On the Long-range Dependence of Packet Round-trip Delays in Internet

Qiong Li David L. Mills*
Department of Electrical and Computer Engineering, University of Delaware

Abstract

Many recent studies on network traffic demonstrate that Long-range Dependence (LRD) is a ubiquitous property of traffic both in a local area network (LAN) and a wide area network (WAN). So the performance of network should be dominated by this property. Packet Round-trip Delay is an important measurement of network performance. In this paper, we present evidence that LRD exists in packet round-trip delays. This discovery has serious implications for understanding the impact on network performance of LRD network traffic, and design of transport control protocols for special applications, i.e., of teleconferencing. Statistical analyses show that the complementary probability distribution of packet round-trip delays decays more slowly than exponential rate; this fact probably justifies the studies on the prediction of the queue length distribution with LRD network traffic. We also tentatively use a multi-queueing system to interpret the existence of LRD in the packet round-trip delay process, which we believe is caused by LRD of Internet traffic.

1 Introduction

In recent years, empirical studies [6, 10, 7] on network traffic both in local area networks (LAN) and wide area networks (WAN) convincingly show that the properties of actual traffic are very different from that predicted by traditional teletraffic models, such as Poisson process. For actually measured traffic, the correlation in traffic can extend to a wide range of different time scales, or mathematically, the correlation function of realistic traffic decays with lag time in the way of power-law, which is the property of so-called long-range dependence (LRD); while for traditional model-generated traffic, its correlation function decays exponentially fast, namely, short-range dependence. An obviously visible physical phenomenon watched in measured traffic is burst, which appearing in the counting processes of traffic is as packets tend to come in clusters. Due to LRD, the burst phenomenon cannot be smoothed out by simply aggregating the traffic in a larger time scale, which means even using very large time units to construct the counting processes, the burst phenomenon can still be observed. This demonstrates that the underlying process of measured traffic is self-similar. Artificial traffic generated by traditional models, on micro structure level, is not very distinguishable in vision from realistic traffic. The aggregating processes of the artificial traffic with increasing time scales, however, will lose the visible burst phenomenon rapidly. So, the burst phenomenon, which keeping in all time scales of the aggregated counting process is called self-similarity, exclusively exists only in measured traffic, and cannot be predicted by traditional models.

One question arose with the finding of LRD (or specially say, self-similarity) in realistic traffic is, what the physical mechanism is, which causes the unique property in modern high speed network traffic. One very recent study in [14, 13] gave an excellent physical explanation of the reason why LRD exists in LAN traffic flow. In this study it shows that, on the source level (distinguished from the link level), by modeling an individual source or source-destination pair as a strictly alternating ON/OFF source (an ON period followed exactly by an OFF period) with heavy-tailed probability distribution of the lengths of both ON and OFF periods, we can get a self-similar aggregated traffic on the link level. This study relating self-similarity of aggregated traffic to individual sources or source-destination pairs provides a deep insight into the physical mechanism of LRD. The key factor claimed in the study, which is responsible to the self-similarity, is the heavy-tailed probability distribution of the lengths of both ON and OFF periods. It is interesting to notice that heavy-tailed probability distribution, such as the Parato distribution, is a prevailing statistical phenomenon found in many aspects related to network traffic [10, 3, 1], such as interarrival times of packets, burst sizes in bytes, file sizes in World Wide Web (WWW), and Unix file systems. In the aforementioned ON/OFF model, an ON period corresponding to a single transmission session time is directly related to the size of the transmitted message. The distribution of the size of the transmitted message is likely to be determined by that of the sizes of available files, which in most cases exhibit a heavy-tailed distribution. So it is reasonable and understandable to assume the distribution of ON periods as a heavy-tailed one. The OFF period is corresponding to the silent period of a source, which is related to factors determined by human behavior and applications' protocols. It is still necessary to do more work to confirm the reasonability of the heavy-tailed assumption of the OFF period. But our primitive research work on human behavior already shows a trace of heavy-tailed property.

