December 1992

On the Dynamics and Significance of Low Frequency Components of Internet Load

Amarnath Mukherjee
University of Pennsylvania

University of Pennsylvania Department of Computer and Information Sciences Technical Report No. MS-CIS-92-83.

On the Dynamics and Significance of Low Frequency Components of Internet Load

Abstract

Dynamics of Internet load are investigated using statistics of round-trip delays, packet losses and out-of-order sequence of acknowledgments. Several segments of the Internet are studied. They include a regional network (the Jon von Neumann Center Network), a segment of the NSFNet backbone and a cross-country network consisting of regional and backbone segments.

Issues addressed include:

(a) dominant time scales in network workload;

(b) the relationship between packet loss and different statistics of round-trip delay (average, minimum, maximum and standard-deviation);

(c) the relationship between out of sequence acknowledgments and different statistics of delay;

(d) the distribution of delay;

(e) a comparison of results across different network segments (regional, backbone and cross-country); and

(f) a comparison of results across time for a specific network segment.

This study attempts to characterize the dynamics of Internet workload from an end-point perspective. A key conclusion from the data is that efficient congestion control is still a very difficult problem in large internetworks. Nevertheless, there are interesting signals of congestion that may be inferred from the data. Examples include (a) presence of slow oscillation components in smoothed network delay, (b) increase in conditional expected loss and conditional out-of-sequence acknowledgments as a function of various statistics of delay, (c) change in delay distribution parameters as a function of load, while the distribution itself remains the same, etc. The results have potential application in heuristic algorithms and analytical approximations for congestion control.

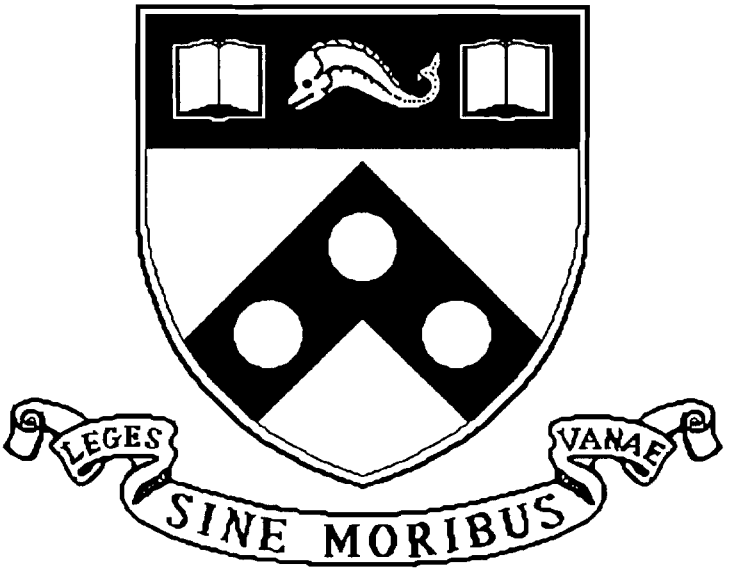
Comments

University of Pennsylvania Department of Computer and Information Sciences Technical Report No. MS-CIS-92-83.

On The Dynamics and Significance of Low Frequency Components of Internet Load

MS-CIS-92-83
DISTRIBUTED SYSTEMS LAB 12

Amarnath Mukherjee



University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

December 1992

Submitted to THE JOURNAL OF INTERNETWORKING: Research and Experience

On the Dynamics and Significance of Low Frequency Components of Internet Load *

Amarnath Mukherjee
Computer and Information Science Department
University of Pennsylvania
Philadelphia, PA 19104

December 7, 1992

(TECHNICAL REPORT NUMBER: MS-CIS-92-83/DSL-12)

Abstract

Dynamics of Internet load are investigated using statistics of round-trip delays, packet losses and out-of-order sequence of acknowledgments. Several segments of the Internet are studied. They include a regional network (the Jon von Neumann Center Network), a segment of the NSFNet backbone and a cross-country network consisting of regional and backbone segments.

Issues addressed include (a) dominant time scales in network workload; (b) the relationship between packet loss and different statistics of round-trip delay (average, minimum, maximum and standard-deviation); (c) the relationship between out of sequence acknowledgments and different statistics of delay; (d) the distribution of delay; (e) a comparison of results across different network segments (regional, backbone and cross-country); and (f) a comparison of results across time for a specific network segment.

This study attempts to characterize the dynamics of Internet workload from an end-point perspective. A key conclusion from the data is that efficient congestion control is still a very difficult problem in large internetworks. Nevertheless, there are interesting signals of congestion that may be inferred from the data. Examples include (a) presence of slow oscillation components in smoothed network delay, (b) increase in conditional expected loss and conditional out-of-sequence acknowledgments as a function of various statistics of delay, (c) change in delay distribution parameters as a function of load, while the distribution itself remains the same, etc. The results have potential application in heuristic algorithms and analytical approximations for congestion control.

Keywords: Round Trip Delay, Workload Characterization, Congestion Control.

*This work was supported in part by the National Science Foundation under Grant 5-24141 and by funds provided by David Farber from a grant from the Corporation for National Research Initiatives.

1 Introduction

Properties of non-stationary offered load on several segments of the Internet are investigated using statistics of end-to-end round-trip delays. The goals of this study were to (a) learn more about offered load and loss characteristics from an end-point perspective, (b) compare and contrast results from different segments of the Internet in order to learn more about the effects of different number of hops, different link speeds and different offered loads, and (c) develop model parameters for round-trip delays in analytical studies. Table 1 summarizes the key results.

The current Internet consists of a large number of hosts, interconnected with each other through a myriad of switches, links (point-to-point, multiple-access) and protocols. These components have a wide range of speeds and buffering capabilities. The need for communication across different administrative domains and the evolution of technology has required the need for inter-operability and this constitutes the large heterogeneous Internet.

Individual host-to-host paths[1] vary in this network, across time and geographical distance. Over local and regional networks, the number of hops is small, but across regional networks (connected through the NSFNet backbone), the number of hops can be large. Tables 2-4 show a few example paths over the Internet. (Since paths in the Internet are not fixed, these are only snapshots in time.)

The environment is too complex and large to admit standard analysis techniques (e.g., Product Form Queueing Networks, Markov Processes or Simulation). Further, for such studies, one needs to know the offered load which is a (not-so-well-known) time varying function. Several studies, for example [3, 10, 31, 36], have, therefore, focussed on workload characterization from measurements at a gateway. Our attempt here is to study the characteristics of end-to-end delay and loss properties over multiple hops from some simple, minimally disruptive measurements. (A more elaborate study involving measurement hooks at each and every point in a path could improve our understanding further, but that is beyond our current resources.) The results have potential application in heuristic algorithms and analytical approximations for congestion control.

At the time we were collecting the data, Sanghi et al [35] were performing a similar study at the University of Maryland. While we were not aware of their study at the time, our experiments are quite similar, although they differ in important details (see Section 2.2). This paper will complement that study without duplicating their results.

It ought to be emphasized that network components are constantly evolving, and with multi-media transport in the horizon, so will the workload. Exact results pertaining to a specific path, therefore, face quick obsolescence. Hopefully, a few general principles presented here will be long-lived; perhaps the most robust of these is the correlation between loss and delay, studied in Section 5.3. See also the distribution of delay, studied in Section 4.

The study was motivated by the following:

1. *Need for round trip delay distribution parameters in analytical studies:* Round-trip delays have been known to cause oscillations in adaptive congestion control studies [2, 11, 16, 30, 40]. In analytical studies, these delays have been assumed deterministic. We were interested in the distribution of delay in order to extend upon the analytical studies, with the hope of deriving better control algorithms.
2. *Potential use of round trip delays for heuristics mechanisms in congestion control:*

[1] We refer to the set of links and switches traversed by a packet as a path. This is not to be confused with a virtual circuit.

1. A few frequencies in the data dominate over the rest.
2. The distribution of delay is approximately a shifted Gamma for all three network paths studied. Its tail is large even for networks with low congestion levels.
3. The shape and scale parameters of the empirical Gamma distribution vary with load and network segment. For instance, for the backbone, the shape parameter ranged from 1.0 during low loads to 6.0 during high loads, although it could go up to 20 for small intervals. For the cross country segment, it ranged from 1.0 to 8.0. For the regional, it was 0.05 to 0.20.
4. There exists a base congestion level that moves slowly with time. The conditional expected packet loss given a statistic of delay (mean, standard deviation, minimum, or maximum of a group of packets) increases with the delay, even when the delay statistics are computed over a packet group where successive packets that are spaced a full one second apart.
5. The correlation between packet loss and delay varies considerably across networks. Over the same network path, it varies with load. Low loads show lower correlations between delay and loss, because losses are rare. Higher loads show higher correlation.
6. The number of out-of-sequence acknowledgments is positively correlated with the different delay statistics for the cross-country segment. Analogous to the packet-loss case, the conditional expectation of fraction of out-of-sequence acknowledgments also increases as a function of delay. This information could potentially be useful for heuristic inference of network state. However, it also implies that it is harder to do bottleneck flow control.

Table 1: Summary of Key Results.

The study in [17] had proposed a delay based congestion control algorithm and had given closed form results for a network with deterministic delays. The principal idea in that study was to adjust flow control parameters based on observed delays. In a deterministic environment, this involved checking if the delay was above its minimum possible value and take appropriate action. While network delays are known to be stochastic, its properties are not well understood, and we were interested in investigating its properties.

Several other recent congestion control studies [9, 14, 15, 27, 32] attempted to classify switch congestion by time scales over which they occur, and proposed mechanisms that address each time scale. For example, in [15], Hui considered packet level, burst level and call level as three time scales with the idea that for high bandwidth-delay-product networks, one could use open loop control for shorter time scales and closed loop control for longer time scales. We had similarly assumed the presence of short term and medium term congestion in order to distinguish between the dynamics that occur within a round-trip time and those that occur over longer periods [9, 27].

We were interested in determining the impact of slower time scales experimentally. Specifically, we were interested in studying the correlations between packet loss and low frequency components of delay.

3. *Need for realistic assumptions in simulations:* Simulation studies (necessarily) involve small network topologies with a limited number of conversations. These studies cannot capture some of the important properties of offered load in long haul networks. We felt it was necessary to determine which assumptions were reasonable and which were not.

Corresponding to the issues above, we learn from this study that:

1. The distribution of delay is a constant plus a Gamma distributed random variable. Its shape and scale parameters change with load.
2. Packet losses and different statistics of delay (average, standard deviation, minimum or maximum) are positively correlated, even when the delay statistic is computed over packets that are spaced a full one second apart (see Section 2.2 for details). The magnitude of correlation depends on the load and the network segment. It could be between 0.4 to 0.7 for the backbone and cross-country segments, or as low as 0.06 for the regional segment. Thus, while delay based congestion avoidance is possible, it could lead to loss of bandwidth.

   This study also finds that a positive correlation exists between out-of-sequence acknowledgments and different delay statistics for the cross-country segment.
3. Smoothed time-series plots of delay indicate the existence of significant low frequency components in the workload. This suggests that there exists a base congestion level that moves slowly with time.
4. The data also suggests the need for improved simulation workloads. For instance, realistic studies of flow control protocols should allow for inclusion of conversation arrival and departure rates, topologies that admit a larger number of hops, a richly multiplexed traffic stream, and dynamic routing if the target network allows for that possibility.

The rest of the paper is organized as follows. Section 2 presents the network model, the measurement procedure adopted, and a preliminary look at end-to-end delay, packet

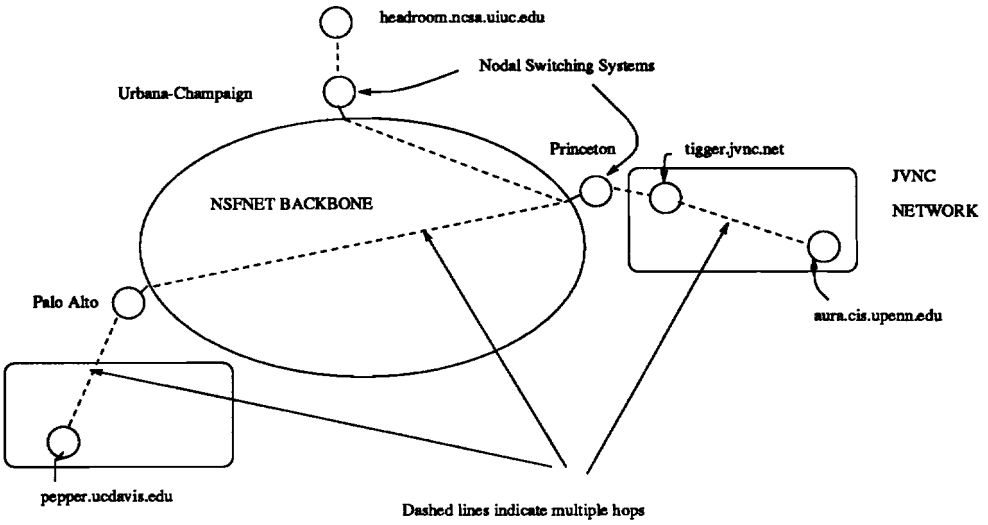


Figure 1: A Sketch of the Internet Showing the Paths Studied.

loss and out-of-order acknowledgment statistics. Section 3 investigates the dominant low frequency components in round-trip delay. Section 4 investigates the distribution of delay. Section 5 presents the properties of loss as a function of delay. In particular, it investigates the conditional expected loss as a function of delay, the correlation between different statistics of delay and loss, and the implications of a given value of correlation in terms of the standard deviations of delay and loss. Section 6 investigates the relationship between out-of-sequence acknowledgments and delay. Section 7 compares the results for a specific network segment over a period of one and a half years. Section 8 discusses related work and Section 9 presents our concluding remarks and lists what we think are important open questions. Finally, Appendices A, B and C give some of the derivations that are mentioned in the text of the paper.

2 Preliminaries

2.1 Network Model

The Internet consists of a large number of machines and is growing with time. Exploration of all possible paths is, therefore, impractical. We studied a total of seven paths, but, in here, we shall focus on only three of them because they provide sufficient similarities and contrast to summarize the results.

The three paths are:

(a) a regional network segment (see Table 2);

(b) a backbone network segment[2] (see Table 3); and

(c) a cross-country network segment consisting of a combination of regional and backbone segments (see Table 4).

A schematic diagram of the Internet and the paths studied is shown in Figure 1.

[2]Due to logistical reasons, the machines from which we collected the data were not the backbone themselves, but fairly close to them.

0	aura.cis.upenn.edu (130.91.6.203)
1	GW.CIS.UPENN.EDU (130.91.6.254)
2	EXTERNAL-DEFAULT.UPENN.EDU (130.91.244.1)
3	JVNCNET-GW.UPENN.EDU (192.84.2.2)
4	cheesesteak1-gateway.jvnc.net (130.94.35.9)
5	ford-gateway.jvnc.net (130.94.34.249)
6	tigger.jvnc.net (128.121.50.145)

Table 2: A sample regional path (over the Jon von Neumann Center Network) from aura.cis.upenn.edu to tigger.jvnc.net. Path traced by traceroute. (Some diagnostic messages deleted for clarity; also the zero-th node, aura.cis.upenn.edu, has been added manually.)

0	tigger.jvnc.net (128.121.50.145)
1	ford-gateway.jvnc.net (130.94.34.249)
2	nss.jvnc.net (192.12.211.1)
3	Pittsburgh.PA.NSS.NSF.NET (129.140.69.8)
4	Urbana_Champaign.IL.NSS.NSF.NET (129.140.76.5)
5	Urbana_Champaign.IL.NSS.NSF.NET (129.140.12.10)
6	linus.ncsa.uiuc.edu (141.142.1.155)
7	headroom.ncsa.uiuc.edu (141.142.22.61)

Table 3: A sample path over the backbone. Measurements were taken between tigger.jvnc.net (near the Princeton Nodal Switching System) and headroom.ncsa.uiuc.edu (near the Urbana-Champaign Nodal Switching System) which were close to the backbone.

0	aura.cis.upenn.edu (130.91.6.203)
1	GW.CIS.UPENN.EDU (130.91.6.254)
2	EXTERNAL-DEFAULT.UPENN.EDU (130.91.244.1)
3	JVNCNET-GW.UPENN.EDU (192.84.2.2)
4	cheesesteak1-gateway.jvnc.net (130.94.35.9)
5	ford-gateway.jvnc.net (130.94.34.249)
6	nss.jvnc.net (192.12.211.1)
7	Ann_Arbor.MI.NSS.NSF.NET (129.140.81.8)
8	Salt_Lake_City.UT.NSS.NSF.NET (129.140.79.17)
9	Palo_Alto.CA.NSS.NSF.NET (129.140.77.15)
10	Palo_Alto.CA.NSS.NSF.NET (129.140.13.10)
11	SU-C3.BARRNET.NET (131.119.252.103)
12	UCD.BARRNET.NET (131.119.9.2)
13	telco-ext.ucdavis.edu (128.120.250.254)
14	bainerfddi-gw.ucdavis.edu (128.120.128.2)
15	* * *
16	* pepper.cs.ucdavis.edu (128.120.56.123)

Table 4: Cross-country path from aura.cis.upenn.edu to pepper.cs.ucdavis.edu, as traced by traceroute on November 9, 1992. A * corresponds to a missed attempt by the traceroute program. Hop number 15 was not identified. Paths traced during June-July 1991 had more than 23 hops. Considerable improvement in terms of number of hops has occurred over this period.

The distinguishing characteristics of these segments are as follows. The regional segment consists of a small number of hops, a small propagation delay, and for the network studied, also has a low load. The backbone path would ideally consist of only nodes in the backbone. For logistical reasons, however, the machines from which the data was collected were close to the backbone, but not on the backbone themselves (see Table 3). We do not, however, expect this to make a significant difference to the conclusions. (In fact, as it turned out, the important conclusions for the backbone and the cross-country segments were found to be the same.) The backbone provides the characteristics of a highly multiplexed traffic stream, the implications of which will be seen in Section 2.3. Its number of hops is expected to be moderate while its propagation delay depends on the choice of end-points.

The cross-country segment, by contrast, has a larger number of hops and for the choice of end-points, has a larger propagation delay. Its workload is a mix of regional and backbone workloads.

2.2 Measurement Procedure

Round-trip delays were measured using the standard Internet Control Message Protocol (ICMP): a packet was sent from a source host, A, to a destination host, B, which responded with an echo packet back to A; the time difference between sending and receiving at A was recorded as an instance of the round-trip delay between A and B.

The goals of this study were to determine the dominant network time scales in wide-area networks, the distributional properties of delay and the packet loss properties of a path as a function of delay. In simplistic terms, this may be stated as follows:

> "What are the signals of congestion, how strong are they, and are there low frequency components that build up and clear out slowly, so that feedback based congestion control schemes could work?"

We first describe the measurement procedure that we adopted and the rationale behind it. Then we shall describe an alternative strategy that was adopted by Sanghi et al [35].

Several issues needed to be addressed in the measurement procedure:

- The injected packets should not cause self-interference, thereby increasing the measured delays. This means that echo packets should not increase the queueing delays of each other, because that would defeat the goal of the study;
- The measurement procedure should induce only a low overhead on regular network traffic. This was of importance because we wished to collect data for days (often a week) at a time;
- There ought to be a way to correlate packet loss statistics with observed delays. One possibility was to transmit a group of packets and correlate the fraction of lost packets in the packet group with the delays of those that were successful. This would work as long as the entire packet group was not lost. In practice, we found it to worked quite well. For instance, during an entire day's worth of measurements over the cross-country segment, only 5 out of 1440 packet groups were lost in their entirety;
- There ought to be ways to determine which statistics of delay were more important indicators of congestion than others. For example, before the experiments were

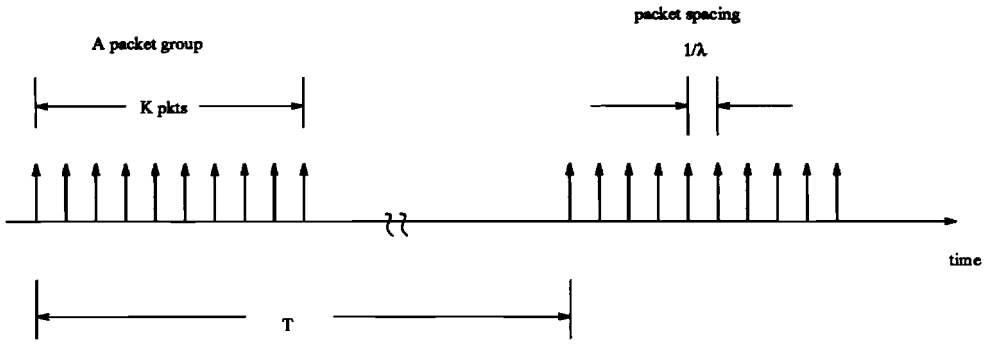


Figure 2: Measurement Procedure.

conducted, we were under the impression that the standard deviation of delay would be a significant indicator of congestion. However, as it turned out, the minimum, the maximum and the average delays usually show a higher positive correlation with packet loss than standard deviation. However, no single statistic consistently shows a higher correlation with loss.

