

Let $\underline{x} = (x_1, x_2, ..., x_n)$ be an observed sequence or a sample path from an Markov source. The k-th order state transition probabilities of the Markov chain can be estimated (as observed in sequence $\underline{x}$) for all $a \in \mathcal{X}, \underline{b} \in \mathcal{X}^k$ as follows.

Let $l^k_{\underline{x}}(a, \underline{b})$ be the number of times state $\underline{b}$ is followed by value a in the sample sequence.

$$l^k_{\underline{x}}(a, \underline{b}) = \sum_{i=1}^{n} \mathbf{I}\{x_i = a, (x_{i-1}, ..., x_{i-k}) = \underline{b}\} \tag{3}$$

Let $l^k_{\underline{x}}(\underline{b})$ be the number of times state $\underline{b}$ is seen.

$$l^k_{\underline{x}}(\underline{b}) = \sum_{i=1}^{n} \mathbf{I}\{(x_{i-1}, ..., x_{i-k}) = \underline{b}\} \tag{4}$$

where $\mathbf{I}(\cdot)$ is the indicator function of an assertion defined as

$$\mathbf{I}\{A\} = \begin{cases} 1 & \text{if assertion } A \text{ is true} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Let $q^k_{\underline{x}}(a|\underline{b})$ be an estimate of the probability that $x_i = a$, given that $(x_{i-1}, ..., x_{i-k}) = \underline{b}$. Let $\underline{b} = (b_1, ..., b_k)$ be a given state of the chain. Then $q^k_{\underline{x}}(a|\underline{b})$ estimates the state transition probability from state $\underline{b}$ to state $(a, b_1, ..., b_{k-1})$. The estimates of the state transition probabilities of the k-th order Markov chain are

$$q^k_{\underline{x}}(a|\underline{b}) = \begin{cases} \frac{l^k_{\underline{x}}(a,\underline{b})}{l^k_{\underline{x}}(\underline{b})} & \text{if } l^k_{\underline{x}}(\underline{b}) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Table 4 shows the estimated state transition probabilities of a 3rd order Markov model for the trace of receiver, `alps`. Since it is a 3rd order model, the next value depends on the 3 previous values and there are 8 different states, one state for each of the possible relevant histories. The table indicates, for each state, the probability of 0 being the next observed value and the probability of 1 being the next observed value. Also shown is the probability of that state being visited.

An issue that needs to be resolved, is to determine the appropriate order (equivalently, the dimensionality of) the Markov Chain. To put it differently, we must estimate how many previous values of the process influence a current value. The order of the Markov chain is the number of previous values which influence the next value.

The problem of estimating the order of a Markov chain has been studied in [17] and [8] *etc.* We shall state these results, without proof, for the sake of completeness.

Let $\underline{x} = (x_1, x_2, ..., x_n)$ be an observed sequence or a sample path from an unknown k^*-th order Markov source. We will assume that although k^* is unknown, we know an upper bound K_0 on its possible values. Techniques for estimating the order without this assumption are also known, but we shall argue later that this assumption is justified in our case.

We define the empirical k-th order conditional **entropy** of $\underline{x}$ as

$$H(q^k_{\underline{x}}) = -\sum_{\underline{b} \in \mathcal{X}^k} q^k_{\underline{x}}(\underline{b}) \sum_{a \in \mathcal{X}} q^k_{\underline{x}}(a|\underline{b}) log q^k_{\underline{x}}(a|\underline{b}) \tag{7}$$

The convention is to set $0 \log 0$ to be equal to 0.

The notion of entropy is fundamental in prediction problems. In particular, the k-th order entropy of a sequence is a measure of randomness of the next variable when the previous k values are known. It is zero if the value of the next variable is completely predictable and is maximized when the next variable is totally random (i.e. has a uniform distribution). It's maximum value is $\log |\mathcal{X}|$.