It seems that the resistance to accepting the existence of LRD in high speed network traffic has been removed. More studies on networking engineering now focus on the influence of LRD in traffic on the performance of networking [4, 5, 9]. These experimental and simulation studies show that LRD, as well as the marginal distribution of a traffic arrival process, has significant effects on all metrics of networking performance, including throughput, packet loss rate, response time, and buffer occupancy. The main reason why LRD affects all the aspects of networking performance is that

*This work is supported by U. S. Army Research Laboratories Cooperative Agreement DAAL01-96-2-0002, DARPA Information Technology Office Contract DABT 63-95-C-0046, NSF Grant NCR 93-01002, and U. S. Navy Contract A303 276-93.

the queueing performance with LRD packet traffic is seriously degraded. Numerical and analytical studies based on models catching LRD property of traffic demonstrate that the tail of a queue length distribution decays much more slowly than the exponential rate; this implies that a packet tends to experience longer delay within networks on average before it gets to its destination than that predicted by traditional models, such as the Markov model. The conclusions drawn from these studies have remarkable implications for the design, control of networks, and tuning of protocols.

Internet is expanding dramatically fast. It is the most complicated collection of networks connected together. Many well-developed and currently developing applications run across Internet to go around the world. Some applications (audio, video) are sensitive to the performance of the whole Internet, or precisely say, the packet delay in Internet. TCP/IP is the most widely used protocol suite. With TCP there are several timeout parameters needed to be carefully tuned, the most sensitive one is the retransmission timeout (RTO) parameter determined dynamically by packet round-trip delay in Internet. The RTO used to indicate when a packet can be assumed lost in the network by a sender and consequently invoke a retransmission event, will dramatically affect the end-to-end behavior of TCP/IP protocol [11]. No doubt, it is necessary to thoroughly understand the characteristics of the packet round-trip delay process (defined in section 2) in Internet, which is the most important metric for evaluating the performance of Internet.

However, there are few studies on the packet round-trip delay process in literature so far. Especially, under current situation with LRD packet traffic, the characteristics of the packet round-trip delay have not been understood very well. In this paper, we carry out an experimental study on the packet round-trip delay process by extensively analyzing the measurements of packet time stamps collected by the Network Time Protocol (NTP) control message packet. Due to the limitation of the current experiment, we focus on the long term behavior of the packet round-trip delay process. The short term behavior of the packet round-trip delay process will be our future topic. The main contribution in this paper is that it presents convincing evidence indicating that the packet round-trip delay process in Internet is a LRD process. If we extend the self-similar model to the packet round-trip delay process, we find that packet going through different routes in Internet will have a different Hurst parameter corresponding to the packet round-trip delay process. Statistical analyses show that the distribution of packet round-trip delay decays more slowly than the exponential rate. If we view the packet round-trip delay process as a tandem of queueing processes along its route, then this fact justifies studies on the prediction of the distribution of queue length with LRD arrival packet traffic.

The rest of the paper is organized as follows: In section 2, we briefly introduce the measurement of packet round-trip delay in our study, and the definition of the packet round-trip process. In section 3, we present the results of statistic analyses on the measured data, show the evidence of LRD in packet round-trip delay process, and the distribution of packet round-trip delay. Section 4 gives a tentative explanation of why the packet round-trip delay process is an LRD process. Finally, in section 5 we conclude this paper, and discuss future work on the topic of the packet round-trip delay process.

2 Measurement and Definition of Packet Round-trip Delay Process

2.1 Measurement of Packet Delay

The calculation of packet delay needs four timestamps, namely, T_1, T_2, T_3, and T_4. All these timestamps are measured by NTP control message packet. When a computer (usually, a time server) sends out a NTP packet, it records the leaving time (T_1) on the outgoing packet. When the packet gets to the peer, the peer records the arrival time (T_2) on the packet too; then the peer passes back the NTP packet and records the leaving time (T_3) on the back packet. When the packet gets back to the sender, the sender records down the receiving time (T_4). So, a packet round-trip delay can be calculated as

$$T = (T_4 - T_1) - (T_3 - T_2)$$

A time server in the Internetworking Research Laboratory at University of Delaware, runs NTP and provides time service to local networks all the time; it also has more than 20 cooperating peers around the whole world. The time server exchanges NTP packets with its peers regularly. For different peers, the exchanged NTP packets go through different routes in Internet. So far, the server has collected the timestamps of all NTP packets it sent out and received since June 1st, 1996. For our purpose in this study we concentrate our analysis on a set of data collected between September and November 1996, by the time server. The set of data is shown in Table 1.