Given the above issues, it was decided to transmit a group of K echo packets every T time units apart. The packets in a group were, themselves, separated by $1/\lambda$ time units. The strategy is schematically shown in Figure 2. For the experiments, the parameter values chosen were: $K = 10$, $T \in \{1, 2\}$ minutes and $\lambda \in \{1, 60\}$ Hertz.

The tradeoffs in the parameter values are obvious:

- T: a smaller T would capture higher frequencies; however, this would come at the expense of overloading the network;
- λ: a larger λ would capture more out-of-sequence acknowledgments, and could show how packet losses increase with λ. It could also give a more accurate delay statistic for a given time point. However, all of this would come at the expense of overloading the network, and introducing self-interference among echo packets as indicated earlier.
- K: a larger K could result in a more accurate delay statistic for a particular time point by increasing the sample size; conversely, it could also decrease the accuracy of delay statistics for a given time point by increasing the time interval over which the measurements are taken. (As an extreme example, consider K to be so large that it takes a whole day to transmit a packet group; this could hardly give an accurate indication of load at 8am.) For a fixed T and λ, increasing K also increases the load on the network.

The measurement procedure adopted by Sanghi et al [35] differ from the one described above in several ways. Their experiment consisted of transmitting a fixed (large) number of packets which were transmitted as fast as the sending machines would allow. In terms of our model, their λ and K were large, and there was no T. Their method has the advantages and disadvantages of a larger λ indicated earlier. The tradeoff between K and the accuracy of delay statistics alluded to before, does not apply to their experiments, because they were addressing individual packet delays rather than packet-group statistics. One still needs a way of associating a delay with a packet loss, because a packet loss, by definition, implies an infinite delay. One possibility is to determine the

relationship between the delay of an immediately preceding successful packet and the packet loss, and that is the avenue they adopted.

Sampling network dynamics with round-trip delays is not without limitations. This is because of Nyquist's sampling theorem which states that the sampling frequency must be at least as large as twice the largest frequency in network dynamics in order to construct its true nature. Sampling at a slower rate, however, could potentially result in aliasing where high frequencies in the original function could appear as low frequencies in the sampled data.

However, measuring at the Nyquist frequency using echo packets would require injecting packets at twice the rate of the fastest change in network dynamics, and this is impractical, because:

(a) injecting packets at that rate is non-trivial;

(b) it would cause overloading and disrupt normal network operations; and

(c) the echo packets will see additional queueing delays due to each other, and these could be larger than the delays they were supposed to measure.

The implication of this is that path-delay measurements using echo packets are not really suitable for measuring the fastest components of network dynamics.

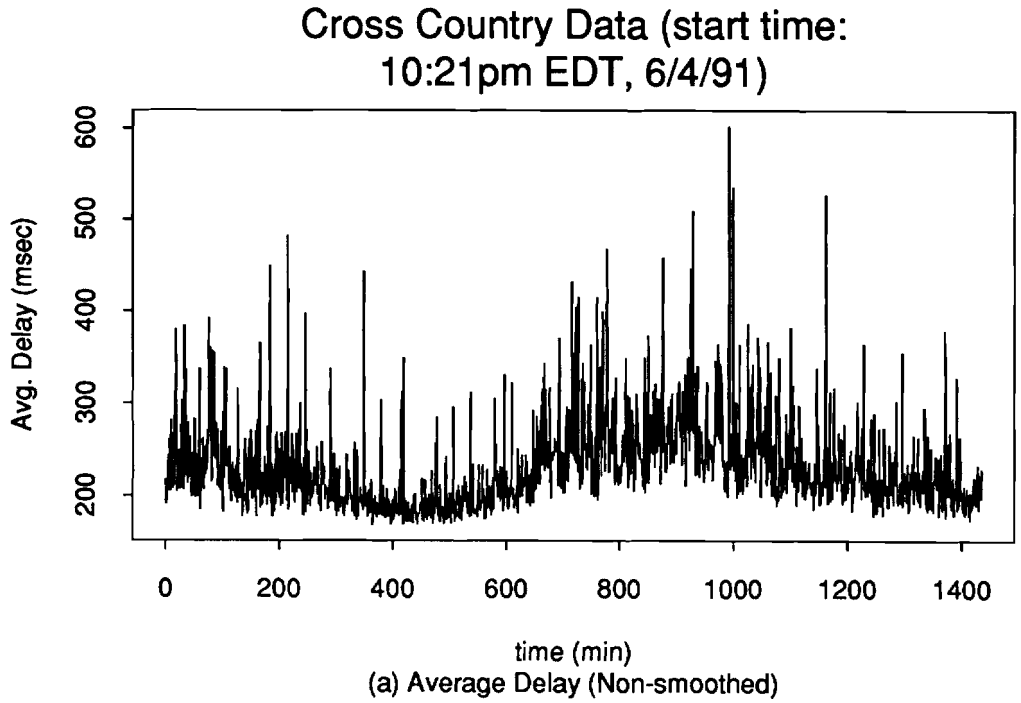
In spite of the possibility of aliasing, however, we shall give in to the temptation of determining dominant low frequency components in the data (Section 3). There are two reasons for this:

- The data shows that the dominant frequencies are the extremely low frequencies. For example, the most dominant frequency in a day is the day itself. Now, for aliasing to occur, the high frequencies that are missed by the Nyquist frequency violation, must be present for the entire length of time of a corresponding low frequency, so as to be sampled as the low frequency. From a practical standpoint, that is extremely improbable, given the dominant low frequencies that were found in the data (see Section 3). Nevertheless, we shall interpret the results of Section 3 with a healthy dose of skepticism, and we ask that the reader does the same. The results in Sections 4-6 are, however, not affected by the Nyquist frequency violation.
- There does not appear to be any simple alternative method for determining the important time scales for a wide-area network path. One possibility is to record each and every event at critical queueing points along a path, and research into ways for combining them. The non-triviality of this approach stems from a possible correlation between queues, and more importantly, the logistics of collecting data simultaneously across several administrative domains.

2.3 A Preliminary Look at Data

Figure 3(a) shows the (raw) estimated average delays as a function of time of day for the cross-country segment. To get a visual feel for patterns in the data, perhaps it is more useful to look at its smoothed version shown in Figure 3(b).

There are several ways to smooth the data. The method we adopted was to take a template of size M, slide it along the figure, and for every point i, compute the average of the values within the bounds of the template. That is, if $d(i)$ are the raw average



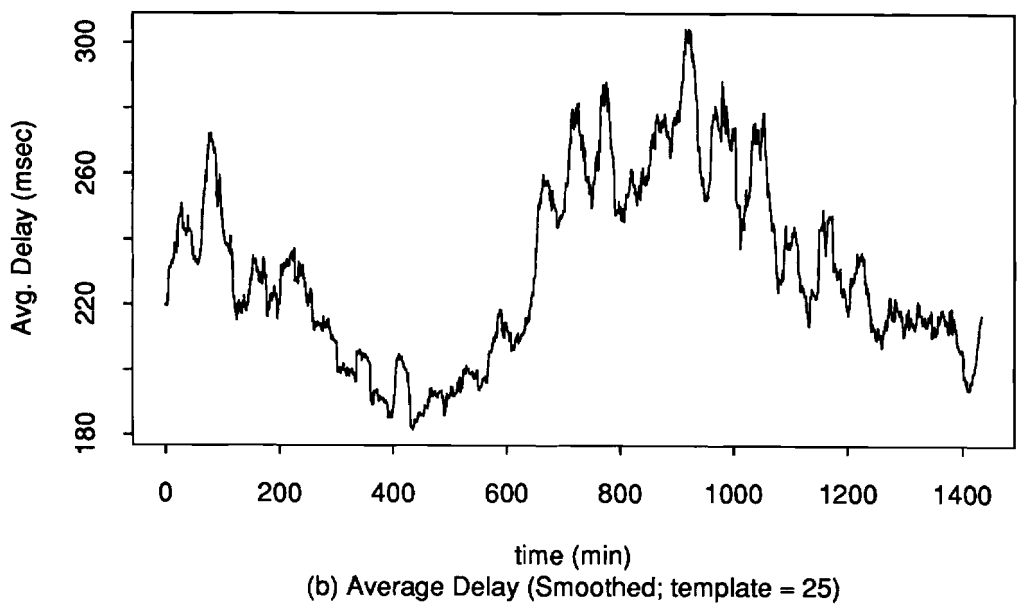


Figure 3: Cross Country Average Delays Over One Working Day.

delays, then the corresponding smoothed values are simply

$$d_{smooth}(i) = \sum_{j=i-\lfloor M/2 \rfloor}^{i+\lfloor M/2 \rfloor} d(j). \quad (1)$$

It is convenient to give an equal weight to the left and right halves of $d(i)$, so M was chosen odd.

This algorithm brings out the low frequency components in the data and has been effectively used in image processing applications. It works well in practice, except at the end-points where the template goes out of the figure. For such cases, we simply extended the values at the left and right end-points. While not a completely satisfactory solution, its consequences are minimal given that it affects only a few peripheral points.

An important parameter of the smoothing algorithm is the template size M. A large M achieves more smoothing but also looses more information. The figures shown in this section are based on judgment and experimentation rather than a concrete algorithm. The reader is assured that smoothing was done only for getting a visual intuition of network dynamics. All analyses that follow were performed on raw data and not on their smoothed versions.

In the rest of this section, we take a preliminary look at the dynamics of delays over several paths. Figure 4(a) shows the average delays over the backbone segment and Figure 4(b) shows the same over the regional segment. Compare these with that of the cross-country segment shown in Figure 3(b). Notice that the patterns are qualitatively different. Over the cross-country segment, the average delay has a marked work-day frequency, interspersed with several other dominant frequencies. Over the backbone, the work-day frequency is more pronounced but other frequencies are relatively less pronounced. We believe that this is due to a high degree of multiplexing over the backbone which together with the law of large numbers, would predict a sharper demarcation of when the network is in heavy use from when it is not. Over the regional segment, the pattern appears to imply that the load is low compared to the network capacity.

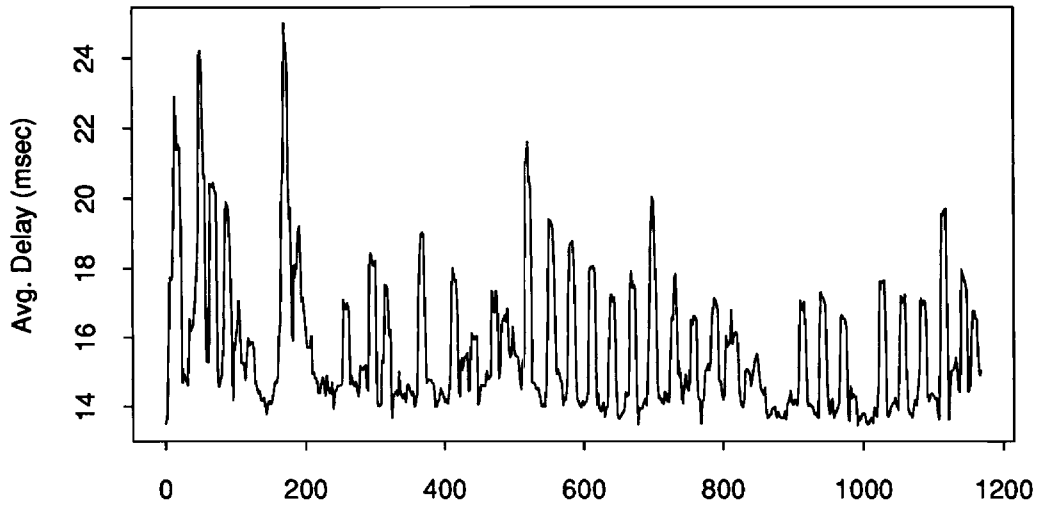
Next, Figure 5 compares the smoothed versions of average, standard deviation, minimum and maximum delays for a specific segment (the cross-country segment). Notice how they are quite similar. Plots for other segments show that the patterns are visually similar for the different delay statistics, as long as the load is not too low; for lightly loaded segments, however, the minimum delay curves are relatively flat. (The figures have not been included here to save space.)

Finally, Figure 6 shows the number of packets lost and the number of out-of-sequence packets for the cross-country segment as a function of time. This data is not smoothed because smoothing was not found to add any intuition. These figures and their relationship with different delay statistics will be studied further in Sections 5 and 6.

3 Dominant Frequencies

In this section, we investigate the dominant network frequencies over a working day. From the data on average delays, it is clear that there are several dominant low frequency components. For instance, the strong diurnal pattern indicates a work day (see Figure 3(b)). There are also several smaller frequencies in the data. We investigate their relative strengths in this section.

(a) Backbone Data (start time: 10:45am EDT, 7/23/91)



(b) Regional Data (start time: 12:42pm EDT, 8/5/91)

Figure 4: Smoothed Average Delays for Backbone and Regional Segments. Notice how the patterns are different. Also compare these with Figure 3(b).

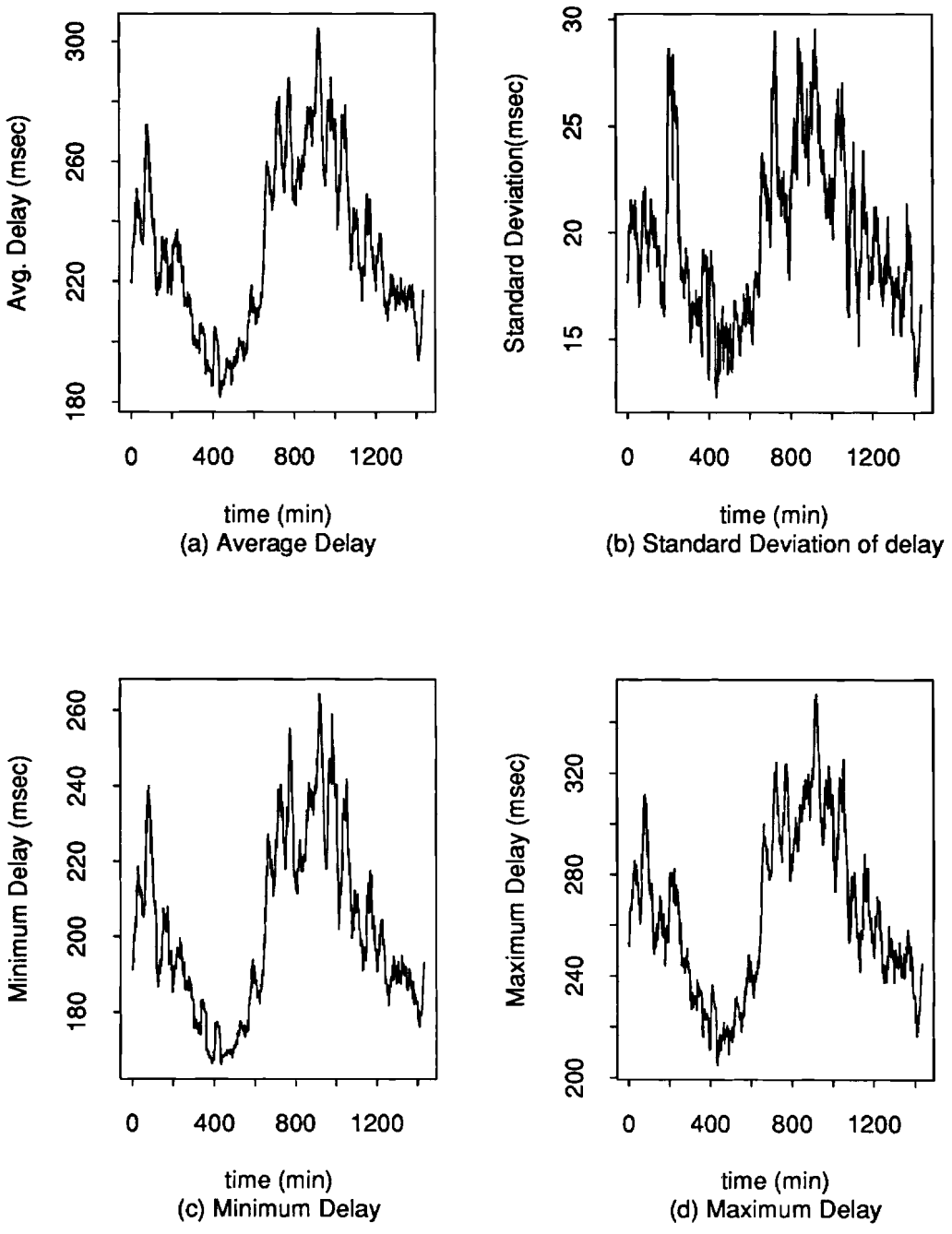


Figure 5: Time Series Plots of Smoothed Delay Statistics for the Cross-Country Segment. Notice how the patterns are similar.

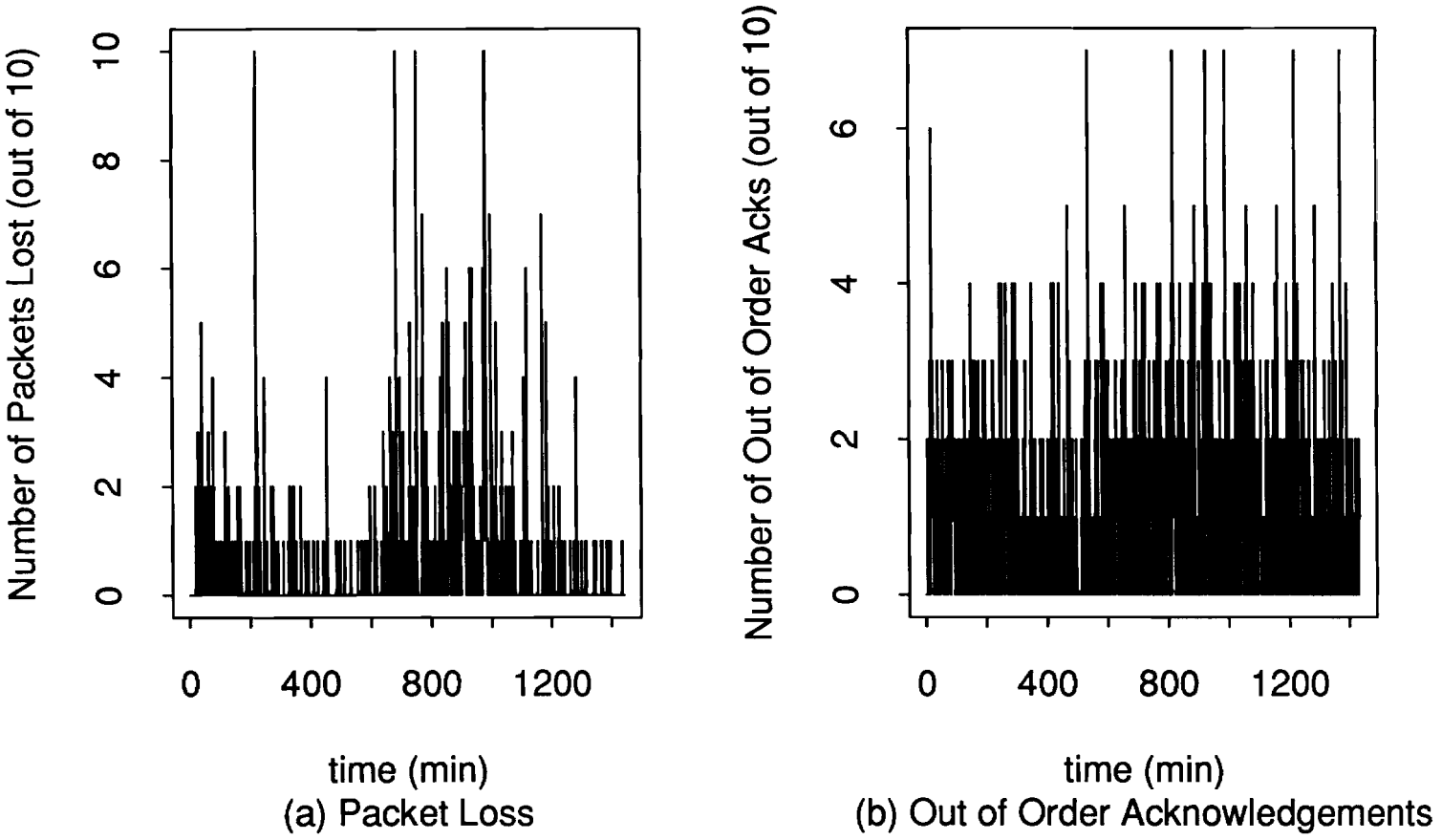


Figure 6: Time Series Plots of Packet Loss and Out-of-Order Acknowledgments for the Cross-Country Segment. The relationship between these and different delay statistics will be studied in Sections 5 and 6.