Finally, the estimator $\hat{k}(\cdot)$ of the Markov order k is defined as

$$\hat{k}(\underline{x}) = \min_{0 \le k \le K_0} \left\{ k \mid H(q_{\underline{x}}^k) - H(q_{\underline{x}}^{K_0}) \le \epsilon_n \right\} \tag{8}$$

where ϵ_n is a threshold that will be specified later.

Thus the estimator $\hat{k}(\cdot)$ employs the philosophy that the shortest history that has (almost) the same predictive power as knowing the entire K_0-past must be the true order of the sequence. Any further knowledge of the past does not make the next variable significantly easier to predict.

Following the analysis in [8], we set

$$\epsilon_n = (2^{K_0+1} + \delta)\frac{\log n}{n} \tag{9}$$

where $\delta > 0$. For this choice of ϵ_n, it has been shown that for all Markov measures P_k of order $k < K_0$,

$$P_k\left(\hat{k}(\underline{X}) < k\right) \le 2^{-nD} \tag{10}$$

where $D > 0$ depends on P_k. Also,

$$P_k\left(\hat{k}(\underline{X}) > k\right) \le C \cdot n^{-\delta} \tag{11}$$

If $\delta > 1$ it can be shown that $\hat{k}(\cdot)$ is strongly consistent for all Markov measures of order $k < K_0$. *i.e.*

$$P_k\left(\lim_{n\to\infty} \hat{k}(\underline{X}) = k\right) = 1 \tag{12}$$

In our analysis, we set $K_0 = 6$, $\delta = 2$ and we have $n = 90,000$ samples for each receiver in the network.

The order of the model for each receiver was estimated by successively calculating the parameters of the model for increasing orders. *i.e.*, starting from order 0, the conditional probability matrix was estimated. Table 5 shows the 0th, 1st and 2nd order Markov models for the same trace. To decide which order was appropriate, the entropy of the variable was calculated at each stage.

The parameters of the Markov models of order less than 6 were calculated for all the receivers. The entropies became almost constant after the 3^{rd} order model, in all cases. Machines, `tove` and `ursa` showed a slightly steeper fall in entropy but in even those cases a threshold of $\epsilon_n = 0.016$ was sufficient to get an order estimate of 3. Table 6 shows how the entropies decreased with increasing orders and the estimated order of the source for a threshold of 0.016.

State	Probability of being in the state	Probability of 0	Probability of 1
000	0.8721	0.9779	0.0221
001	0.0208	0.6112	0.3888
010	0.0142	0.8819	0.1181
011	0.0102	0.2710	0.7290
100	0.0208	0.9278	0.0722
101	0.0036	0.4198	0.5802
110	0.0102	0.8109	0.1891
111	0.0481	0.1539	**0.8461**

Table 4: 3rd order Markov Model

Thus it can be concluded from the above analysis, a Markov model of order 3 is adequate to model the data for all receivers.

This result also justifies, in retrospect, our choice of $K_0 = 6$. Had any of the traces exhibited temporal dependencies of orders greater than 6, then, with very high probability, its conditional entropy would have continued to decrease beyond $k = 3$ and $\hat{k}(\cdot)$ would have estimated its order to be the maximum permitted value of K_0.

Looking at the state transition matrix in Table 4 we observed that the probability of a 1 occurring in state 111 is 0.846. This is considerably greater than the probability of observing a 1 in the single-state Markov chain (0.0827). This means that the likelihood of loss is much greater if loss has happened in the recent past. Once the system is in lossy state, it does tend to remain in that state. Additionally, if the lossy state has continued over a couple of time intervals, the probability of it continuing further increases.

6 Conclusions and Future Work

This paper described the results of experiments on packet loss in the MBone, a multicast network in widespread use. Measurements were taken at 11 distinct geographical locations in Europe and the US, for 2 hours during the day and 1 hour during the night.

It was found that almost 70% of the packets were not received by one or more sites. The "backbone" loss between two main routers was found to be significantly smaller than that near the "edges" of the multicast network. A negligible number of packets were lost at the receiving hosts themselves.