The set of data we choose includes five different routes in Internet, two of them are within the United States, The three others go outside the United States to Sweden, Australia, and Chile respectively. We choose the special measurement periods simply because all the peers and the time server in out lab run continually in these periods; no shutdown and reboot, or other abruptly events happened. So, the data collected in these periods are perfectly continuous, and there are no interruptions caused by computer clocks' losing synchronization. The time server tries to exchange NTP packets with its peers regularly; however, the regularity is broken sometimes by losing of NTP packets in Internet. Fortunately, losing the NTP packet is a rare event, so the degradation of the regularity does not essentially affect our conclusion.

2.2 Definition of Packet Round-trip Delay Process

A packet round-trip delay is the sum of delays on each subnet link traversed by the packet. Each link (or hop) delay in turn consists of four components, including processing delay, queueing delay, transmission delay and propagation delay. Fixed the packet length and the route, the packet round-trip delay only changes with the queue delay, which in Internet is changed with the fluctuation of traffic. So, for a fixed route in Internet, a packet round-trip delay $T(t)$ is a random variable at time t. $T(t)$ describes the process of a packet round-trip delay.

Peer Address	Location	Measurement Period	Total Number of Packet	Maximum Delay (second)	Minimum Delay (second)	Average Interval (second)
192.5.41.40	USA	Sept., 1996	38146	0.5986	0.0137	67.95
192.43.244.18	USA	Sept., 1996	32425	7.3313	0.0564	79.94
192.67.12.101	Australia (AU)	Sept., 1996	35044	13.060	0.2761	73.96
192.36.143.150	Sweden (SE)	Sept., 1996	53427	6.1593	0.1029	48.51
149.83.8.200	Chile (CL)	Sept., 1996	11046	5.1299	0.5777	234.66

Table 1: Qualitative description of the set of data used in the analysis in section 3

In this paper, our goal is to investigate the characteristics of a time series T_i obtained by discretizing $T(t)$ with t. Simply, T_i is a sample process of $T(t)$ at $t = t_i$, where $i = 1, 2, 3, \ldots$. In our case, t_i is the time at witch the NTP packet leaves the time server. With each NTP packet we can get a sample of $T(t)$ for the route between its source and destination at its leaving time. In the set of data we used, the sampling time t_i for each route is not distributed along t strictly in even interval, the reason just as mentioned above. In order to simplify the question, we assume t_i is evenly distributed along t with average interval calculated from original t_i series.

3 Statistic Analysis of Packet Round-trip Delay Process

Various techniques for estimating the intensity of long-range dependence in a time series are available [12]. We choose a simple, but effective one to do the estimation, which is called variance-time plot method (or aggregated variance method). The variance-time plot is obtained by plotting $\log\{var[T^{(m)}]\}$ against $\log(m)$, where for each $m = 1, 2, \ldots$, dividing the original time series $T = \{T_i, i >= 1\}$ into blocks of size m, and averaging within each block, we can get the aggregated process $T^{(m)} = \{T^{(m)}(k)\}$, $k = 1, 2, \ldots$, is the index of blocks. Then, if T has (short-range or) no dependence

$$\mathrm{Var}[T^{(m)}] \sim m^{-1}, \quad \text{as} \quad m \longrightarrow \infty \tag{1}$$

the slope obtained from the variance-time plot should be equal to -1 (this is the slope of the reference line in Figure 3, below); while with long-range dependence, T can be characterized by

$$\mathrm{Var}[T^{(m)}] \sim m^{-\beta}, \quad \text{as} \quad m \longrightarrow \infty, \;\; 0 < \beta < 1 \tag{2}$$

so the slope obtained from the variance-time plot should deviate from -1. Hurst parameter H is commonly used to measure the intensity of LRD, and is related to the parameter β in (2) by

$$H = 1 - \beta/2, \quad 1/2 < H < 1 \tag{3}$$

For short-range dependence processes, $H = 1/2$.