Dominant frequencies may be identified using a discrete Fourier transform. Suppose $d(i)$ is the average delay at time i, $i = 0, 1, 2, \cdots, N-1$. Then the discrete Fourier transform of $d(i)$ is defined as:

$$D(k) = T \sum_{i=0}^{N-1} d(i)\, e^{-\frac{j2\pi ik}{N}}, \tag{2}$$

where T is the time difference between two sampling points. In our experiments T was either 1 or 2 minutes.

Now, suppose f_k, defined as k/NT, is the kth frequency in the data. (The corresponding time period is NT/k.) Then the discrete Fourier transform of f_k is defined as

$$D(f_k) = T \sum_{i=0}^{N-1} d(i)\, e^{-j2\pi f_k iT}. \tag{3}$$

It can be shown that $D(k)$ and $D(f_k)$ are identical:

$$\begin{aligned} D(f_k) &= T \sum_{i=0}^{N-1} d(i)\, e^{-j2\pi f_k iT} && (4)\\ &= T \sum_{i=0}^{N-1} d(i)\, e^{-j2\pi (\frac{k}{NT})iT} && (5)\\ &= T \sum_{i=0}^{N-1} d(i)\, e^{-\frac{j2\pi ik}{N}}. && (6) \end{aligned}$$

$D(f_k)$ maps f_k to the set of complex numbers. Its magnitude, or gain, is defined as:

$$|D(f_k)| = \left\{ (\mathrm{Re}D(f_k))^2 + (\mathrm{Im}D(f_k))^2 \right\}^{1/2}. \tag{7}$$

Using (7), one may determine the magnitudes of different frequencies in the data. One need not use (3) directly to generate the Fourier Transform $D(f_k)$; there are well known fast fourier transform algorithms for this, and that is what we used. In the figures that follow, we shall plot k on the x-axis instead of f_k, for convenience.

Figure 7 shows a scatter plot of gain versus frequency for the cross-country data. For this curve, T was one minute and N was 1440 minutes (1 day). Therefore, the lowest frequency in the figure corresponds to a day and the highest frequency corresponds to a minute. For any point k on the x-axis, the corresponding frequency is k/NT (or equivalently, the time period is $1/f_k = NT/k = 1440/k$ minutes).

From the figure, the most dominant time period over a work day is the day itself ($k = 1$). This matches our intuition from the data in Figure 3(b). There are several other low frequency components which have relatively large gains compared to the rest of the frequencies. Most others have small gain components.

We shall investigate the dominant frequencies in backbone and regional segments, shortly. Let us first ask ourselves the following question: Does the most dominant frequency, or a few dominant ones, explain most of the low frequency trends in the delay curve in Figure 3(b)? To get an intuitive visual feel for what the dominant frequencies mean, we plot the delays as predicted by dominant frequencies in Figure 8. In the background is shown the smoothed round-trip delay curve of 3(b), for comparison.

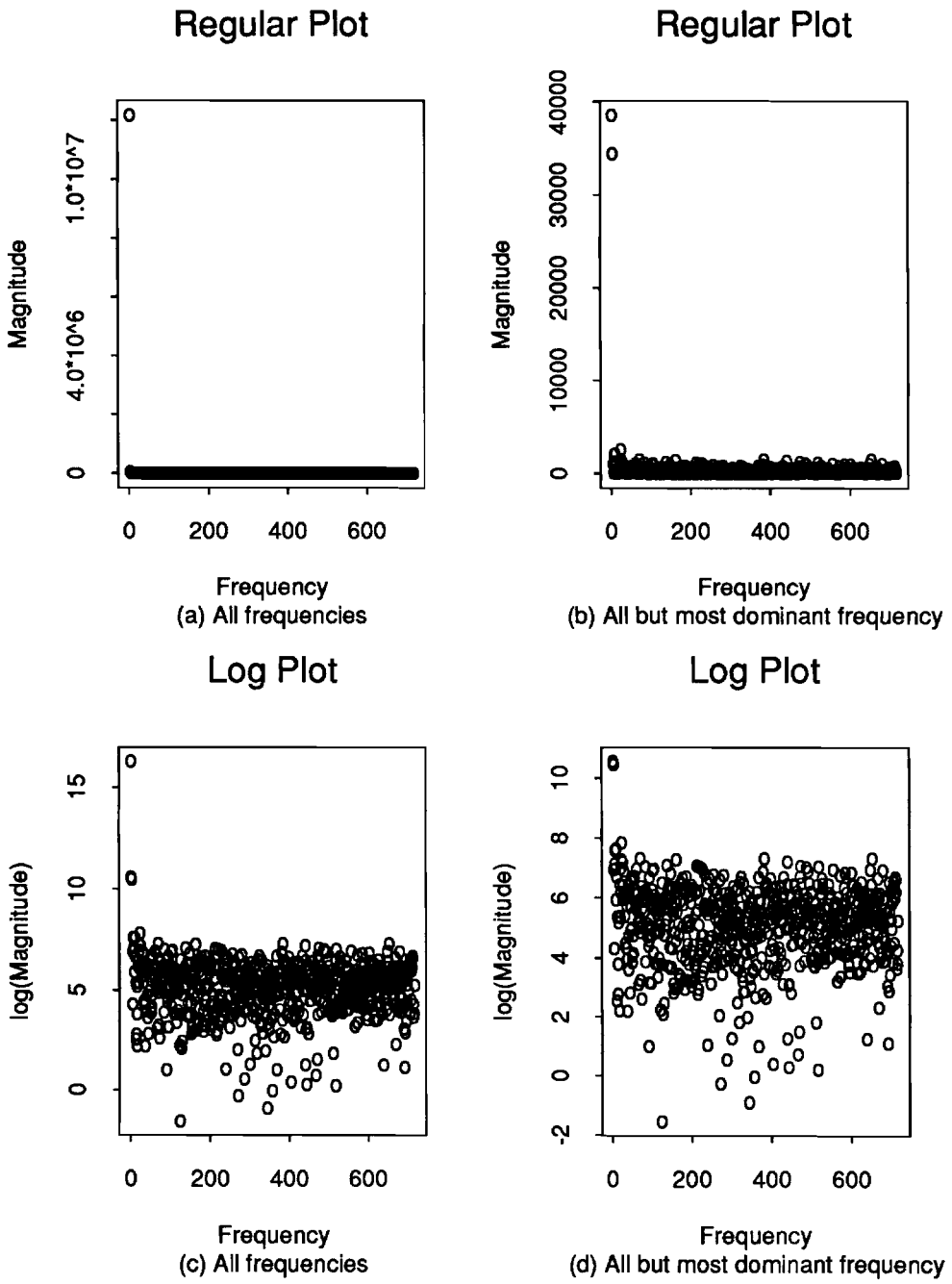


Figure 7: Gain versus Frequency for Cross-Country Data: (a) All frequencies; (b) Zoom into all but the most dominant frequency; (c) and (d): Same figures plotted in log scale. Notice that (i) the most dominant frequency is 1 (= 1 day), and this has a much larger gain than the rest, and (ii) there are several other (low) frequencies which might be included in the final model; the rest are essentially noise.

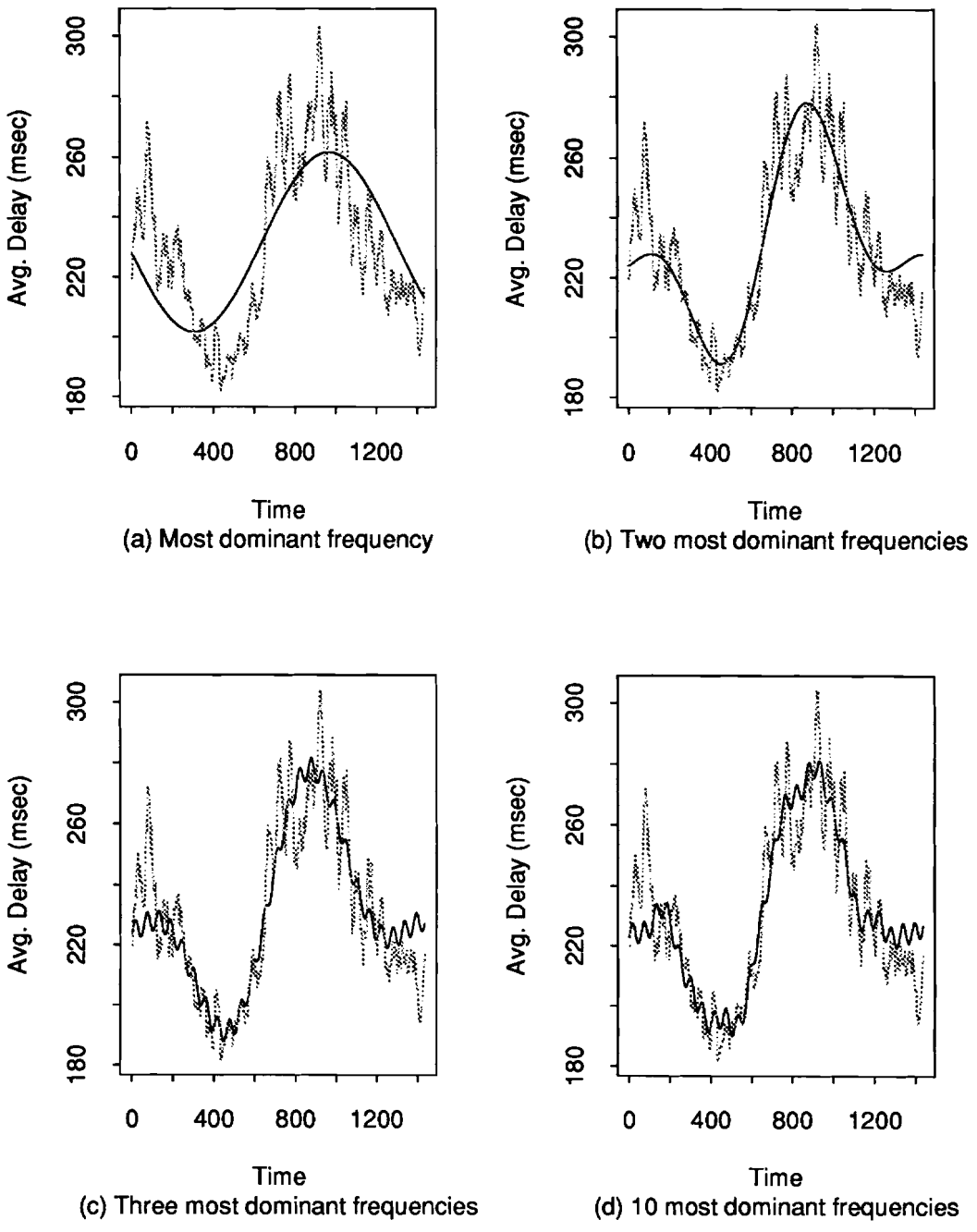


Figure 8: Average delay as a function of dominant frequency components.

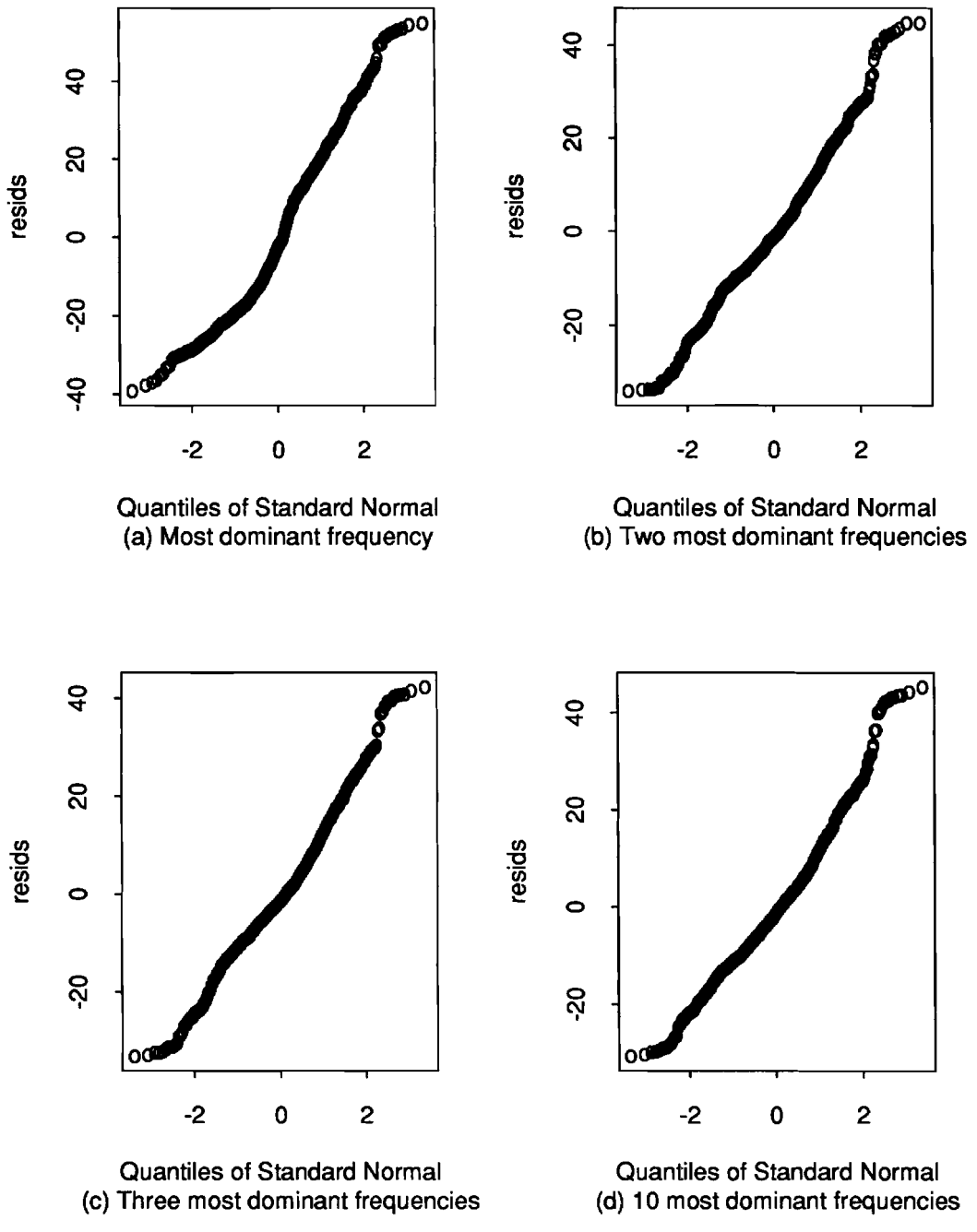


Figure 9: Quantile-Quantile Plot of residuals with respect to Normal Quantiles for the regression.

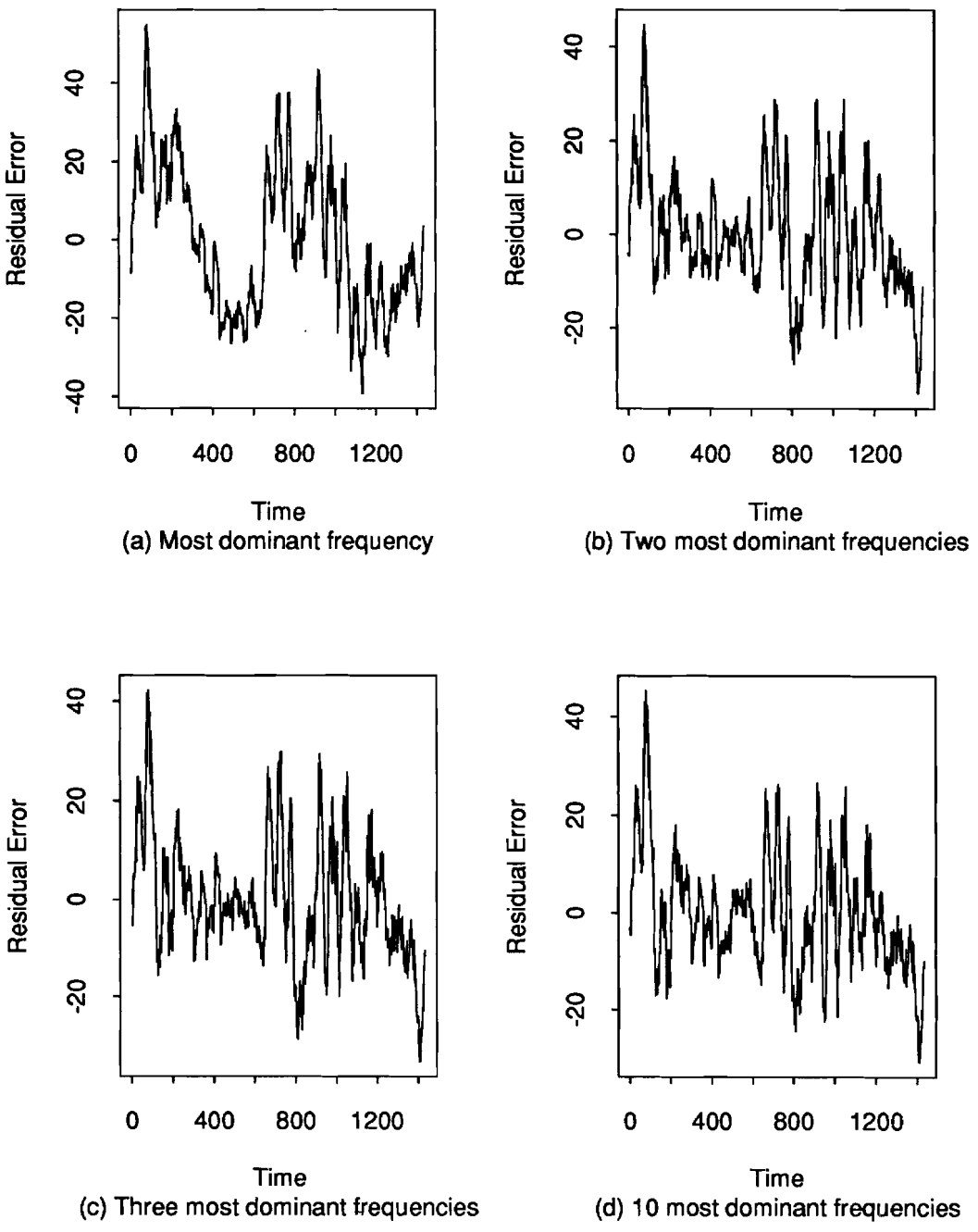


Figure 10: Residual errors in regression as a function of time.

There are several ways to plot the dominant frequencies; Figure 8 was plotted using the following method: given a set of dominant frequencies $\mathcal{F}$, let us assume that the roundtrip delay $d(t)$ can be modeled by[3]:

$$d(t) = \sum_{f_i \in \mathcal{F}} \left[a_i \cos(\frac{2\pi f_i t}{N}) + b_i \sin(\frac{2\pi f_i t}{N}) \right] + \epsilon(t), \tag{8}$$

where $\epsilon(t)$ are assumed to be independent Normally distributed errors. Since we have measured values for the average delays, $d(t)$, we may obtain a_i and b_i, $i/NT \in \mathcal{F}$, by a simple linear least squares regression[4] and this is what is shown in Figure 8.