Markov Model of order 0, entropy = 0.285301

State	Probability of being in the state	Probability of 0	Probability of 1
		0.9173	**0.0827**

Markov Model of order 1, entropy = 0.162640

State	Probability of being in the state	Probability of 0	Probability of 1
0	0.9173	0.9734	0.0266
1	0.0827	0.2948	0.7052

Markov Model of order 2, entropy = 0.150984

State	Probability of being in the state	Probability of 0	Probability of 1
00	0.8929	0.9767	0.0233
01	0.0244	0.5830	0.4170
10	0.0244	0.8523	0.1477
11	0.0583	0.1744	0.8256

Table 5: Markov Models of order 0, 1 and 2

Workstation Name	1-st order entropy	2-nd order entropy	3-rd order entropy	6-th order entropy	Estimated order (threshold=0.016)
alps	0.16264	0.15243	0.15098	0.14957	1
anhur	0.10579	0.10406	0.10365	0.10326	1
cedar	0.11549	0.10994	0.10928	0.10821	1
collage	0.05875	0.04682	0.04521	0.04300	2
erlang	0.19229	0.16685	0.16087	0.15550	2
float	0.18194	0.16837	0.16610	0.16406	2
law	0.13726	0.12865	0.12780	0.12659	1
tove	0.50142	0.45436	0.43580	0.42512	3
ursa	0.46624	0.39565	0.37682	0.36810	3
willow	0.25206	0.23241	0.23108	0.22975	2
zen	0.30919	0.28370	0.28032	0.27883	2

Table 6: Entropies and the Estimated Order

Looking at the number of consecutive losses seen by a receiver, a large number of single losses were observed in the traces. Also, a few extremely long bursts (greater than 100 consecutive packets) were seen. These long lossy bursts contribute heavily to the total loss, despite their small numbers (22% in one trace).

The packet loss at a receiver was modeled as a discrete-time, binary-valued Markov chain. A method of estimating the state transition matrix for Markov chains of a given order was discussed. Using "entropy" as a measure of uncertainty in predicting whether the next packet will be lost, given the history of the last n packets, we found that an eight-state Markov chain ($n = 3$) is adequate to model the temporal correlation in loss. The state transition matrix showed that the probability of a solid block of loss continuing is high (0.8) as compared to the basic probability of loss occurring anywhere in the trace (0.08).

Further analyses of the collected data would give better insight into the character of MBone packet loss. In particular, it would be interesting to estimate the burstiness of loss along each link of the transmission tree making it possible to see how the temporal correlations of loss vary in different parts of the multicast network. Combining the loss information for all the receivers, one could construct a vector-valued, discrete-time Markov chain model which would characterize both the spatial and the temporal correlations of the packet loss.

7 Acknowledgements

We would like to thank Sanjeev Khudanpur for his help with the Markov chain model and for various timely pointers and insights, Henning Schulzrinne and Vinay Kumar for their immense help in conducting the experiments, and our colleague Sue Bok Moon for her support and encouragement.

We are grateful to Phillipe Nain of INRIA, Chuck Cranor of Washington Univ., St.Louis, Ron Erlich of Columbia Univ., Deborah Estrin of Univ. of Southern California , Edward Knightly of Univ. of California at Berkeley, Simon Lam of Univ. of Texas, Jorg Liebeherr of University of Virginia at Charlottesville, Stephen Pink of SICS (Swedish Institute of Computer Science), Sweden, Larry Peterson of Univ. of Arizona, Kishor Trivedi of Duke Univ., Tatsuya Suda of Univ. of California at Irwine, Jon Reid of University of Illinois at Urbana-Champaign, Satish Tripathi of Univ. of Maryland at College Park, Peter Wan of Georgia Institute of Technology and Raj Yavatkar of Univ. of Kentucky at Lexington for providing MBone-capable computer accounts which allowed us to take the measurements.

References

[1] P. Bhagwat, P. Misra, S. Tripathi, "Effect of Topology on Performance of Reliable Multicast Communication," *Proc. IEEE Infocom 94,* (Toronto, June 1994), pp. 602 – 609.