3.1 LRD in Packet Round-trip Delay Process

To get intuitively understanding of a packet round-trip delay process, we plot packets' go-trip delays against back-trip delays in Figure 1 and round-trip delays against packets' leaving times

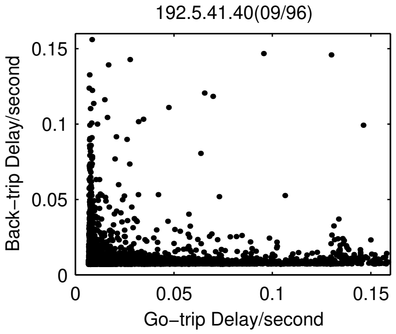

Figure 1: Packet go-trip delay vs. back-trip delay measured on route USNO.

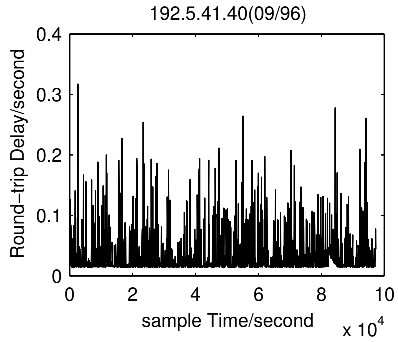

Figure 2: Packet round-trip delay vs. sampling time measured on route USNO.

(sampling times) in Figure 2. These figures show only a portion of the measured data in one route (192 .5 .41 .4).

We apply the variance-time plot technique to the whole set of data described in the previous section. To describe easily thereafter, we refer the route from our lab to machine 192 .5 .41 .4 as USNO route, to 192 .43 .244 .18 as NIST route, to 192 .67 .12 .101 as AU (Australia) route, to 192 .36 .143 .150 as SE (Sweden) route, and to 146 .83 .8 .200 as CL (Chile) route. (a)-(e) of Figure 3 show the variance-time plots of the five routes. It is interesting to notice that all five variance-time curves show asymptotic slopes that are distinctly different from -1 (dash dotted line). Using least square linear fit and omitting both the first five points and the last five points on each curve, we can easily estimate the slope (β) of the curves and calculate the corresponding Hurst parameters (H) by (3) as shown in table 2.