Alternately, one could obtain the a_i's and b_i's directly from the fast fourier transform operation. Either way, there are going to be residual errors from the frequencies omitted. We shall show only the least squares fit, because the latter minimizes the residual errors in the least squares sense.

How good is the regression fit in Figure 8? For instance, are the error residuals ($\epsilon(t)'s$) Normally distributed? Further, are they independently distributed?

Figure 9 shows the Quantile-Quantile plots of the residuals with respect to the Standard Normal Distribution, for each of the four cases in Figure 8. (A Quantile-Quantile plot is a standard way to check if two distributions match. It is explained in more detail in Section 4 where the distribution of delay is investigated. The key, for now, is to check if the plot is a straight line.) It appears that the fit is reasonably close to a straight line. However, Figure 10, which plots the error residuals as a function of time, shows that there is a pattern in the error, implying that they are correlated. Notice, however, that the correlations decrease somewhat, as more frequencies are added in the regression, (compare for instance, Figure 10(a) with 10(d)). This suggests that while a few dominant frequencies visually capture most of the diurnal pattern in the data (Figure 8) and seem to yield Normally distributed errors (Figure 9), they do not yield uncorrelated errors, and hence are not statistically sufficient. This may be improved by including more frequencies in the set of dominant frequencies $\mathcal{F}$.

Experiments with only the peak period of the workday showed the presence of several other relatively strong frequencies. However, these frequencies were relatively weak during the night, and therefore, did not show high gains overall. When added to $\mathcal{F}$, regression over the entire data set does not add much weight for these frequencies because doing so would increase the errors during other periods of time. They are, however, present during peak hours. The figures have not been included here to save space.

Figure 11 shows the relative magnitudes of different frequencies in the backbone and the regional segments. We observe that:

(a) Several low frequency components dominate the backbone segment, similar to its cross-country counterpart.

[3]The sampling period T, for this particular data was 1 minute, and is therefore suppressed.

[4] From the data in Figure 3(b), it might appear that a phase angle may have to be carried along explicitly in the regression, i.e., $d(t)$ must be of the form:

$$d(t) = \sum_{f_i \in \mathcal{F}} \left[a_i \cos(\frac{2\pi f_i t}{N} + \phi_i) + b_i \sin(\frac{2\pi f_i t}{N} + \phi_i) \right] + \epsilon(t),$$

where ϕ_i is the phase angle corresponding to f_i. However, as shown in Appendix B, the need for ϕ_i is obviated by the form (8).

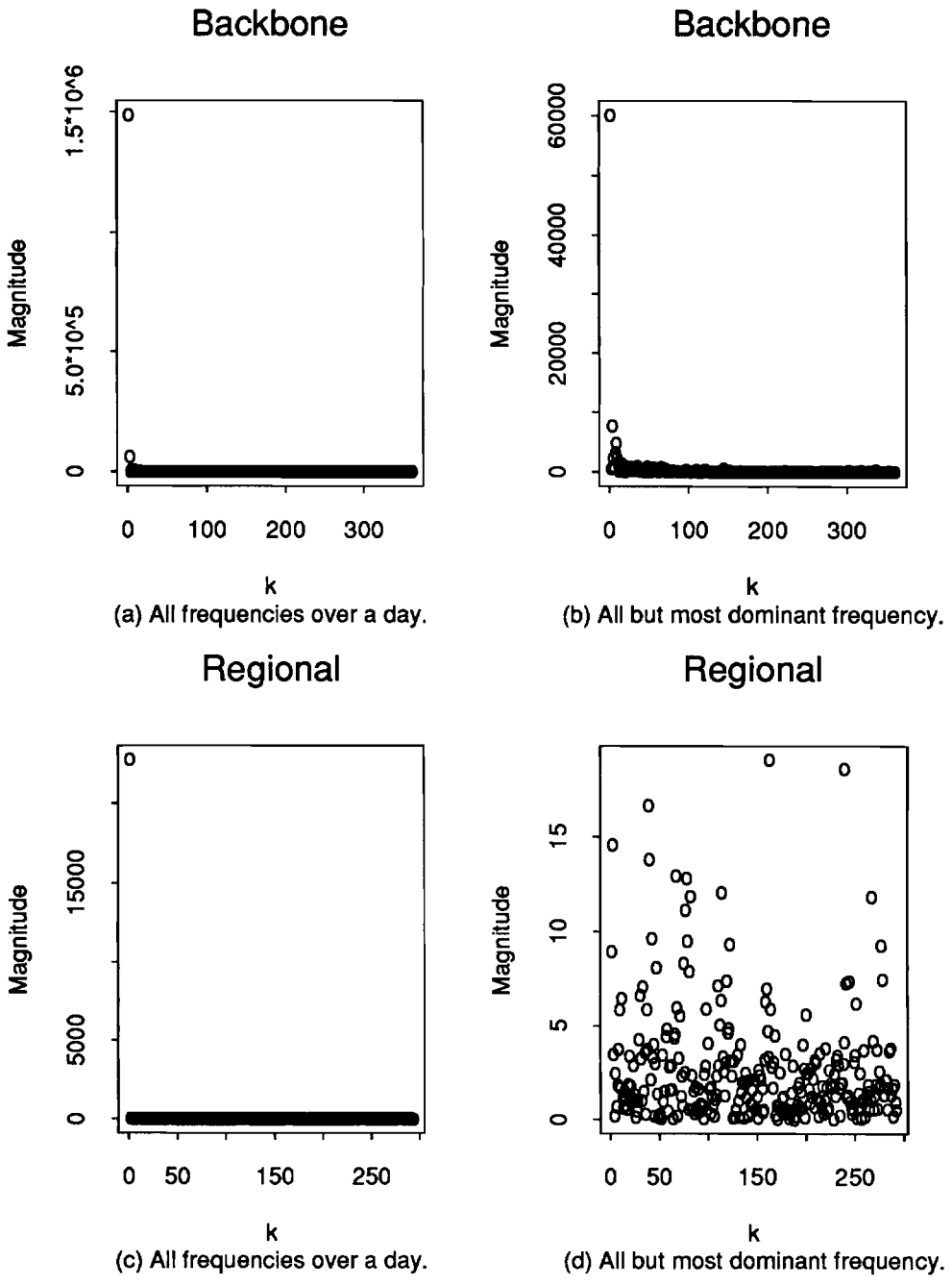


Figure 11: Magnitude versus Frequency for the Regional and Backbone Segments. The frequency corresponding to k is $f_k = k/NT$. T was 2min for these figures.

(b) The regional segment too has a strong diurnal frequency (Figure 11(c)), which was not obvious from Figure 4(b). However, unlike the backbone and cross-country segments, the other dominant frequencies are not necessarily at the low end of the frequency spectrum. (See Figure 11(d).) The chances of aliasing, indicated in Section 2.2, is, however, higher at this end of the frequency spectrum, so these frequencies ought to be interpreted with caution.

To summarize, what we learn from the frequency domain analysis is that the average delay over the cross-country and backbone segments have several dominant low frequencies. The most dominant frequency, which is the work-day itself, has a gain that is several orders of magnitude larger than any of the rest. However, there are other strong frequency components in the data, as seen from Figure 8. The most dominant frequency over the regional segment also corresponds to the work-day. However, the other relatively dominant dominant frequencies do not follow the same pattern as the backbone or the cross-country segments.

Shoch and Hupp [37], and Fowler and Leland [10] have reported a similar strong diurnal pattern in packet arrival rates over a LAN, and over an external campus gateway, respectively. It is interesting to note the similarities in the frequency domain between these studies and wide-area end-to-end delays.

Dominant low frequencies in end-to-end delays are potentially useful, if it can be established that there exists a relationship between the low frequency components of delay and congestion. We shall address that topic in Sections 5 and 6.

4 Distribution of Delay

This section investigates the distribution of end-to-end delay. It is important to keep in mind that the environment of the Internet is non-stationary and hence a sampling technique for estimating the distribution is at best an approximation. Nevertheless, given the relatively low amplitudes of high frequency components in Figure 7, one should be able to get an approximate idea of the distribution, if the time interval over which the distribution is estimated is kept short.

One piece of information that comes from this estimation process is that the delays are approximately Gamma-distributed. This is true for all network paths studied (regional, backbone and cross-country) and for all observed loads, although the accuracy of the approximation varies. The conclusions presented in this section are based on rigorous testing of over 72 hours of data.

4.1 Motivation

Why is the distribution of delay important? We believe there are several reasons:

(a) Distribution of end-to-end delay is an important component for modeling studies of feedback based congestion control where feedback delay is known to be an important component [2, 11, 16, 30, 33, 40] in performance. Since this distribution has hitherto been unknown, it has usually been assumed deterministic [2, 11, 30], essentially for lack of anything better. Knowledge of the actual delay-distribution could make the models more accurate, show its impact on performance, and enable investigation of new and improved congestion control algorithms.

(b) The Gamma distribution has two parameters, scale and shape. Given an end-to-end delay characterization by a two-parameter random variable, is it possible to enhance congestion detection algorithms based on an estimate of these two parameters? We do not yet know the answer to this question, but it served as a motivation for rigorous testing of the distribution, once we knew it was potentially a Gamma.

4.2 Fitting An Empirical Distribution.

4.2.1 Some Examples and Hypothesis.

Figures 12(a) and 12(b) show three example delay histograms computed over consecutive one-hour periods and consecutive 10-minute periods, respectively. Notice that (a) they move with time, (b) their tails are quite large, and (c) they look roughly similar to a Gamma density function, shifted by a constant. The similarity with a Gamma is better for the hourly histograms than for the 10-minute ones. Most likely, this is because there are more sample points for the former, hence the less jaggedness. The constant shift along the x-axis is the length of the shortest measured path during the one hour or 10-minute interval. We think the large tails warrant further investigation, especially if quality of service guarantees are to be provided in future networks.

Based on the above curves, we make the following hypothesis:

Hypothesis 1 *Let X be the random variable representing delay. Let c be a constant (equal to the shift in the x-axis in the histograms). Then $X - c$ is approximately Gamma distributed.*

To test this hypothesis, assume that Z is a Gamma distributed random variable with unit scale and shape s, i.e., its density function is given by:

$$f_s(z) = \frac{z^{s-1}e^{-z}}{\Gamma(s)}. \tag{9}$$

We shall find that X is of the form

$$X = mZ + c, \tag{10}$$

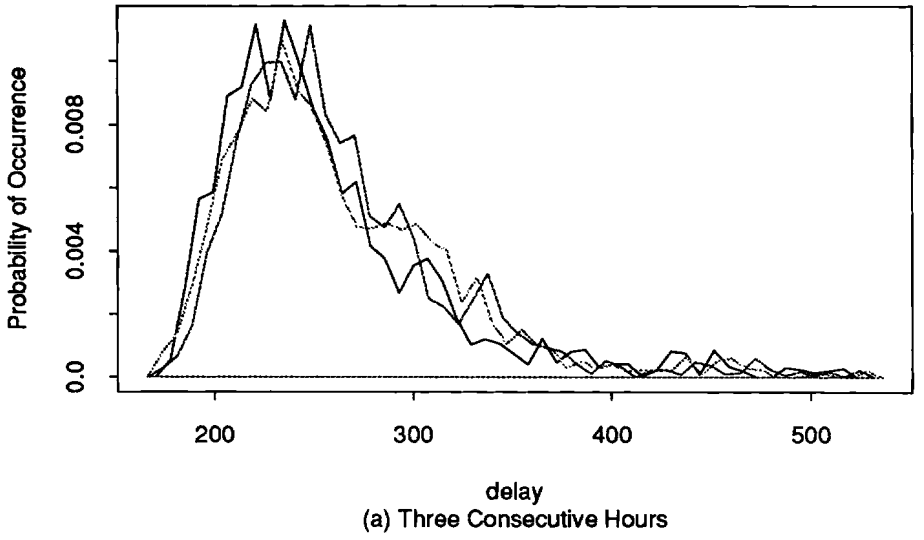
where m is the scale parameter of the Gamma distribution. (Recall that a Gamma density function with scale m and shape s is given by $f_{s,m}(x) = \frac{(z/m)^{s-1}e^{-z/m}}{\Gamma(s)}$.)

4.2.2 Quantile-Quantile Plots and Validation of Hypothesis.

Suppose $F(x)$ and $G(x)$ are the empirical and theoretical distributions, respectively. It is required to test if $F(x) = G(x)$.

A standard test of whether a distribution matches another is to check if their quantile-quantile plot is a straight line. The idea is to plot the observed quantiles versus the corresponding theoretical quantiles.

Suppose y_i is the empirical q_ith quantile, i.e., $F(y_i) = q_i$. For the same q_i, suppose x_i is such that $G(x_i) = q_i$, i.e., x_i is the q_ith quantile of the theoretical distribution. If $F(x) = G(x)$, the plot of y_i's versus x_i's will be a straight line with slope one and passing through the origin. The plot of y_i's versus the x_i's is known as a quantile-quantile plot.



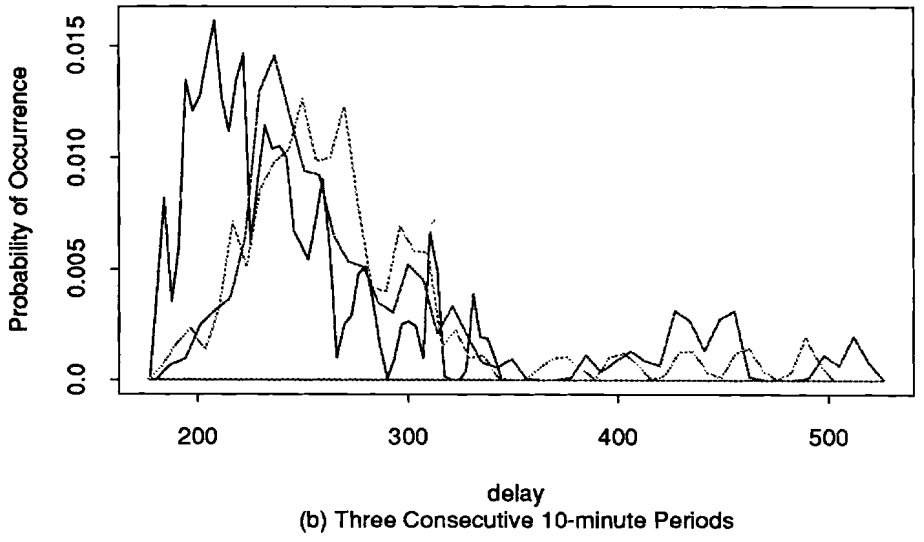


Figure 12: Example Delay Densities Over the Cross-Country Segment.

In Appendix A, it is shown that if X is Gamma-distributed with a form given by (10), then the quantile-quantile plot of the distributions of X and Z will be a straight line with slope m and intercept c. Therefore, it is sufficient to plot the quantiles of the observed distribution with that of a unit scale Gamma distribution, (9), and check for a straight line fit.

The next issue is the time interval over which the empirical distribution is computed. The smaller the interval, the less likely that the distribution could have shifted much. However, it also means a smaller sample size. The solution adopted was to test over a range of time intervals, from ten minutes to one hour.

We checked the quantile-quantile plots of empirical marginal distributions of delay versus the unit-scale Gamma distribution for over 72 hours of data for the three network segments combined. This was done using the S package [1]. Figures 13-15 show three representative plots, one for each network path. For the backbone and the cross-country segments, the straight line fit is reasonably good implying that the delay distribution is approximately a Gamma. For the regional network segment, however, the fit is not that good. (Although the quantile-quantile plot is a reasonable straight line, the density functions do not look similar; most likely this is due to the small shape parameter, which causes the unit-scale Gamma to be almost an impulse function.)

From the quantile-quantile plots, we observe that:

- the slope is not necessarily one; its value gives the scale parameter;
- the intercept is not necessarily zero; it is equal to the length of the shortest observed path. (If the path does not change, then the intercept, c, equals the constant propagation delay plus sum of packet service times along the path.)

4.2.3 Goodness of Fit

When is match with the Gamma distribution good? When is it not so good? This section summarizes the findings[5].

- **Backbone:**
  - match is good for low to medium values of shape (less than 5); both Q-Q plot comparison and visual matching of densities show the density to be similar to a Gamma.
  - match is not so good for large values of shape ($\approx$ 20). Empirical density function could be jagged at points. Q-Q plots show almost a straight line fit, but visual matching of densities is not as good as in the previous case.
  - Coefficient of variation can go up to 1.2 during low loads; it is lower (0.6 or less) during high loads! This is because the coefficient of variation of a Gamma random variable is $1/\sqrt{s}$.
- **Cross Country:**
  - match between the empirical and Gamma distributions is the best for this configuration.
  - shape parameter is in the medium range for all cases. This may in fact be the reason for the good fit.
- **Regional:**

[5]This summary is based on volumes of graphs which are not included here due to obvious space constraints.

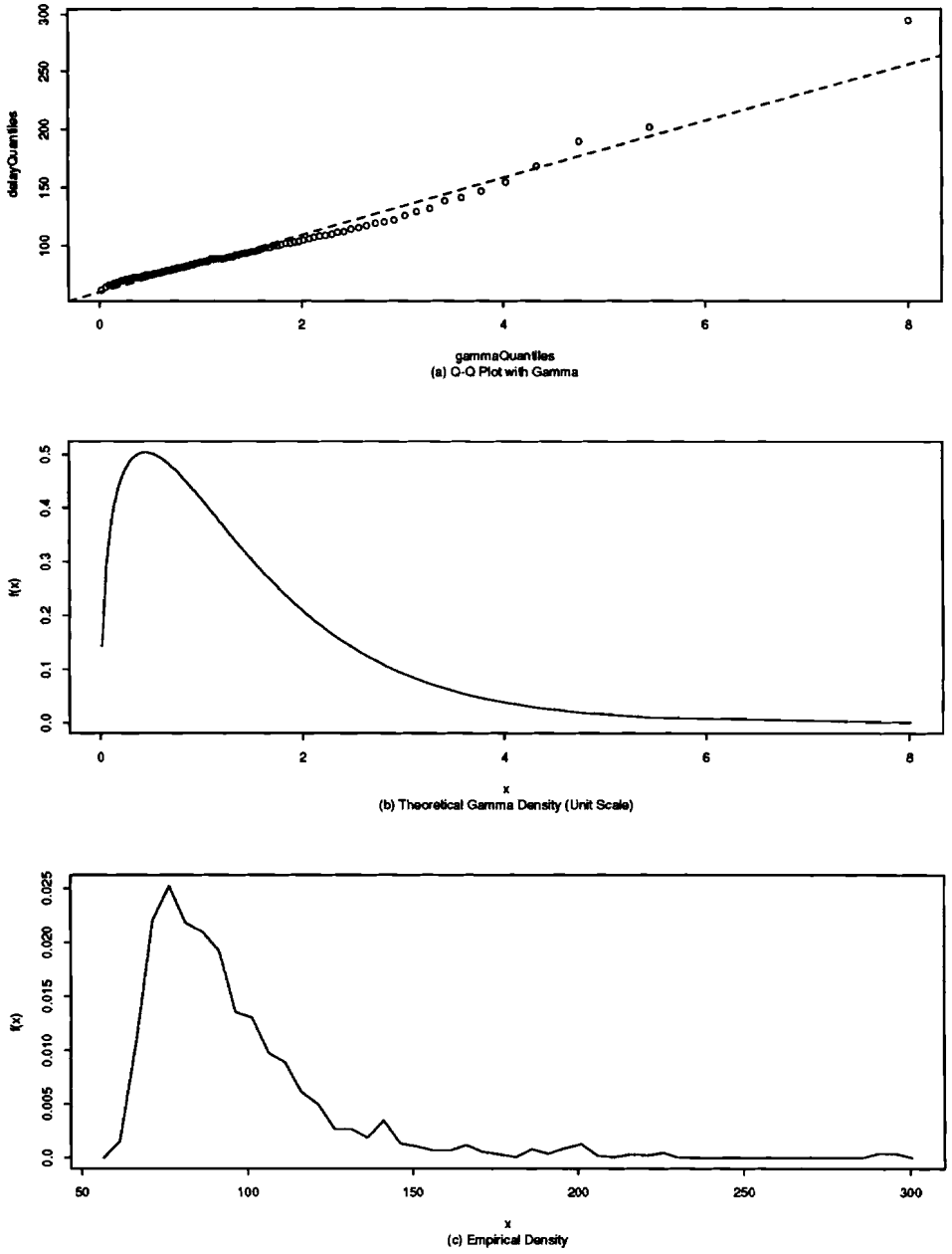


Figure 13: Distribution of delay over the backbone (time interval = 1 hour). (a) Quantile-Quantile plot of empirical and unit-scale Gamma distributions. Observe that it is almost a straight line; (b) The theoretical, unit-scale Gamma density function. Its shape is estimated from the data; (c) The empirical density function.