[2] J.C. Bolot, "End-to-End Packet Delay and Loss Behavior in the Internet," *Proc. 1993 ACM SIGCOMM Conf.,* (Sept. 1993, San Francisco), pp. 289-298.

[3] J. Bolot, T. Turletti, "A Rate Control Scheme for packet video in the Internet," *Proc. IEEE Infocom94,* pp. 1216 - 1223, June 1994.

[4] J. Bolot, H. Crepin, and A Vega Garcia, "Analysis of Audio Packet Loss in the Internet," *Proc. 1995 Workshop on Network and Operaing System Support for Audio and Video,* pp. 163 - 174.

[5] R. Braudes, S. Zabele, "Requirements for Multicast Protocols," RFC 1458, May 1993.

[6] S. Casner and S. Deering, "First IETF Internet Audiocast," *ACM Computer Communication Review,* Vol. 22, No. 3 (July 1992), pp. 92 - 97.

[7] S.E. Deering and D.R. Cheriton, "Multicast Routing in Datagram Internetworks and Extended LANs", *ACM Trans. on Computer Systems*, 8:85–110, May 1990.

[8] L. Finesso, C-C Liu and P. Narayan "The Optimal Error Exponent for Markov Order Estimation" , submitted for publication to *IEEE Trans. on Information Theory*

[9] R. Frederick, "nv", Manual Pages, Xerox Palo Alto Research Center.

[10] V. Jacobson and S. McCanne, "vat", Manual Pages, Lawrence Berkeley Laboratory, Berkeley, CA.

[11] V. Jacobson and S. McCanne, "Using the LBL Network 'Whiteboard", Lawrence Berkeley Laboratory, Berkeley, CA.

[12] V. Jacobsen, "Multimedia Conferencing on the Internet," *Tutorial Notes - ACM Sigcomm94,* (London, Sept. 1994).

[13] S. Karlin, "A First Course in Stochastic Processes"

[14] B. Kedem, "Binary Time Series"

[15] M. Macedonia and D. Brutzman, "MBone Provides Audio and Video Across the Internet," *IEEE Computer Magazine,* April 1994, pp. 30 -35.

[16] S. McCanne and V. Jacobsen, "VIC: Video Conference," UC Berkeley and Lawrence Berkeley Lab, Software available via ftp://ftp.ee.lbl.gov/conferencing/vic.

[17] N. Merhav, M. Gutman and J.Ziv, "On the Estimation of the Order of a Markov Chain and Universal data Compression", *IEEE Trans. on Information Technology*, vol 35, No 5, Sept 1989

[18] S. Paul, K. Sabnani, D. Kristol, "Multicast Transport Protocols for High-Speed Networks," *Proc. 1994 IEEE Int. Conf. Network Protocols*, (Boston, Oct. 1994).

[19] S. Pingali, D. Towsley, J. Kurose, "A Comparison of Sender-Initiated and Receiver-Initiated Reliable Multicast Protocols", *Proc. 1994 ACM SIGMETRICS Conf.*, 221 – 230, May 1994.

[20] S. Ramakrishnan and B. N. Jain, "A Negative Acknowledgement with Periodic Polling Protocol for Multicast over LANs," *Proc. IEEE Infocom'87*, pp 502–511, Mar-Apr 1987.

[21] R. Ramjee, J. Kurose, D. Towsley, "Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks," *Proc. INFOCOM'94*.

[22] H. Schulzrinne, "Voice Communication Across the Internet: a Network Voice Terminal," Technical Report, Dept. of Computer Science, U. Massachusetts, Amherst MA, July 1992. (available via anonymous ftp to gaia.cs.umass.edu in pub/nevot/nevot.ps.Z)

[23] R. Yavatkar, L. Manor, "End-to-End Approach to Large Scale Multimedia Dissemination," *Computer Communications*, Vol. 17, No. 3 (March 1994), pp. 205 – 218.

[24] R. Yavatkar and L. Manoj, "Optimistic Approaches to Large-Scale Dissemination of Multimedia Information", *Proc. ACM Multimedia '93*, August 1993.