It is surprised to find that all Hurst parameters > 0.75, which

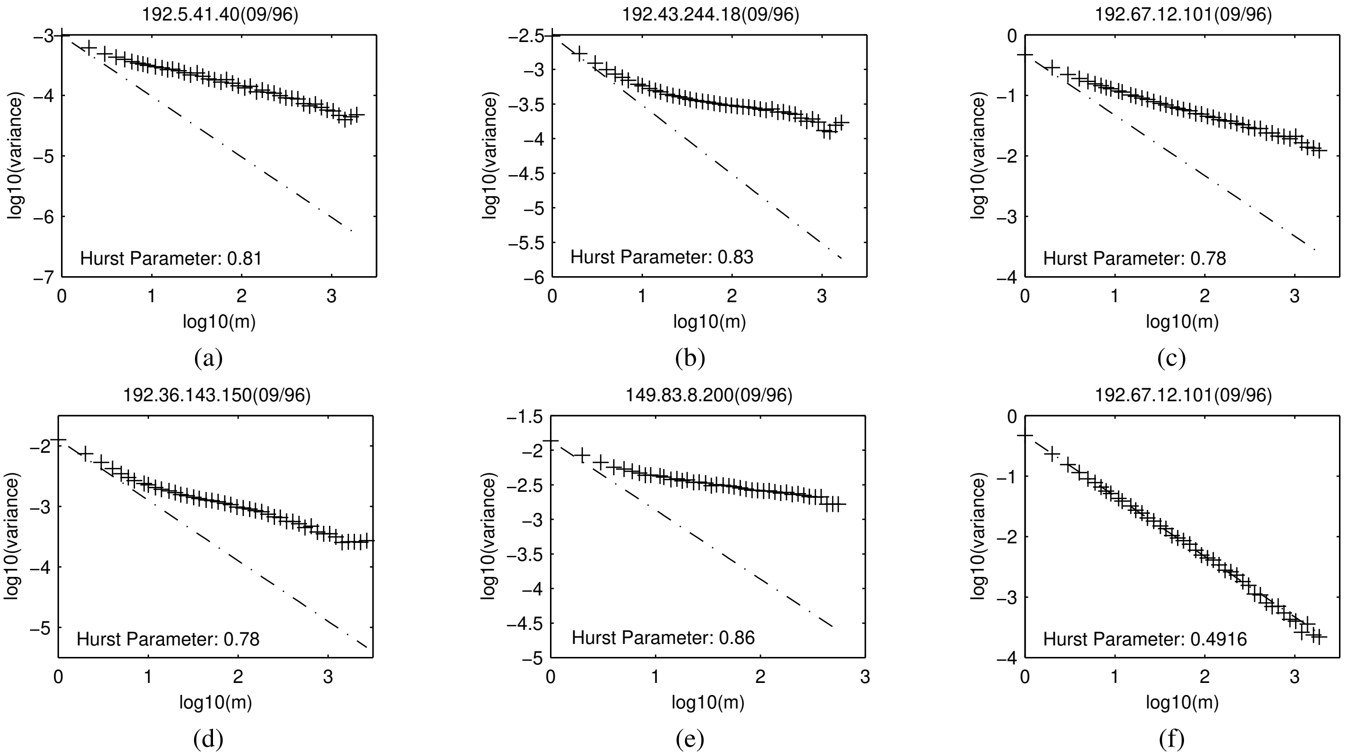

Figure 3: Variance-time plot for route (a) USNO, (b) NIST, (c) AU, (d) SE, (e) CL, (f) AU(reshuffled).

Route	USNO	NIST	AU	SE	CL
β	0.38	0.34	0.44	0.44	0.28
H	0.81	0.83	0.78	0.78	0.86

Table 2: β and H of the five routes obtained by least square linear fit.

suggests that every packet round-trip delay process in the five routes shows a strong intensity of LRD. However, not all five variance-time curves show perfect straight lines. Variance-time curves for route USNO and AU ((a), (c)) being almost linear, implies that packet-round trip delay processes in the two routes can be considered as exactly self-similar more than as asymptotically self-similar; while for route NIST, SE, CL ((b), (d), (e)), these asymptotically linear variance-time curves suggest that the packet round-trip delay processes in these routes are asymptotically self-similar. In order to further demonstrate that LRD really exists in the measured time series data, we randomly reshuffle the time series measured in route AU, namely, randomly change the data's orders in the time series, and then apply the variance-time plot to the reshuffled time series. The result is shown in (f) of Figure 3. It can be seen , in this case the variance-time plot is consistent with the dash dotted line. We believe that (a)-(e) provide convincing evidence of the existence of LRD in packet round-trip delay process.

3.2 Influence of Sampling Interval on H

As described above, the measured packet round-trip delay time series is only a sampled process with an average sampling interval. If the original packet round-trip process is LRD, then its power spectral density is of the form

$$S_T(\omega) \sim |\omega|^{-\gamma}, \;\; \gamma = 2H - 1, \quad \text{as} \quad \omega \longrightarrow 0 \tag{4}$$

So, as $\omega \to 0$, $S_T(\omega)$ diverges to infinity ($1/f$- noise phenomenon). In order to approximately capture the divergence feature near $\omega = 0$, the sampled process should have a long enough sampling period. In our case the sampling period is a month. Within a certain sampling period, the shorter the sampling interval, the better the sampled process to reflect the characteristics of the original process.

In our experiment the minimum sampling interval is limited by the experimental mechanism (determined by the polling interval of the time server in our lab). We cannot get a better sampled process with a shorter sampling interval from the measured data set. Oppositely, we can withdraw subsets of data from the original measured data set to construct sampled processes with different sampling intervals longer than the original one. With the withdrawn subsets of data we can check with the influence of a sampling interval on H. (a) and (b) of Figure 4 show H against the sampling interval L (L is the ratio of the sampling interval of the subset process and the original one) plots for route USNO and AU. We can see that with the increase of L, H tends to decrease, the vibration of H in the plots probably is caused by the estimation error introduced by the variance-time plot method. It is worth to point out that even with sampling intervals longer than half an hour (corresponding to $L > 25$ in both plots), the sampled processes still can show LRD property, this has special implication in Internet (see section 4).

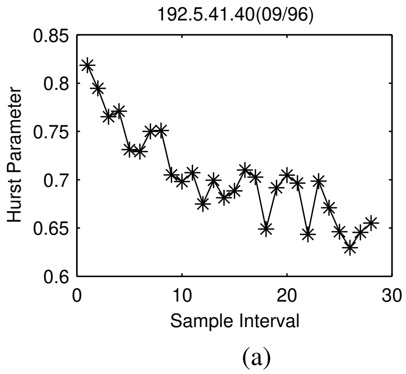

Figure 4: Hurst parameter vs. sampling interval for route (a) USNO, (b) AU.