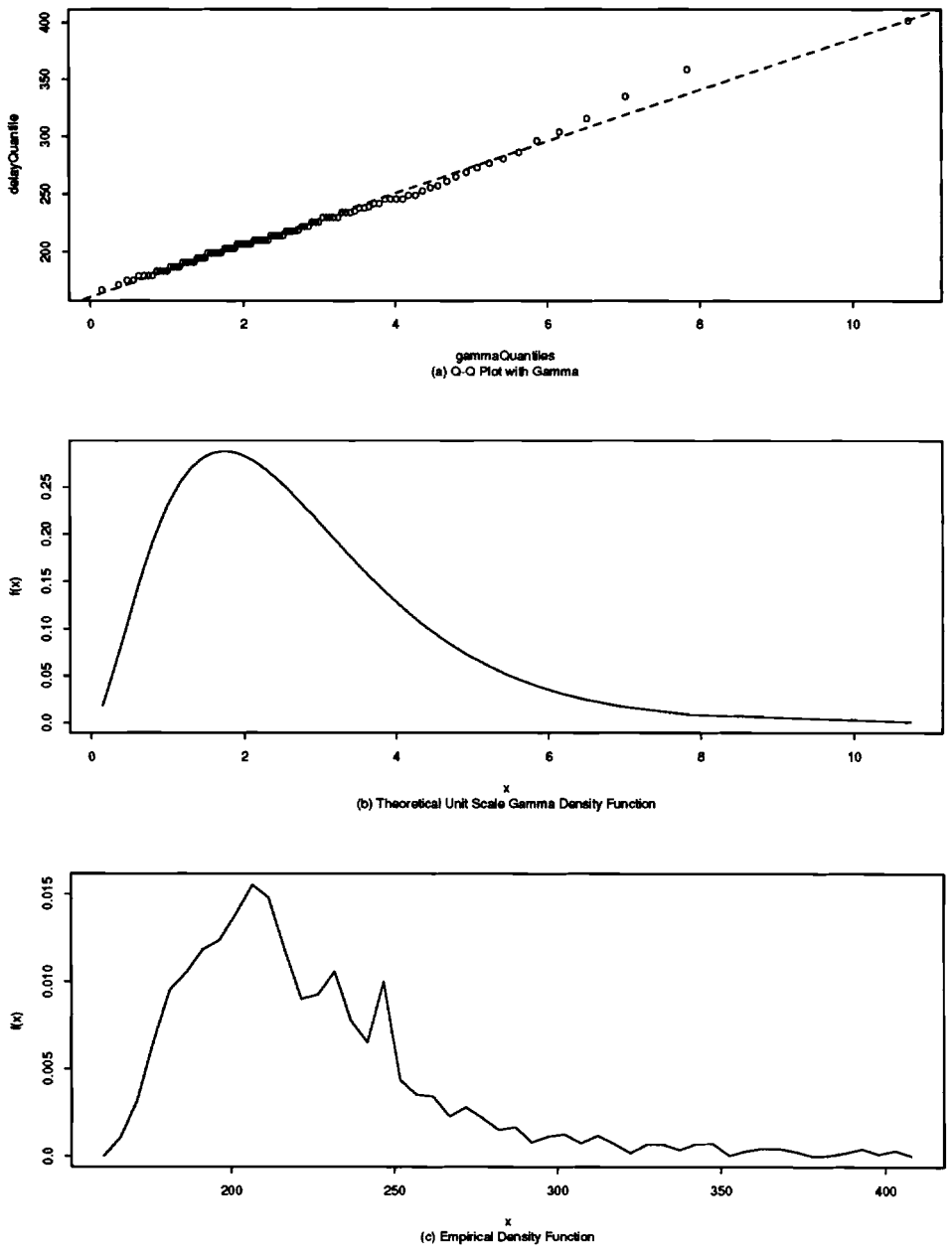


Figure 14: Distribution of delay over the cross-country segment.

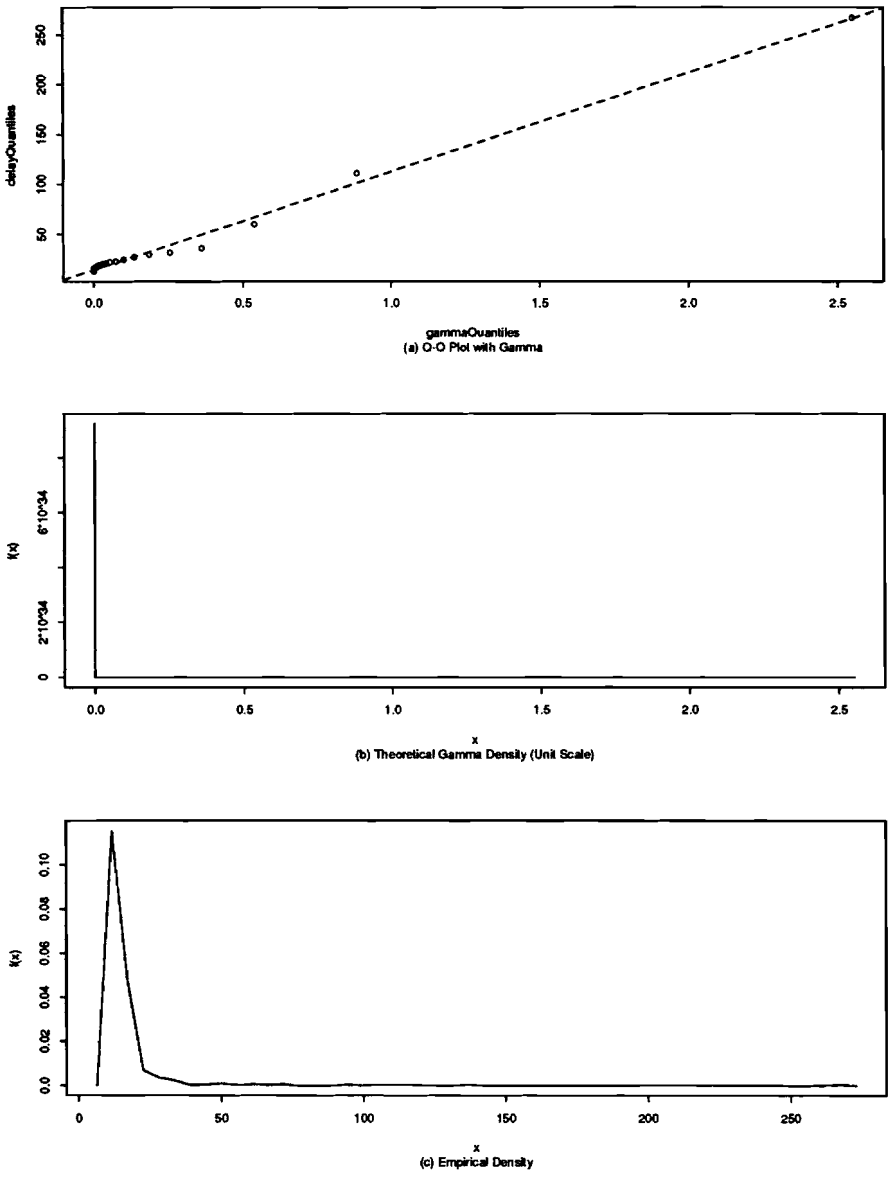


Figure 15: Distribution of delay over the regional segment.

- shape is much smaller; it ranges from 0.06 to 0.20.
- load is low; coefficient of variation is high;
- quantile-quantile fit with Gamma is good, but empirical and theoretical densities look different.

4.2.4 Variation of Shape and Scale Parameters Across Network Paths and Network Loads.

From Figure 16, the shape and scale parameters vary considerably across networks, and across time over the same network. It is small for small underloaded networks and increases for large and/or overloaded network.

For the Jon von Neumann Center Network (JVNC), where the load is low and the number of hops is small, s was between 0.05 and 0.20, although for most cases, it was less than 0.10. For the backbone, s varied from approximately 1.0 during low loads to 6.0 during high loads. For smaller sampling periods, s could go up to as large as 20. For the cross-country segment, s varied from less than 1.0 to 8.0

The variation of s with load makes intuitive sense: an increase in s implies a more spread in the distribution, which is to be expected at high loads. What is interesting is that the distribution is still approximately a Gamma.

A possible application for this is in modeling studies of congestion control where one may assume that the delay is Gamma-distributed and study the performance of feedback control as a function of the shape (s), scale (m) and shift (c) parameters. Relative sensitivities of control versus shape, for instance, might show new insights on control algorithms.

A second application could be along the lines of a delay based congestion control algorithm. For instance, given that s increases with load, and that the distribution of delay is a Gamma, is it possible to devise new algorithms for congestion detection and avoidance based on an estimation of s?

4.3 Why is the Distribution a Gamma?

Why is the Distribution a Gamma? We do not know the precise answer to this question. However, we do know that:

(a) The Gamma distribution is a versatile distribution in that several important distributions that are known to model nature well (e.g., Exponential, Erlang and χ-Square), are special cases of the Gamma distribution.

(b) End-to-end delays are positive random variables with empirical histograms that are non-symmetric, and with long tails. This is captured well with a Gamma Distribution (at least better than, say, a Normal Distribution, which was found to be inadequate). Possible reasons for the large tail include a First-Come-First-Served (FCFS) scheduling and a systematic batching of traffic [6].

Perhaps a more critical question is, why is the tail so large, even in lightly loaded segments, and what can be done to prevent it. If the reason is FCFS scheduling, the solution lies in using a different scheduling discipline, e.g., Fair Queueing [5]; if, however, it is due to some form of traffic synchronization [6], or perhaps something else, a more serious investigation would be required to prevent it from occurring.

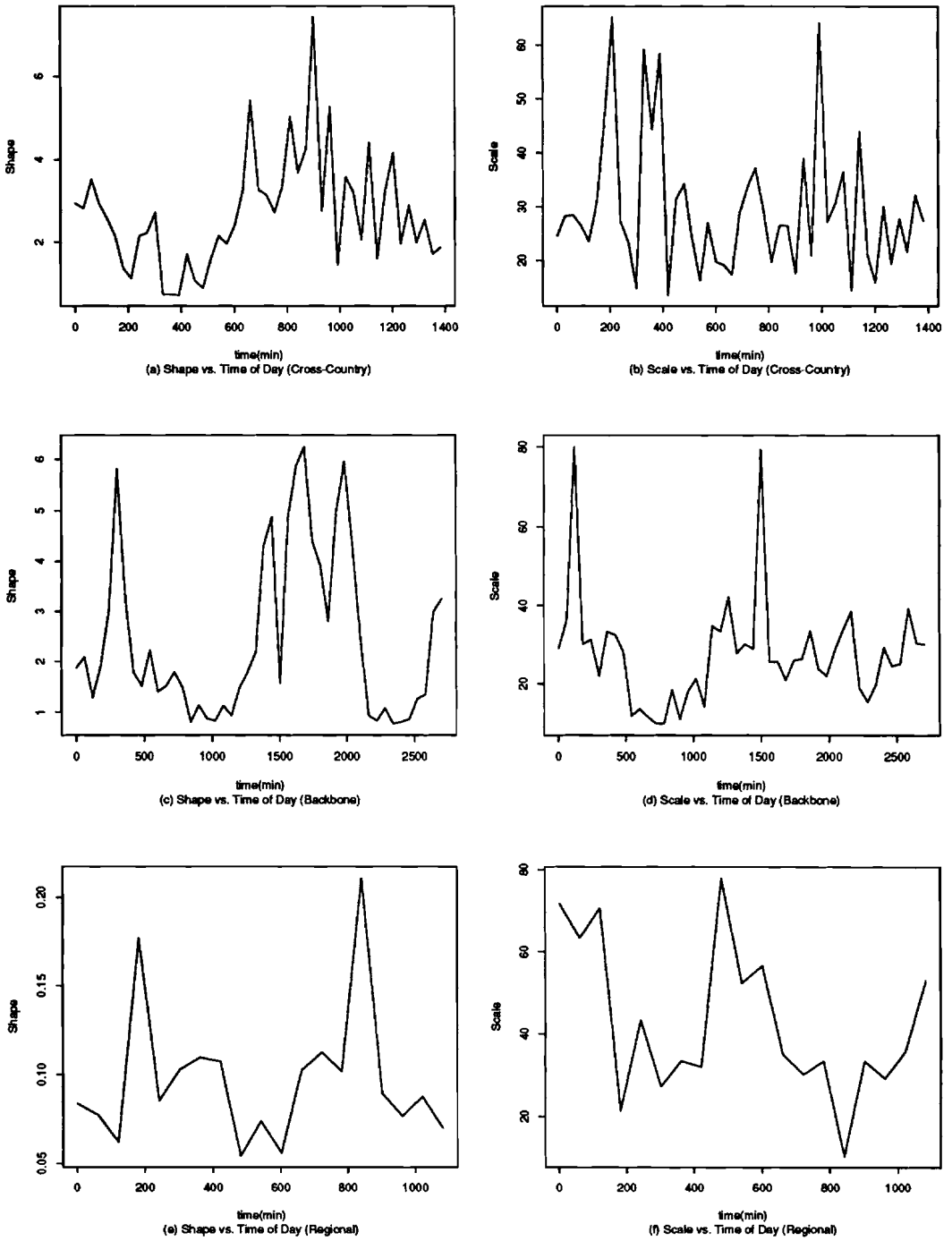


Figure 16: Scale and Shape Parameters as a Function of Time for the Three Network Segments. The distribution parameters shown here were computed over 30 minute intervals.

5 Loss vs. Delay

This section investigates the correlation between sampled end-to-delay and observed packet loss.

Why should low frequency components of delay be interesting? Do they point to a base level congestion along a path that may be effectively tracked with a feedback mechanism? If this were so, then even in high speed networks, the congestion information could be useful (in spite of delay in the feedback path).

Most feedback based schemes proposed in the literature, e.g., [9, 16, 20, 22, 25, 27, 32, 33, 34], would benefit if low frequency components of delay were to show a correlation with packet loss. For example, the Decbit strategy [33] of averaging the number of congestion signals over an interval and adjusting the window for the following interval assumed a correlation of traffic across intervals. Low frequency components were also assumed for the Dynamic Time Window flow control protocol [9] where transport endpoints adapted their burstiness as a function of (low frequency) switch loads along their paths.

The principal goal of a feedback based congestion control strategy is to constrain the high frequency components induced by a transmitter, as a function of load measured over a previous control interval. A correlation between the control intervals is therefore important for such a scheme to be effective. A new algorithm reported by Ramamurthy and Sengupta [34] is perhaps, the most direct approach for exploiting this correlation.

Correlation across time intervals, off course, implies the presence of dominant low frequencies in load variation. Figures 3(b), 4(a) and 4(b), in Section 2.3, suggest that, that indeed is the case. In this section, we investigate the relationship between the low frequency components of delay and observed packet loss.

5.1 Base congestion level.

Suppose that at time t, k out of n packets return successfully and the rest $\ell = n-k$ packets are lost in the network. Let the observed delays of the k packets be $\{d_1, d_2, \cdots, d_k\}$, and let $d_{av}(t)$, $d_{sdev}(t)$, $d_{min}(t)$ and $d_{max}(t)$ be the sample mean, standard deviation, minimum and maximum of this set, respectively.

Figures 17(a)-17(d) show the scatter plots of the number of lost packets versus d_{av}, d_{sdev}, d_{min}, and d_{max}, respectively, for the cross-country segment. They were computed as follows. For each time t, if there were measured values of loss ($\ell(t)$) and delays, $d_{av}(t)$, $d_{sdev}(t)$, $d_{min}(t)$, and $d_{max}(t)$, then the scatter plot of, say packet loss versus average delay, was plotted by marking the point $(\ell(t), d_{av}(t))$ in Figure 17(a).

Several observations can be made from Figures 17(a)-17(d):

(a) There is no simple linear relationship between packet loss and delay. (We use 'delay' as a generic term for d_{av}, d_{sdev}, d_{min}, and d_{max}.) In fact, for a given value of packet loss, there is a wide spectrum of possible delay values.

(b) Similarly, for a given value of delay, there is a spectrum of possible packet losses. Fortunately, the zero loss case seems to be the preferred one.

(c) There is a shift in the *distribution* of packet loss as delay increases. To see this, observe that the scatter plots shift to the right, somewhat, for larger values of packet loss.

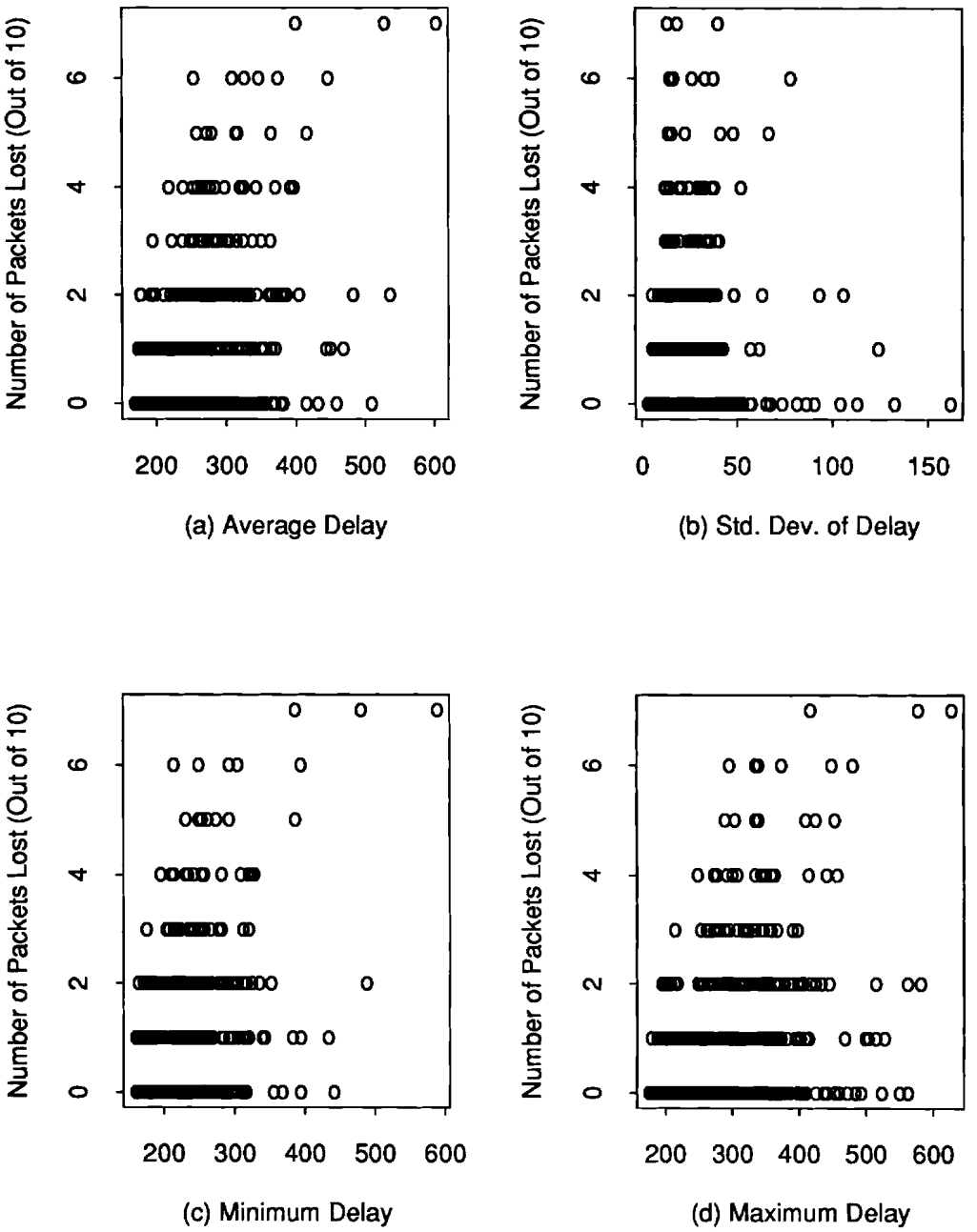


Figure 17: Scatter plot of Loss versus Average, Standard Deviation, Minimum and Maximum Delay for the Cross-Country Segment.