3.3 The Distribution of Packet round-trip Delays

The probability distribution of packet round-trip delays is an important aspect for understanding the characteristics of a packet round-trip delay process. By investigating the distribution of packet round-trip delays, we can get some insights into the queueing performances along the packet route (see section 4). (a)-(e) of Figure 5 show the statistic results of the aforementioned data set. All plots are drawn in log-log scale. In each plot, the dashed curve corresponds to an exponential distribution with the mean equal to that of the packet round-trip delays in the investigated route. The solid curve is the distribution of $P(\textit{round-trip delay} > x)$. The Two dotted lines indicate that 10% and 50% of all measured data points are to the right of the respective lines, the two dash dotted lines have the same meaning, but correspond to the exponential distribution only for the purpose of reference. From these plots we can see that, the distribution of packet round-trip delays in Internet is very different from the exponential distribution, and in a wide range of x it decays more slowly than the exponential rate. With x beyond a certain value, the distribution of packet round-trip delay drops rapidly; this phenomenon is believed to be caused by the finite buffer capacity in Internet and the losing of long delay packets. These findings have significant implication to the reason of LRD in a packet round-trip delay process in Internet (see section 4).

4 On the Reason of LRD in a Packet Round-trip Delay Process

As we have noted in section 2, a packet delay in Internet consists of four components. Being fixed the length and routing of a packet, the packet delay only changes with the queue delays on the route. In Internet a packet route generally includes a tandem of many queues (which as many as hops), so the packet round-trip delay process is mainly determined by the queueing performances along the route. Let express $T(t)$ as the form

$$T(t) = T_0 + \triangle T(t) \tag{5}$$

where T_0 is the constant part of a packet round-trip delay. For the measured data, T_0 is simply the minimum packet round-trip delay for a route. In this case, a packet with the minimum round-trip delay means when the packet goes through all the links along its route it does not meet any other packet ahead of it; all the queues in the nodes along the route at the moment when the packet passes them are empty. This situation is not a rare event in Internet according to the measured data. $\triangle T(t)$ is the varying part of a packet round-trip delay, which is the sum of all the queue delays along its route

$$\triangle T(t) = \sum_{i=1}^{N} Q_i(t) \tag{6}$$

where N is the total number of queues along its route, $Q_i(t)$ is the delay of the packet experiencing in queue i.

Suppose we apply the Kleinrock independence approximation [2] to Internet, which means we adopt an /M/M/1 queueing model for each link regardless of the interactions among traffics on different links and LRD in traffic streams, then the complementary distribution of Q_i (suppose Internet is in steady state,$P(Q_i > x) = lim_{t\to\infty} P(Q_i(t) > x)$) is of the form

$$P(Q_i > x) = e^{(-\gamma_i x)}, \ \ \gamma_i = \mu_i(1-\rho_i) \tag{7}$$

where μ_i is the service rate on link i, ρ_i is the corresponding utilization factor. The complementary distribution of $\triangle T$, namely, $P(\triangle T > x) = lim_{t\to\infty} P(\triangle T(t) > x)$, is of the hypo-exponential form

$$P(\triangle T > x) = \sum_{i=1}^{N} a_i e^{(-\gamma_i x)} \tag{8}$$

where

$$a_i = \prod_{j=1, j\neq i}^{N} \frac{\gamma_i}{\gamma_j - \gamma_i}, \quad \sum_{i=1}^{N} a_i = 1,$$

(8) shows that $\triangle T$ should decay exponentially fast. However, Figure 5 indicates that $\triangle T$ decays more slowly than exponential rate. This fact implies the Kleinrock independence approximation may not be valid in the case of Internet.