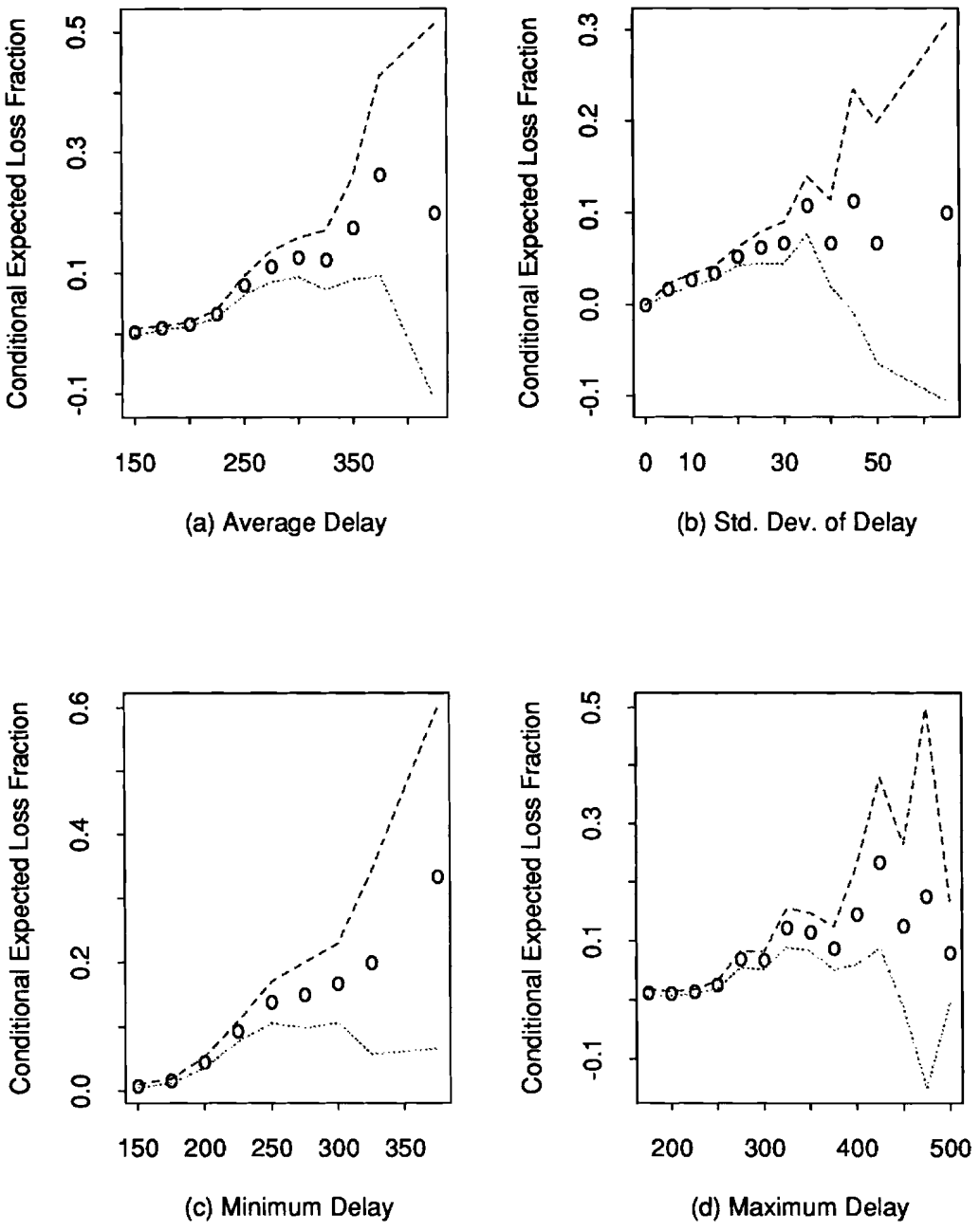


Figure 18: Conditional Expected Loss ($E[L|d]$) versus Average, Standard Deviation, Minimum and Maximum Delay for the Cross-Country Segment. Dashed lines indicate the 90% Confidence Intervals. Notice that the confidence interval is quite tight for low to medium range of the delay spectrum, but gets wide at the higher end. We have let the confidence interval trail into the negative domain, although the conditional expectation can never be negative.

It may be argued, given observation (a) above, that a significant proportion of packet losses are caused by some phenomenon other than congestion, see [35]. While we cannot ascertain if that is true from our experiments, observation (c) suggests that there does exist a congestion-component in packet loss, as has widely been believed. (See for example, [16].)

To validate this, we investigate the conditional expectation of loss given a certain value of delay. Let $L = \ell/n$ be the fraction of packets lost in a packet group, and let the conditional expectations with respect to delay be defined as $E\{L = l/n|d_{av}\}$, $E\{L = l/n|d_{sdev}\}$ $E\{L = l/n|d_{min}\}$ and $E\{L = l/n|d_{max}\}$. We shall explore the properties of $E\{L = l/n|d_{av}\}$, $E\{L = l/n|d_{sdev}\}$ $E\{L = l/n|d_{min}\}$ and $E\{L = l/n|d_{max}\}$ as a function of d_{av}, d_{sdev}, d_{min}, and d_{max}, respectively.

Figure 18 shows $E\{L = l/n|d_{av}\}$, $E\{L = l/n|d_{sdev}\}$ $E\{L = l/n|d_{min}\}$ and $E\{L = l/n|d_{max}\}$ as a function of d_{av}, d_{sdev}, d_{min}, and d_{max}, respectively. These figures were computed in a straightforward way, as follows. Given a delay value, d, (where d is one of d_{av}, d_{sdev}, d_{min}, and d_{max},) let $L_1, L_2, \cdots, L_n$ be the set of measured loss fractions. Then an estimate of the conditional expectation given d is:

$$E\{\hat{L}|d\} = \frac{1}{n}\sum_{i=1}^{n} L_i, \tag{11}$$

and its sample variance is

$$\sigma^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} {L_i}^2\right) - E\{\hat{L}|d\}^2. \tag{12}$$

Finally, the $100(1-\alpha)\%$ confidence interval for the conditional expectation is given by:

$$\left(E\{\hat{L}|d\} - t_{1-\alpha/2;n}\frac{\sigma}{\sqrt{n}},\ \ E\{\hat{L}|d\} + t_{1-\alpha/2;n}\frac{\sigma}{\sqrt{n}}\right), \tag{13}$$

where $t_{1-\alpha/2;n}$ is the $(1-\alpha/2)$th quantile of a t-distribution with n degrees of freedom.

In practice, one needs to compute the conditional expectations over a neighborhood of delays, instead of computing them over singleton d's. This is similar to using buckets for computing histograms, so as to increase the number of samples in a bucket. The conditional expectations in Figures 18(a), 18(c), and 18(d) were computed using a bucket of 25msec, while that in Figure 18(b) was computed using a bucket of 5msec.

From Figures 18(a)-18(d), we observe that the conditional expected loss, given delay, does indeed have an increasing trend. For smaller values of delay, there is no appreciable loss, but as delay increases, so does the expected packet loss, as one would expect if there exists a base congestion level. The scatter plots in Figures 17(a)-17(d) indicate, however, that the correlation between delay and packet loss could be low. We shall compute the correlations in Section 5.2.

The conditional expectations of packet loss with respect to delay for the backbone and regional segments were found to have a similar increasing pattern. The figures are not shown here to save space.

5.2 Correlation Coefficients

The correlation coefficient, ρ_{ij}, between two random variables X_i and X_j is defined as

$$\rho_{ij} = \frac{{\sigma^2}_{ij}}{\sigma_i \sigma_j},$$

Network Path	Average Delay	Std. Dev.	Minimum	Maximum
Cross Country (June 4)	0.4937719	0.1832922	0.5240888	0.425907
Backbone (July 27)	0.5598968	0.3251563	0.7132696	0.4585366
Regional (August 5)	0.06565072	0.04922241	0.02348277	0.2721359

Table 5: Correlations Between Packet Loss and Different Statistics of Delay.

where ${\sigma^2}_{ij}$ is Covariance (X_i, X_j).

Sample correlations between packet loss and different functions of delay for the three network segments is shown in Table 5. Each of these was computed over approximately one working day's worth of data.

We observe that:

- The correlations for backbone and cross-country segments are greater than that for the regional segment. We believe this is because losses in the regional segments are rare.
- For the backbone and cross-country segments, the correlation between average, minimum and maximum delays and packet loss are larger than that between standard deviation and packet loss; this is somewhat surprising;
- For the regional segment, the correlation values are extremely low, implying that the network is lightly loaded. (The next subsection discusses this in more detail.) The maximum delay statistic, however, shows a somewhat more tangible positive correlation with packet loss.

The observed correlations were also found to change with the time of day and whether it was a weekday or a weekend. We shall address this in more detail in Section 5.4.

5.3 Magnitude of Correlation: Implications

Without loss of generality, assume that $d = d_{av}$. The correlation between loss and average delay computed over an entire working day was found to be between 0.49 and 0.56 over the backbone and cross-country segments and much less for the regional network. An important question is, can the congestion state of the network be inferred accurately from average roundtrip delays? Stated differently, what is the implication of the magnitude of correlation on inference of congestion state?

To answer this question, consider the Venn-diagram of the sample space of loss ($\mathcal{L}$) and average delay ($\mathcal{D}$) shown in Figure 19. For a given value of delay, if there is no packet loss, let the sample point be in $\bar{\mathcal{L}} \cap \mathcal{D}$. However, if there is a loss, let the sample point be in either $\mathcal{D} \cap \mathcal{L}$ if the loss is caused by congestion, or in $\bar{\mathcal{D}} \cap \mathcal{L}$ if the loss is caused by some other phenomenon.

Now, let D and L be random variables representing delay and packet loss, respectively. Let D have two orthogonal components D_ℓ and $D_{\bar{\ell}}$, where $D_\ell \in \mathcal{D} \cap \mathcal{L}$ and $D_{\bar{\ell}} \in \bar{\mathcal{L}} \cap \mathcal{D}$. Similarly, let L have two orthogonal components L_d and $L_{\bar{d}}$, where $L_d \in \mathcal{L} \cap \mathcal{D}$ and $L_{\bar{d}} \in \bar{\mathcal{D}} \cap \mathcal{L}$. Then, without loss of generality, let[6]

$$D = D_\ell + D_{\bar{\ell}}, \tag{14}$$

[6]Alternately, one may assume that $D = \alpha_1 D_\ell + \alpha_2 D_{\bar{\ell}}$. In that case, one needs to replace $\sigma^2(D_\ell)$ in the sequel, with ${\alpha_1}^2\sigma^2(D_\ell)$, and $\sigma^2(D_{\bar{\ell}})$ with ${\alpha_2}^2\sigma^2(D_{\bar{\ell}})$. The key idea, however, remains the same, so we drop the additional notational complexity.

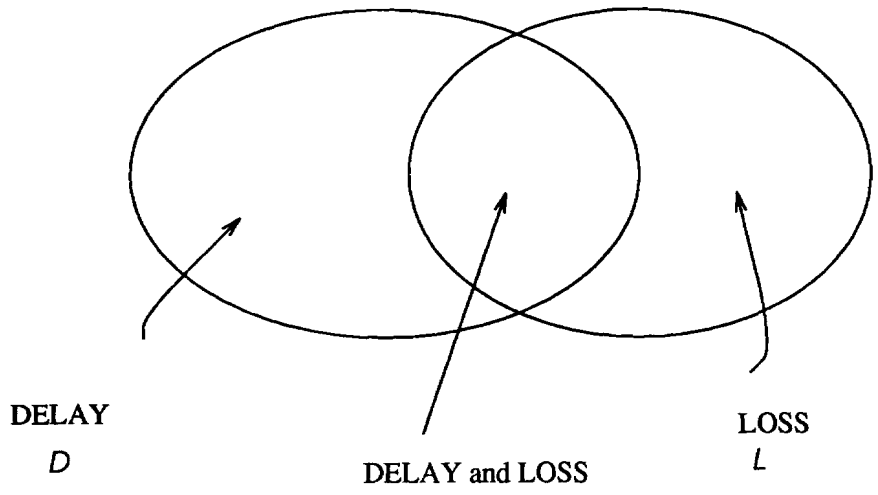


Figure 19: Sample space of loss and average delay

and

$$L = L_d + L_{\bar{d}}. \tag{15}$$

In words, D has two orthogonal components D_ℓ and $D_{\bar{\ell}}$, and L has two orthogonal components L_d and $L_{\bar{d}}$. From the assumptions on D_ℓ and L_d, we may assume that the correlation between D_ℓ and L_d is one. Let us further assume that they are linearly related:

$$D_\ell = aL_d + b, \tag{16}$$

where a and b are constants.

How accurately can roundtrip delays predict loss? The following lemma gives a fundamental limit in prediction accuracy, governed by the standard deviations of D_ℓ, $D_{\bar{\ell}}$, L_d and $L_{\bar{d}}$.

Lemma 1 *The correlation between D and L is given by*

$$\mathrm{cor}(D, L) = \frac{a\,\sigma^2(L_d)}{\sqrt{a^2\sigma^2(L_d) + \sigma^2(D_{\bar{\ell}})}\sqrt{\sigma^2(L_d) + \sigma^2(L_{\bar{d}})}}. \tag{17}$$

The proof is in Appendix C.

Implication:

Let us assume, for exposition purposes, that all losses are caused by congestion (as measured by the average delay), i.e., no losses belong to the domain $\bar{D} \cap L$, and therefore, $\sigma^2(L_{\bar{d}}) = 0$. In that case, (17) simplifies to

$$\mathrm{cor}(D, L) = \frac{a\,\sigma(L_d)}{\sqrt{a^2\sigma^2(L_d) + \sigma^2(D_{\bar{\ell}})}}. \tag{18}$$

From (16), $a\sigma(L_d) = \sigma(D_\ell)$. Therefore, (18) further simplifies to

$$\mathrm{cor}(D, L) = \frac{\sigma(D_\ell)}{\sqrt{\sigma^2(D_\ell) + \sigma^2(D_{\bar{\ell}})}}. \tag{19}$$

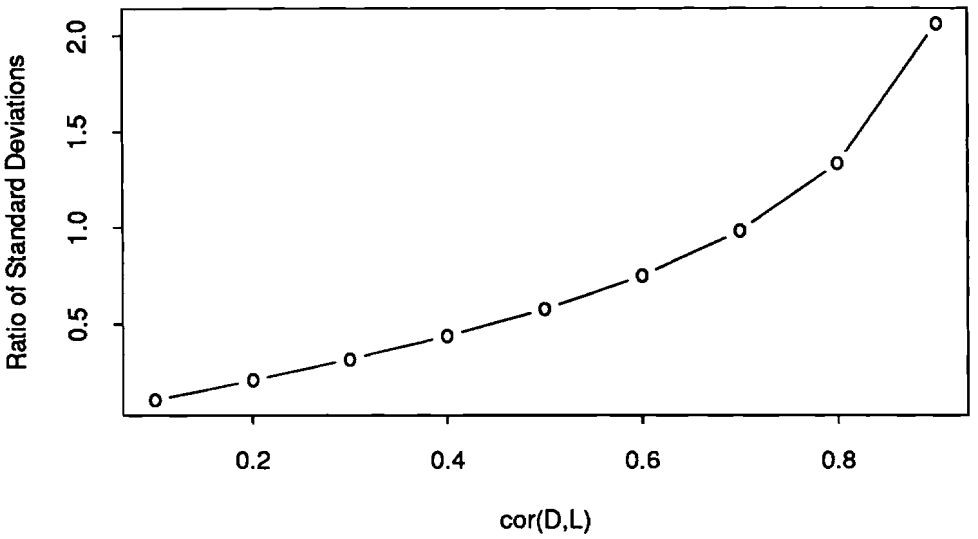


Figure 20: $\sigma(D_\ell)/\sigma(D_{\bar{\ell}})$ versus $\text{cor}(D, L)$

From (19), $\text{cor}(D, L)$ depends on the relative magnitudes of $\sigma(D_\ell)$ and $\sigma(D_{\bar{\ell}})$. In particular, a large average delay is likely to cause a loss if $\sigma(D_\ell)/\sigma(D_{\bar{\ell}})$ is large. The question is how well can average delay predict loss? Denoting $\text{cor}(D, L)$ by ρ, we get from (19):

$$\begin{aligned} \sigma^2(D_\ell) &= \rho^2(\sigma^2(D_\ell) + \sigma^2(D_{\bar{\ell}})) \\ \Rightarrow \quad \frac{\sigma(D_\ell)}{\sigma(D_{\bar{\ell}})} &= \frac{\rho}{\sqrt{1-\rho^2}}. \end{aligned} \tag{20}$$

Figure 20 shows the relative magnitudes $\sigma(D_\ell)/\sigma(D_{\bar{\ell}})$ for different values of ρ. We observe that $\sigma(D_\ell)$ is smaller than $\sigma(D_{\bar{\ell}})$ for ρ less than 0.7. For the backbone and cross-country segments, ρ was in the range 0.49 to 0.56 for the average delay case, and 0.52 to 0.71 for the minimum delay case; consequently, $\sigma(D_\ell)$ can be expected to be much smaller than $\sigma(D_{\bar{\ell}})$ for the average delay case and somewhat larger for the minimum delay case. For instance, for $\rho = 0.5$, $\sigma(D_\ell)$ is $0.58\,\sigma(D_{\bar{\ell}})$.

The implication is that $\sigma(D_{\bar{\ell}})$ dominates $\sigma(D_\ell)$ for the range of correlations measured. This is most likely because there are enough buffers along the way. For example, if there are k hops along a path, then the delay, d, is a sum of $2k$ individual queueing delays, plus the propagation delay, i.e.,

$$d = \sum_{i=1}^{2k} d_i + P, \quad d_i > 0, \;\; P \geq 0 \tag{21}$$

where d_i are the individual random delays and P is the propagation delay. A loss will only occur if a particular queue (out of the $2k$ constituents) overflows, i.e., $d_i > d_{i,max}$, for some i. For small d, there will exist many alternative possibilities for which none of the queues would overflow, and a few possibilities for which some would. (See (21).) This would explain the low correlation between packet loss and delay when losses and/or delays are low.

For larger d's, however, the number of alternative situations when queues could build

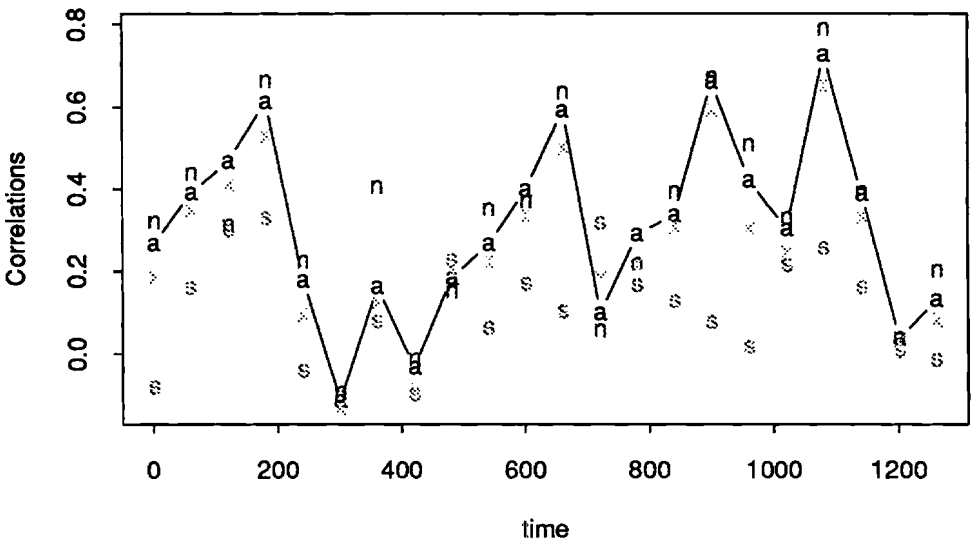


Figure 21: Correlation between loss and different statistics of delay as a function of time of day, for the cross-country segment. 'a' stands for average, 's' for standard deviation, 'n' for minimum and 'x' for maximum delay. For convenience in comparing the correlations, the points marked 'a' are joined together by a line. All points above the line are instances when a statistic other than d_{av} yielded a higher correlation. Conversely, all points below the line indicate when they yielded a lower correlation than d_{av}.

up and overflow, increases, and the alternatives when no queue overflows, decreases.(See (21).) This explains why correlations increase as the delay and/or loss increases.