The autocorrelation function of $Q_i(t)$ has no explicit analytic form even under the condition of the Kleinrock independence approximation and the assumption of infinite buffer capacity at each link node. The reality in Internet is that the buffer capacity at each

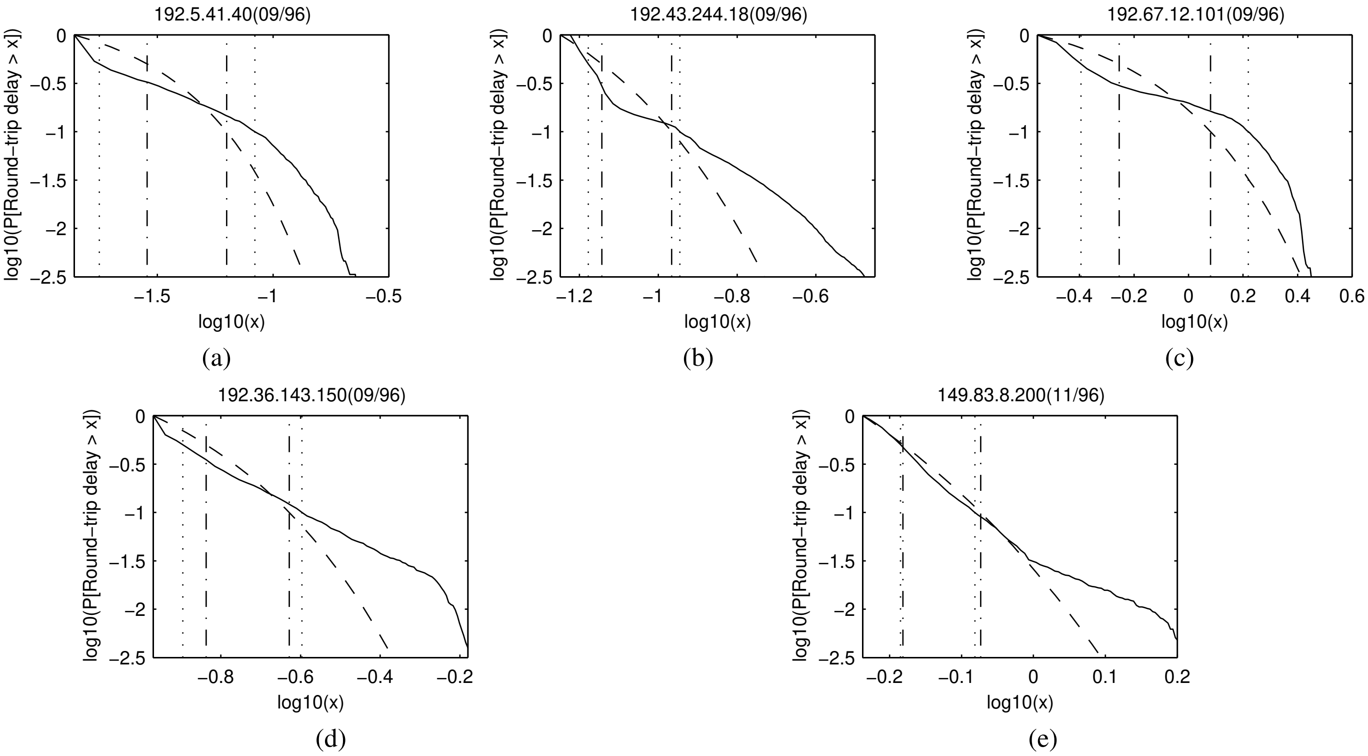

Figure 5: 5 Distribution of packet round trip delays for route (a) USNO, (b) NIST, (c) AU, (d) SE, (e) CL.

link node is limited, so whenever the queue i is empty or full the autocorrelation of $Q_i(t)$ should be broken beyond a certain lag time in case that the arrival packet traffic is a short-range dependence process. Then $Q_i(t)$ should not behave as an LRD process, so does $\triangle T(t)$. However, if the arrival packet traffic is an LRD process, the autocorrelation of $Q_i(t)$ should not be entirely erased by events of queue emptying and buffer filling. The structure of the autocorrelation function may be interrupted by these events, but the dependence in $Q_i(t)$ should be kept in some degree even beyond these events' happening. This is likely the reason why a packet round-trip delay is an LRD process. In Figure 1 and Figure 2 we can find that the queue emptying event happens frequently (which corresponds to the point with minimum packet delay); Figure 4 indicates that even with a sampling interval longer than half an hour (with a probability near 1 that a queue emptying event will happen within such a long interval), the sampled packet round-trip delay process still shows long-range dependence. This fact further explains that the LRD in arrival packet traffic possibly causes the LRD in packet round-trip delays.