In terms of the Venn-diagram in Figure 19, Lemma 1 says that most of the sample points would fall in $\bar{\mathcal{L}} \cap \mathcal{D}$ for low values of delay, while a relatively larger number of points would fall in $\mathcal{D} \cap \mathcal{L}$ for larger values of delay.

To summarize:

- Increased buffering and other strategies that reduce loss probabilities also decrease the correlation between loss and delay.
- The good news is that the observed correlation between loss and delay is low, because it implies low congestion.
- The not so good news is that implicit congestion detection is hard given the low correlation, and will get harder, as lower loss probabilities are desired by future applications. Explicit signaling of congestion and/or resource reservation, or a guaranteed smooth flow of traffic as in Stop-and-Go queueing [12] may be necessary to ensure low loss.

5.4 On Heuristics for Congestion Detection and Avoidance

Suppose, one were interested in heuristic algorithms for congestion avoidance. Which of minimum, maximum, average or standard deviation of delay should one use for adjusting flow control parameters?

To answer this question, we plot the relative magnitudes of the correlations between packet loss and d_{av}, d_{sdev}, d_{min} and d_{max} as a function of time. Figure 21 shows these

correlations for the cross-country segment. Each point was computed on an interval's worth of data, where the interval size was arbitrarily chosen as one hour. (We have played around with other values as well.) We observe that the magnitude of the correlations change with time. Also, there is no given statistic of delay that has a consistently higher correlation with packet loss than others.

Similar figures for the other network segments, and over the cross-country segment over different days, indicate the same phenomenon: that none of the statistics of delay considered, yields a consistently higher correlation than others. An important open question is, can one find a function of the delay statistics, d_{av}, d_{sdev}, d_{min}, d_{max}, or perhaps other order-statistics of delay, that would improve upon and show higher correlations than any of the individual components. If there exists such a function, it could lead to better heuristics for congestion inference. This subject is currently under investigation.

6 Out of Order Sequence of Acknowledgments

This section investigates the correlation between sampled end-to-delay and out-of-order sequence of acknowledgments. Recall that packets in each packet-group were separated by a spacing of $1/\lambda$ time units. Out-of-order return of acknowledgments were found to be present only for larger values of λ. In particular, out-of-sequence acknowledgments were observed for $\lambda = 60$Hz, but not for $\lambda = 1$Hz, which is to be expected, given that all round-trip delays were less than one second.

In an effort to minimize disruption of regular network service, our experiments were performed at a maximum rate of $\lambda = 60$Hz, and that too, only for one specific segment, the cross-country network.

Analogous to the packet loss case in Section 5, suppose that at time t, ℓ out of n acknowledgments return out of order, m of them return in order and the rest ($k = n - \ell - m$) are lost. Let the observed delay values of the $\ell + m$ packets be $\{d_1, d_2, \cdots, d_{\ell+m}\}$, and let $d_{av}(t)$, $d_{sdev}(t)$, $d_{min}(t)$ and $d_{max}(t)$, be the sample mean, standard deviation, minimum and maximum of this set, respectively.

Figures 22(a)-22(d) show the scatter plots of number of out-of-order acknowledgments versus d_{av}, d_{sdev}, d_{min} and d_{max} respectively. These were computed in a manner similar to that in Section 5. We observe that the patterns are similar to those of Figure 17 in Section 5.1.

The following observations from Figure 22, are also similar to the packet-loss observations:

- There is no simple linear relationship between out-of-order acknowledgments and delay. For a given value of number of out-of-order acknowledgment, there is a wide spectrum of possible delay values.
- There is a shift in the *distribution* of number of out-of-order packets as delay increases. To see this, observe that the patterns shift to the right as the number of out-of-sequence acknowledgments increases.

Therefore, as in Section 5, we study the conditional expectation of the fraction of out-of-order acknowledgments given a certain value of delay. Let $R = \ell/n$ be the fraction of acknowledgments that were out-of-order, and let the conditional expectations with respect to $d_{av}, d_{sdev}, d_{min}$ and d_{max} be defined as $E\{R = \ell/n | d_{av}\}$, $E\{R = \ell/n | d_{sdev}\}$,

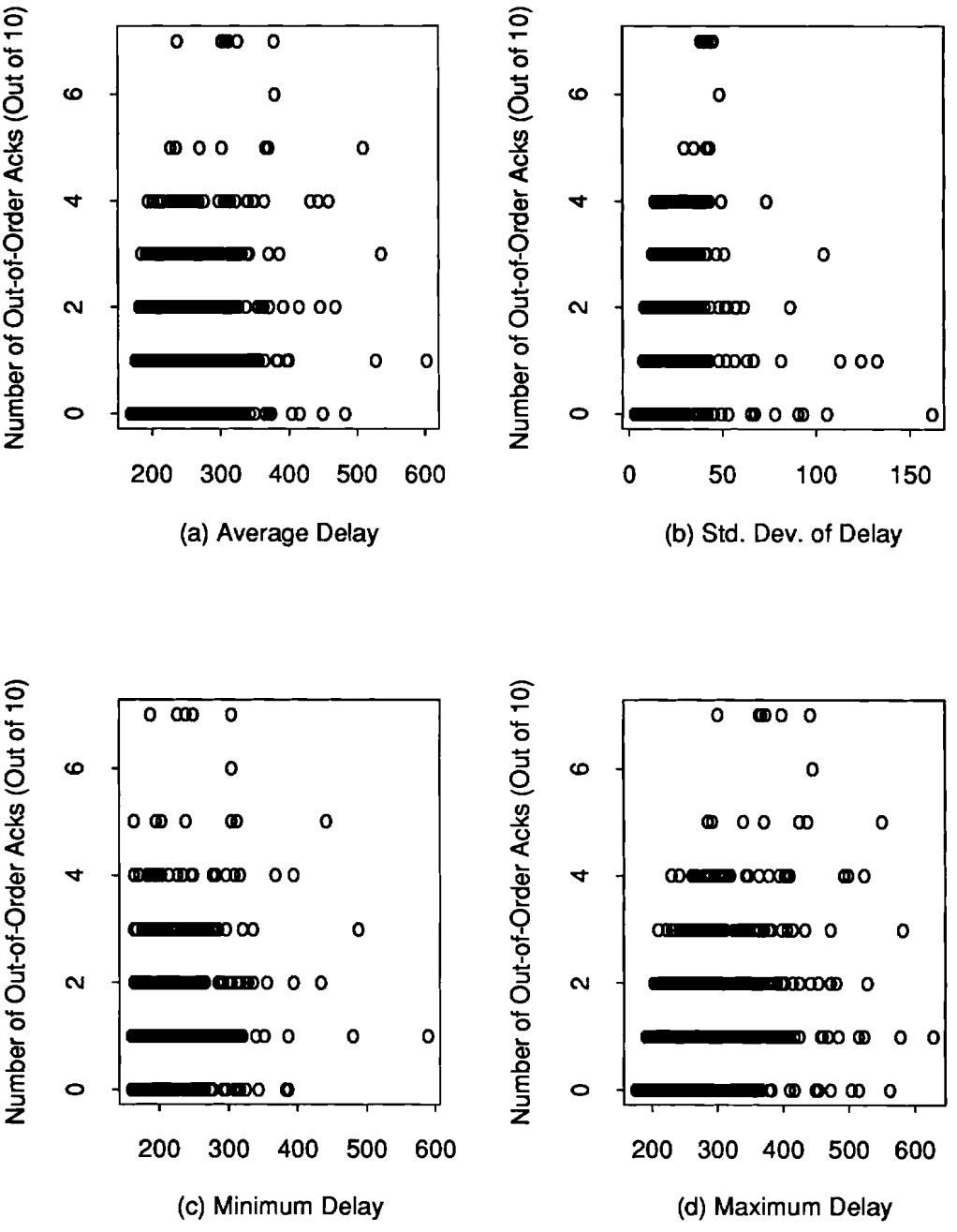


Figure 22: Scatter plot of Out-of-Order Acknowledgments versus Average, Standard Deviation, Minimum and Maximum Delay for the Cross-Country Data.

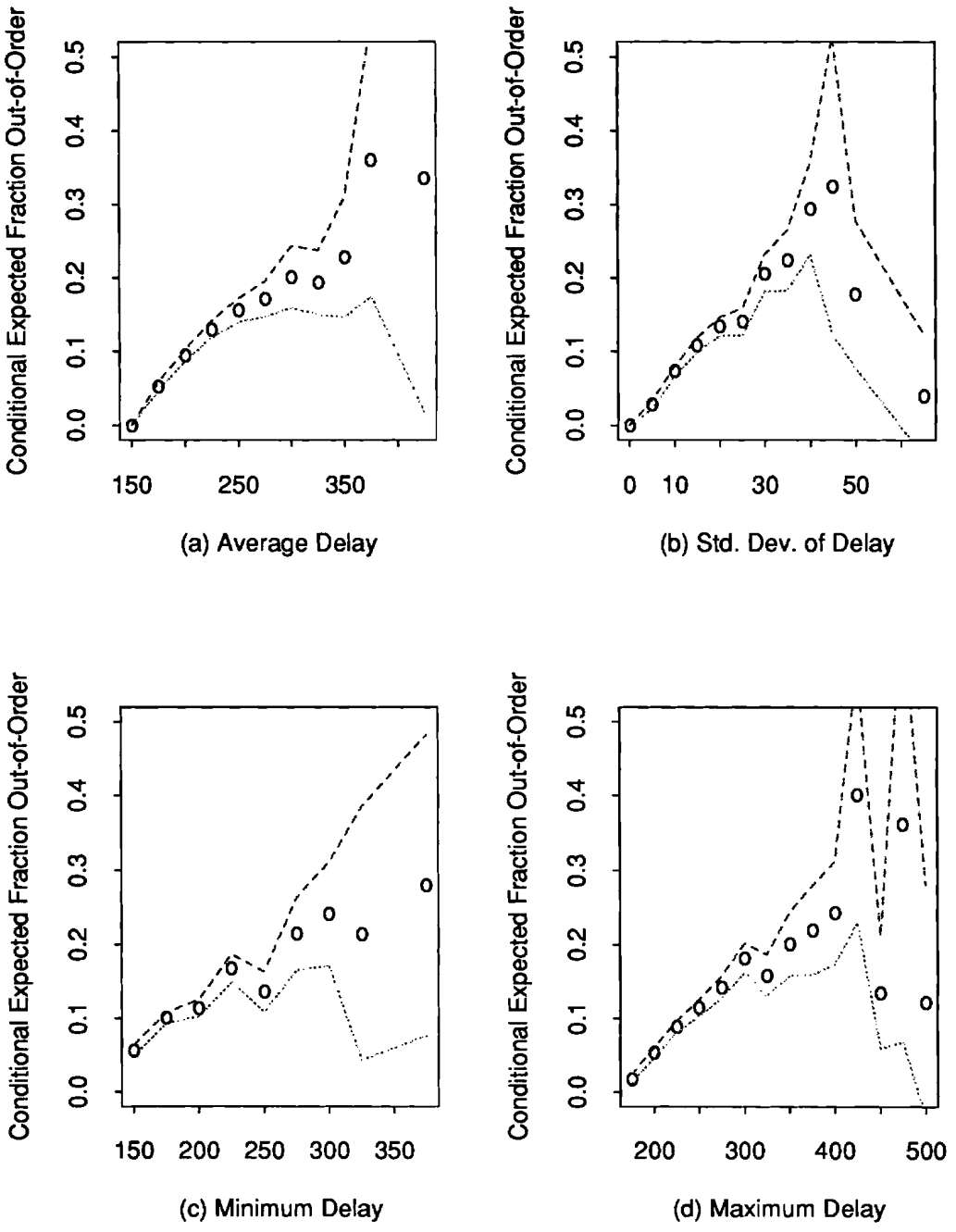


Figure 23: Conditional Expectation of Out-of-Order Acknowledgments versus Average, Standard Deviation, Minimum and Maximum Delay for the Cross-Country Data. Dashed lines indicate the 90% Confidence Intervals. Notice that the confidence intervals are quite tight for low to medium range of the delay spectrum, but get wide at the higher end.

	Average Delay	Std. Dev.	Minimum	Maximum
Out-of-Order Acks	0.3615411	0.3224647	0.2765262	0.3977756

Table 6: Correlations Between Out-of-Order Acknowledgments and Different Statistics of Delay for Cross-Country Data.

$E\{R = \ell/n|d_{min}\}$ and $E\{R = \ell/n|d_{max}\}$, respectively. We wish to study the properties of $E\{R = \ell/n|d_{av}\}$, $E\{R = \ell/n|d_{sdev}\}$, $E\{R = \ell/n|d_{min}\}$ and $E\{R = \ell/n|d_{max}\}$ as a function of $d_{av}, d_{sdev}, d_{min}$ and d_{max}, respectively.

Figures 23(a)-23(d) show the estimated $E\{R|d_{av}\}$, $E\{R|d_{sdev}\}$, $E\{R|d_{min}\}$ and $E\{R|d_{max}\}$ as a function of $d_{av}, d_{sdev}, d_{min}$ and d_{max}, respectively. Also given are the 90% confidence intervals around the conditional expectations. These were all computed by an algorithm similar to (11)-(13) in Section 5.1. Notice that the conditional expectation of out-of-order acknowledgments increases with delay. For smaller values of d_{av}, d_{sdev}, d_{min}, and d_{max}, the expected out-of-sequence acknowledgments is low, but it increases as the delay increases. We believe, this is due to multiple source-destination paths.

We also observe from Figures 23(a)-23(d) that the confidence intervals are reasonably tight for a low to medium range in the delay spectrum, but get large at the high end, implying that there were an insufficient number of sample points at the high end. One way to decrease the uncertainty would be to observe the network for a longer period of time, so that more samples fall in that region.

The correlation coefficients between out-of-order acknowledgments and delays are summarized in Table 6. Again, notice that the correlation values, while not negligible, are quite small, as in the packet-loss case. The discussion of Section 5.3, therefore, also applies here.

7 Comparison of Results Across Time

A comparison of results across different points in time reveals a surprising similarity in the statistical properties of delay and loss, studied so far. For instance, we studied the cross-country segment for seven days during the week of July 27, 1992, and then again for a day on November 10, 1992, and the correlation properties of packet-loss and delay, the distribution of delay, etc. all show similar trends.

A major difference has been in the absolute values of the delays, where a significant improvement has taken place. For instance, for the November 10, 1992 data, the raw numbers range from 84msec to 380msec, whereas previously, it could be anywhere from 170msec to 600msec. We believe this is due to the NSFNet T3 backbone which has replaced the T1 backbone. The number of hops for the cross-country segment also seems to have reduced significantly, so this could be a factor in the improvement as well.

Given that the load on the segment is lower, we find, as we would expect from Section 5, that the absolute values of the correlations between packet loss and delay are slightly lower as well. The difference between them is roughly 0.1.

8 Related Work

This section discusses previous work. Since different studies have been conducted with different goals in mind, and have used different methodologies, we have decided on the following taxonomy: Section 8.1 discusses other round-trip delay studies; Section 8.2 discusses different workload characterization studies for Internet data traffic; Section 8.3 discusses models for data networks that have been used in recent studies; Section 8.4 discusses various experimental and theoretical studies of congestion control. Several studies do not admit to a unique classification. In such cases, we have broken ties somewhat arbitrarily.

8.1 Round Trip Delay Studies

Round trip delay studies have had several applications:

- *Goal: Understanding Internet Workload:* Sanghi et al [35] have recently reported measurements of round-trip delays over the Internet. Their measurement strategies are somewhat different from ours, and has its advantages and disadvantages. (See Section 2.2.) All results and analysis in our paper (e.g., distribution of delay, correlation properties of loss with different statistics of delay, correlation properties of out-of-order packets and delay, dominant frequencies in the data, etc.) are new.
- *Goal: Distributed Time Synchronization:* Mills [24] has developed the Network Time Protocol for time synchronization across machines separated by non-negligible delays. The goal of this study was to accurately estimate a fixed relative clock offset between a pair of machines in spite of random communication delays. The Network Time Protocol addresses randomness in round-trip delays by (a) exchanging a set of (time-stamped) messages that yield an estimated (offset, delay) tuple, and (b) by synchronizing clock offsets based on a minimum-delay filter. Mills shows that this leads to the minimum error in clock offsets.

  Mills has also reported on Internet Delays as a function of packet size etc. [23].
- *Goal: Timer Adjustment:* Several papers have addressed timer adjustment in TCP [16, 19, 39]. The first two papers model round-trip delays as a first order Auto-Regressive Moving Average process. The third paper shows some of the difficulties of timer adjustment in TCP.
- *Goal: Report On NSFNet Backbone Statistics:* Every month, the National Science Foundation, in conjunction with Merit Corporation, publishes statistics (minimum, average and maximum) of pairwise delays between backbone nodes (Network Switching Systems) on the Internet.

8.2 Workload Characterization Over the Internet

Workload characterization studies over the Internet fall into three categories, conversation level, packet level and NSFNet backbone level.

- *Conversation Level Studies:* Caceres et al [3] and Paxson [31] have studied the properties of TCP/IP conversations from packet traces collected at several campus gateways. Schmidt and Campbell [36] have reported statistics on conversations and IP datagrams from data collected at NSFNet and CSOnet gateways.

These studies give individual conversation characteristics and conversation arrival rates at the gateways.

From the perspective of running simulations for congestion control, these studies are invaluable. For modeling of large networks, however, the conversation arrival rates may need further exploration because the round-trip delays (in Section 2.3) show considerably different characteristics from regional, to backbone to cross-country networks. Individual conversations properties most likely will remain the same.

- *Packet Level Studies/LAN/Gateway:* Leland and Wilson [21] and Fowler and Leland [10] studied packet arrival statistics at the BellCore gateway. Interestingly, they too reported presence of low frequency fluctuations in in average packet arrival rates at their gateway (see for instance, Figure 1, pp. 1142 in [10]). The round-trip delays in Section 3 reinforce this observation and suggests that path-congestions too have slow build-up and clear-out components. It ought to be possible to exploit this information for heuristic congestion control and adaptive routing algorithms.

  Erramilli and Forys [6] have reported large queueing delays in switches in spite of moderate utilizations. They ascribe it to traffic synchronization which is an undesirable batching of workload caused by interaction between control and data traffic. The large tails that we observed in the delay-distributions may in fact be due to some form of traffic synchronization that we do not completely understand. We think this is an issue that needs attention.

  In other measurements, Jain and Routhier [18] reported that successive packets on a Local Area Network belonged to the same end-to-end transmission entity. Moghul [28] has reported existence of similar localities at the level of processes. Feldmeier [7] has used locality of traffic for routing caches in gateways. Gussella [13] has studied local area network traffic and has reported on packet inter-arrival and size distributions. Shoch and Hupp's landmark paper [37] is perhaps one of the earliest papers on traffic measurements over a Local Area Network.

- *Traffic Characteristics of the T1 NSFNet BackBone:* A recent study by Claffy et al [4] has reported on the traffic characteristics over the NSFNet backbone. The study gives long-term growth of traffic volume, packet size statistics for different protocols, distribution of median delays across different backbone nodes, link utilization statistics, and link error and nodal downtime statistics.

8.3 Models for Network Traffic and Round Trip Delays (for Congestion Control Studies)

- *Models for a single conversation:* A simple fluid model explaining rate regulation in window flow control was proposed by Jacobson [16]. The key observation was that, under certain assumptions, successive packets transmitted by a sender are spaced apart by an amount equal to the bottleneck's processing time. (The assumptions included the sender using window flow control, and being in steady state, i.e., in a state where it was able to transmit a full window's worth of packets.) Moghul [29] and Zhang et al [41] have reported that there is a potential for returning acknowledgments getting batched together in the presence of interfering traffic. This phenomenon is similar in spirit to the study by Erramilli and Forys [6] mentioned earlier, but differ in details. Keshav [20] has proposed a flow con-

trol algorithm based on the Jacobson model, but refined to estimate a random bottleneck capacity.

We had originally planned on experimenting with these ideas using round-trip delays, but out-of-order return of acknowledgments precluded such experiments.

- *Effect of propagation delay on simulation results:* Floyd and Jacobson [8] have observed that, presence of propagation delays in dynamic window flow control simulations may give unexpected results. They showed that periodicity in transmissions (caused by window flow control) and deterministic round-trip delays caused certain transmitters to win out over others. The phenomenon causing this was referred to as phase-effect.

  While clearly a problem that needs to be addressed in simulations, our experiments show that it may be of less concern in real network operation. This is because the delays observed in our experiments show enough variability which should break the systematic pattern causing phase effect. However, as network speeds go up, this conclusion may need to be re-evaluated, assuming window flow control is continued to be used.

- *Delay based heuristics for congestion control:* Several researchers, e.g., [17], have proposed using round-trip delays for end-point flow control, a proposal especially attractive in heterogeneous networks because it is free of explicit signalling mechanisms.

  Our results indicate that while conditional expected loss increases with delay (Figures 18), the correlation between delay and loss is low, and is going to be that way, for low loss networks. Any such scheme can, therefore, at best be a congestion-avoidance algorithm. Care must be taken to ensure that transmitters are not overly pessimistic.

8.4 Congestion Control: Analysis and Simulation

Several performance studies on congestion control algorithms have been reported recently [2, 11, 26, 25, 30, 38, 40]. Several of them have reported that feedback delay plays an important role in stability and fairness.

Most authors, however, have assumed a fixed (deterministic) delay because the actual distribution was not known. One contribution of this paper is an empirical observation on the distribution of delay — that it is a Gamma (Section 4). Whether or not the distribution is provably a Gamma is an open question, but there is strong empirical evidence that it is at least approximately so.

It will be interesting to investigate if there are algorithms that will guarantee (or bound) oscillations and ensure fairness (perhaps, in some weak form?), given this delay distribution.

9 Concluding Remarks

We have presented a detailed study of low frequency components of Internet end-to-end delays, their distribution, their relationship with loss, and their relationship with out-of-sequence acknowledgments. The key results are summarized in Table 1 of Section 1.

Instead of restating the results, we take a look at some of the open questions. The observations in the paper indicate that there is a real need for further investigation in, and a clear understanding of, flow control protocols. For instance, consider the following issues (this list is by no means exhaustive; the reader is welcome to fill in his/her own):

- *Incorporation of low frequency components in simulations:* Most simulation studies assume a small (usually fixed) number of conversations. Conversation arrivals and departures must be incorporated into the simulations in order to get a true picture of how flow control protocols would perform in practice.
- *Need for a clear understanding of how to adjust flow control parameters as a function of load:* Consider, for instance, a leaky bucket scheme for flow control in high speed networks. Given that there is slowly varying component in the workload, how exactly should the leaky-bucket parameters (bucket size, token flow rate) be adjusted, and how often do they need to be adjusted? Similarly, for a time-window scheme, how exactly should the time-window parameters be adjusted? The linear-increase/exponential-decrease algorithms [9, 16, 33] or the algorithms proposed in [11, 25] are some possibilities, but can one do better?
- *Need for better heuristics for congestion avoidance:* As indicated in Section 5, the correlations between loss and different statistics of delay are low; the exact values depend on network load. Further, there is no one single winner among d_{av}, d_{sdev}, d_{min} and d_{max}, in terms of being able to signal congestion best all of the time. An important question then is, does there exist a function that would enhance the magnitude of the correlations? Further, given that there will be a demand for lower network losses in the future, implying even lower correlations, would such a strategy still work? Or must there be explicit signalling protocols and/or reservation based schemes for congestion control?

  A related open question is, can one use the information that round-trip delays are approximately Gamma-distributed to one's advantage? For example, can one make flow control choices based on parameter estimates of the Gamma distribution?
- *Need for prevention of large tails in the delay distribution:* The delay distributions in Section 4 have a long tail. Even though the mass at the tail is small, it is not negligible. What is the reason for this, and how should one prevent it? For instance, are the long tails due to First-Come-First-Served scheduling? Or are there some phenomena out there that we do not completely understand (e.g, traffic synchronization [6])?
- *Need for realistic assumptions in flow control studies:* Examples of this are incorporation of delay distributions in analytical studies and incorporation of out-of-sequence acknowledgments in the development of heuristics for flow control.

With that, we conclude the body of this paper.

Acknowledgments

We would like to thank several people who were of help to us in the preparation of this paper. Sabyasachi Basu and Albert Broscius provided valuable input during the early stages of this research. Brendan Becker, Charles Catlett and Joel Replogle provided us with accounts on their machines for collecting the data. Mark Foster and Ira Winston, in our systems lab, were extremely helpful and patient with us during the experiments. Chuck Kalmanek, Srinivasan Keshav and Biswanath Mukherjee provided

valuable feedback. We would also like to extend our special thanks to Richard Becker, John Chambers and Allan Wilks for the S programming language, Alfred Aho, Brian Kernighan and Peter Weinberger for the Awk programming language, and the people who created the Shell programming environment on UNIX. Their environments made this work a pleasure rather than a chore. The responsibility for the contents of the paper, of course, remains with the author.

A Relationship between quantile-quantile plot of X and Z from Section 4.

Lemma 2 *Let Z be a unit-scale gamma-distributed random variable and X be the observed random variable. Suppose that the quantile-quantile plot of the CDF of X and Z is a straight line with slope m and intercept c. Then $X - c$ is a gamma-distributed random variable with scale m and shape s.*

Proof:

By hypothesis, the quantile-quantile plot of X and Z is a straight line with slope m and intercept c. Define the random variable

$$Y = \frac{X - c}{m}. \tag{22}$$

Then, the quantile-quantile plot of Y versus Z should be a straight line with unit slope and zero intercept. I.e.,

$$f_{s,Y}(y) = \frac{y^{s-1}e^{-y}}{\Gamma(s)}.$$

Therefore,

$$Pr\{y \leq Y \leq y + \Delta y\} = \frac{y^{s-1}e^{-y}}{\Gamma(s)}\Delta y \tag{23}$$

$$\Rightarrow Pr\{y \leq \frac{X - c}{m} \leq y + \Delta y\} = \frac{y^{s-1}e^{-y}}{\Gamma(s)}\Delta y. \tag{24}$$

The left hand side may be simplified to:

$$Pr\{my + c \leq X \leq my + c + m\Delta y\} = Pr\{x \leq X \leq x + \Delta x\}. \tag{25}$$

Using (22), the right hand side of (24) becomes

$$\frac{\left(\frac{x-c}{m}\right)^{s-1} e^{-\frac{x-c}{m}}}{\Gamma(s)} \frac{\Delta x}{m} \tag{26}$$

From (25) and (26), the density function of X is given by

$$f_{s,m,X}(x) = \frac{\left(\frac{x-c}{m}\right)^{s-1} e^{-\frac{x-c}{m}}}{m\Gamma(s)}. \tag{27}$$

Now, let $X' = X - c$. Substituting in (27), we have

$$f_{s,m,X'}(x') = \frac{\left(\frac{x'}{m}\right)^{s-1} e^{-\frac{x'}{m}}}{m\Gamma(s)}. \tag{28}$$

I.e., X' is a standard Gamma-distributed random variable with scale parameter m and shape parameter s. Since $X = X' + c$, where c is a constant, X has the same distribution as X' but shifted by the constant c. ∎

B Proof of Footnote in Section 3

From the data in Figure 3(b), it might appear that a phase angle may have to be carried along explicitly in the regression in Section 3, i.e., $d(t)$ must be of the form:

$$d(t) = \sum_{f_i \in \mathcal{F}} \left[a_i \cos\left(\frac{2\pi f_i t}{N} + \phi_i\right) + b_i \sin\left(\frac{2\pi f_i t}{N} + \phi_i\right)\right] + \epsilon(t), \tag{29}$$

where ϕ_i is the phase angle corresponding to f_i. However, from a regression standpoint, ϕ_i is irrelevant, because for every i, it can be shown that

$$a_i \cos\left(\frac{2\pi f_i t}{N} + \phi_i\right) + b_i \sin\left(\frac{2\pi f_i t}{N} + \phi_i\right) = a'_i \cos\left(\frac{2\pi f_i t}{N}\right) + b'_i \sin\left(\frac{2\pi f_i t}{N}\right),$$

for some a'_i and b'_i.

Proof:

$$\begin{aligned}
& a_i \cos\left(\frac{2\pi f_i t}{N} + \phi_i\right) + b_i \sin\left(\frac{2\pi f_i t}{N} + \phi_i\right) \\
&= a_i \left[\cos\left(\frac{2\pi f_i t}{N}\right)\cos(\phi_i) + \sin\left(\frac{2\pi f_i t}{N}\right)\sin(\phi_i)\right] + b_i \left[\sin\left(\frac{2\pi f_i t}{N}\right)\cos(\phi_i) + \cos\left(\frac{2\pi f_i t}{N}\right)\sin(\phi_i)\right] \\
&= [a_i \cos(\phi_i) + b_i \sin(\phi_i)] \cos\left(\frac{2\pi f_i t}{N}\right) + [b_i \cos(\phi_i) + a_i \sin(\phi_i)] \sin\left(\frac{2\pi f_i t}{N}\right) \\
&= a'_i \cos\left(\frac{2\pi f_i t}{N}\right) + b'_i \sin\left(\frac{2\pi f_i t}{N}\right),
\end{aligned}$$

for some a'_i and b'_i. I.e., the need for a phase angle is obviated by the form (8).

∎

C Proof of Lemma 1 from Section 5.3

Lemma 1 : *The correlation between D and L is given by*

$$\mathrm{cor}(D, L) = \frac{a\,\sigma^2(L_d)}{\sqrt{a^2\sigma^2(L_d) + \sigma^2(D_{\bar{\ell}})}\sqrt{\sigma^2(L_d) + \sigma^2(L_{\bar{d}})}}. \tag{30}$$

Proof:

From (14) and (16),

$$D = aL_d + b + D_{\bar{\ell}}. \tag{31}$$

Now,

$$\begin{aligned}
\mathrm{cor}(D, L) &= \frac{E[DL] - ED\,EL}{\sigma_D \sigma_L} \\
&= \frac{E\left[(aL_d + b + D_{\bar{\ell}})(L_d + L_{\bar{d}})\right] - E[(aL_d + b + D_{\bar{\ell}}]\,E[L_d + L_{\bar{d}}]}{\sqrt{a^2\sigma^2(L_d) + \sigma^2(D_{\bar{\ell}})}\sqrt{\sigma^2(L_d) + \sigma^2(L_{\bar{d}})}}
\end{aligned} \tag{32}$$

Simplifying the numerator,

$$
\begin{aligned}
& aE[{L_d}^2] + aE[L_d L_{\bar{d}}] + b(E[L_d] + E[L_{\bar{d}}]) + E[D_{\bar{\ell}} L_d] + E[D_{\bar{\ell}} L_{\bar{d}}] \\
& \qquad -E[(aL_d + b + D_{\bar{\ell}})\, E[L_d + L_{\bar{d}}] \qquad (33) \\
= \;& aE[{L_d}^2] + aE[L_d]E[L_{\bar{d}}] + b(E[L_d] + E[L_{\bar{d}}]) + E[D_{\bar{\ell}}]E[L_d] + E[D_{\bar{\ell}}]E[L_{\bar{d}}] \\
& \qquad -\left\{aE^2[L_d] + aE[L_d]E[L_{\bar{d}}] + b(E[L_d] + E[L_{\bar{d}}]) + E[D_{\bar{\ell}}]E[L_d] + E[D_{\bar{\ell}}]E[L_{\bar{d}}]\right\} \qquad (34) \\
= \;& a\left(E[{L_d}^2] - E^2[L_d]\right) \\
= \;& a\sigma^2(L_d) \qquad (35)
\end{aligned}
$$

(34) follows from (33) by the independence of L_d and $L_{\bar{d}}$, $D_{\bar{\ell}}$ and L_d, and $D_{\bar{\ell}}$ and $L_{\bar{d}}$. From (32 and (35),

$$
\mathrm{cor}(D, L) = \frac{a\,\sigma^2(L_d)}{\sqrt{a^2\sigma^2(L_d) + \sigma^2(D_{\bar{\ell}})}\sqrt{\sigma^2(L_d) + \sigma^2(L_{\bar{d}})}}. \qquad (36)
$$

∎

References

[1] Becker, R.A., J.M. Chambers and A.R. Wilks, "The new S language, a programming environment for data analysis and graphics," *Wadsworth and Brooks/Cole*, 1988.

[2] Bolot, J., and A. U. Shankar, "Dynamical behavior of rate-based flow control mechanism," *Computer Communication Review,* pp. 35-49, April 1990.

[3] Caceres,R., P.B.Danzig, S.Jamin and D.J.Mitzel, "Characteristics of wide-area TCP/IP conversations," *Proc of ACM Sigcomm,* pp 101-112, Zurich, Sept 1991.

[4] Claffy, K.C., G.C. Polyzos and H.W. Braun, "Traffic characteristics of the T1 NSFNet backbone," *Proc. of the IEEE Infocom,* to appear.

[5] Demers, A., S. Keshav and S. Shenker, "Analysis and simulation of a fair queueing algorithm," *Proc. ACM Sigcomm,* pp 1-12, Sept 1989.

[6] Erramilli, A. and L.J. Forys, "Oscillations and chaos in a flow model of a switching system," *IEEE Journal on Selected Areas in Communications,* pp 171-178, February, 1991.

[7] Feldmeier, D. "Improving gateway performance with a routing-table cache," *Proc. IEEE Infocom,* March 1988.

[8] Floyd, S. and V. Jacobson, "On traffic phase effects in packet-switched gateways," *Journal of Internetworking: Research an Experience,* Vol. 3, pp. 115-156, 1992.

[9] Faber, T., L.H. Landweber and A. Mukherjee, "Dynamic time windows: packet admission control with feedback," *Proc. ACM Sigcomm,* pp 124-135, Baltimore, August 1992.

[10] Fowler, H.J., and W.E. Leland, "Local area network traffic characteristics, with implications for broadband network congestion management," *IEEE Journal on Selected Areas in Communications,* pp 1139-1149, September, 1991.

[11] Fendick, K., M. Rodriguez and A. Weiss, "Analysis of a rate based congestion control strategy with delayed feedback," *Proc. ACM Sigcomm,* pp 136-148, Baltimore, August 1992.

[12] Golestani, S. J., "A Stop-and-Go queueing framework for congestion management," *Proc of ACM Sigcomm,* pp 8-18, September 1991.

[13] Gusella, R., "Characterizing the variability of arrival processes with indexes of dispersion," *IEEE Journal on Selected Areas in Communications*, pp. 203-211, February 1991.

[14] Hahne, E.L., C. R. Kalmanek, S. P. Morgan, "Fairness and congestion control on a large ATM data network with dynamically adjustable windows," ITC 13, Copenhagen, June 1991.

[15] Hui, J.Y., "Resource allocation for broadband networks," *IEEE Journal on Selected Areas in Communications,* pp 1598-1608, Dec 1988.

[16] Jacobson, V., "Congestion avoidance and control," *Proc. of the ACM Sigcomm,* pp 314-329, Aug. 1988.

[17] Jain, R., "A delay-based approach for congestion avoidance in interconnected heterogeneous computer networks," *Computer Communication Review,* pp. 56-71, 1989.

[18] Jain, R., and S. Routhier, "Packet Trains: Measurements and a New Model for Computer Network Traffic," Tech Report MIT/LCSTM 292, Department of Electrical Engg, MIT, Nov 1985.

[19] Karn, P. and C. Partridge, "Improving round-trip estimates in reliable transport protocols," *ACM Trans. Comp. Systems,* pp 364-373, Nov. 1991.

[20] Keshav, S., "A control-theoretic approach to flow control," *Proc. of the ACM Sigcomm,* Zurich, pp. 3-15, September 1991.

[21] Leland, W.E. and D. Wilson, "High time resolution measurement and analysis of Lan traffic: implications for Lan interconnection," *Proc. IEEE Infocom,* 1991.

[22] Mishra, P., and H. Kanakia "A hop-by-hop rate-based congestion control scheme," *Proc. ACM Sigcomm,* Baltimore, pp. 112-123, August 1992.

[23] Mills, D.L., "Internet delay experiments," Request for Comment: RFC-889, Network Information Center, SRI International, December 1983.

[24] Mills, D.L., "Internet time synchronization: the network time protocol," *IEEE Trans. Commun.,* pp. 1482-1493, October 1991.

[25] Mitra, D. and J. Seery, "Dynamic adaptive windows for high speed data networks: theory and simulations," *Proc. ACM Sigcomm,* pp 30-37, Philadelphia, September 1990.

[26] Mitra, D., "Asymptotically optimal design of congestion control for high speed data networks," *IEEE Trans. Commun.,* pp. 301-311, February 1992.

[27] Mukherjee, A., L.H. Landweber and T. Faber, "Dynamic time windows and generalized virtual clock: combined closed-loop/open-loop congestion control," *Proc. IEEE Infocom,* Florence, pp 322-332, May 1992.

[28] Moghul, J.C., "Network locality at the scale of processes," *Proc. ACM Sigcomm,* Zurich, pp. 273-284, September 1991.

[29] Moghul, J.C., "Observing TCP dynamics in real networks," *Proc. ACM Sigcomm,* Baltimore, pp. 305-317, August 1991.

[30] Mukherjee, A. and J.C. Strikwerda, "Analysis of dynamic congestion control protocols — a Fokker-Planck approximation," *Proc. ACM Sigcomm,* Zurich, pp. 159-169, September 1991.

[31] Paxson, V., "Measurements and models of wide area TCP conversations," Technical Report LBL-30840, Lawrence Berkeley Laboratory, University of California at Berkeley, May 1991.

[32] Ramamurthy, G., and R.S. Dighe, "Distributed source control: a network access control for integrated broadband packet networks," *IEEE Journal on Selected Areas in Communications,* pp. 990-1002, September 1991.

[33] Ramakrishnan, K.K., and R. Jain, "A Binary Feedback Scheme for Congestion Avoidance in Computer Networks with a Connectionless Network Layer," *ACM Trans. Computer Systems,* vol. 8, pp 158-181, May 1990.

[34] Ramamurthy, G., and B. Sengupta, "A predictive congestion control policy for high-speed wide-area networks," Technical Report 92-C007-4-5011-1, NEC Research Laboratories, Princeton, May 1992.

[35] Sanghi, D., A. Agrawala, O. Gudmundsson and B. Jain, "Experimental assessment of end-to-end behavior on Internet," *Proc. of the IEEE Infocom,* 1993, (to appear).

[36] Schmidt, A. and R. Campbell, 'Internet protocol traffic analysis with applications for ATM switch design," Technical Report TR92-1735, Dept. of Computer Science, University of Illinois, 1992.

[37] Shoch, J.F., and J.A. Hupp, "Measured performance of an Ethernet local network, "*Commun. ACM,* pp. 711-721, December 1980.

[38] Wang, Y.T., and B. Sengupta, "Performance analysis of a feedback congestion control policy under non-negligible propagation delay," *Proc of ACM Sigcomm,* pp. 149-157, Zurich, September 1991.

[39]

[40] Zhang, L., "A New Architecture for Packet Switching Network Protocols," Ph.D. thesis, MIT, Dept. of Electrical Engg. and Computer Science, Aug. 1989.

[41] Zhang, L., S. Shenker and D.D. Clark, "Observations on the dynamics of a congestion control algorithm: the effects of two-way traffic," *Proc. ACM Sigcomm,* Zurich, pp. 133-147, September 1991.