Assuming the arrival packet process is an LRD process, Norros [8] has derived an asymptotic lower bound for the probability distribution of $P(Q_i > x)$ by driving a deterministic service time queue with an fractal Brownian motion process , as

$$P(Q_i > x) \sim e^{(-cx^{2-2H})} \tag{9}$$

where H is Hurst parameter, c is a constant determined by H and the traffic source number. From (9) we see that the complementary distribution of a queue length decays more slowly than strictly exponential rate, this result is consistent with the finding of the distribution of packet round-trip delays.

5 Conclusions

The finding of the long range-dependence both in LAN and WAN network traffic made it essential that network engineers understand the impact of LRD on network performances. Internet is a huge collection of various networks connected together. Many applications (audio, video, etc.) run across Internet is very sensitive to the performances of the whole Internet. A packet round-trip delay is an important metrics of the performance of Internet.

In this paper, we analyze the sampled packet round-trip delay processes measured by NTP in the environment of Internet. The main findings of our statistical analysis of the measurements on five different Internet routes for a period of a month are that: (i) The packet round-trip delay processes of the five routes in Internet are all LRD processes; they are either exactly self-similar or asymptotically self-similar; the intensity of LRD (or self-similarity) measured in terms of Hurst parameter H is different for different routes. (ii) The measured packet round-trip delay process is only a sampled process with certain sampling interval; H tends to decrease with the increase of sampling interval for the reason that with a longer sampling interval more error is introduced into the sampled process in sense of comparing with the original process; it is worth to notice that even with the sampling interval longer than half an hour the sampled packet round-trip delay process still shows the property of LRD; (iii) The probability distribution of $P(\textit{round-trip delay} > x)$ decays with x more slowly than an exactly exponential rate. Based on our findings, we think that the LRD in a packet round-trip delay process is caused by the LRD in the arrival packet traffic in Internet.

More study on the packet round-trip delay process in Internet is necessary in future. In this paper, due to the limitation of measurements we cannot directly investigate the property of sampled packet round-trip delay processes with smaller sampling intervals than the current minimum one. Perhaps the sampled packet round-trip delay process with a smaller sampling interval is more interesting to understand the performance of Internet. This work is currently being carried out in our lab. Some applications are sensitive to the packet one-way (go-trip or back-trip) delay (such as teleconferencing). It is believed that a packet one-way delay process is similar to a round-trip delay process. In order to accurately measure a packet one-way delay, we must synchronize the clocks of the packet source and destination to some degree of accurate. This can be done by running NPT both in the packet source and destination, but NTP must run continuously long enough to bring the two clocks into synchronization. This work is also currently being carried out in our lab.

References

[1] M. G. Baker, J. H. Hartman, M. D. Kupfer, K. W. Shirriff, and J. K. Ousterhout. Measurements of a distributed file system. In *Proceedings of the Thirteenth ACM Symposium on Operating System Principles*, pages 198–212, Pacific Grove, CA, October 1991.

[2] Dimitri Bertsekas and Robert Gallager. *Data Networks*. Prentice Hall, Englewood Cliffs, New Jersey 07632, 2nd edition, 1992. see page 211-213.

[3] M. E. Crovella, M. S. Taqqu, and A. Bestavros. Heavy-tailed probability distributions in the world wide web. Preprint, Submitted for publication in the book "A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions", 1996.

[4] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Transactions on Networking*, 4(2):209–223, April 1996.

[5] M. Grossglauser and Bolot J-C. On the relevance of long-range dependence in network traffic. *ACM/SIGCOMM Computer Communications Review*, 26:15–24, 1996.

[6] W. E. Leland, M. S. Taqqu, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, February 1994.

[7] D. M. Lucantoni, M. F. Neuts, and A. R. Reibman. Methods for performance evaluation of vbr video traffic models. *IEEE/ACM Transactions on Networking*, 2(2):176–180, April 1994.

[8] I. Norros. A storage model with self-similar input. *Queueing System*, 16:387–396, 1994.

[9] K. Park, G. Kim, and M. Crovella. The effects of traffic self-similarity on tcp performance. Preprint, available at: http://www.cs.bu.edu/faculty/crovella/papers.html, 1996.

[10] V. Paxson and S. Floyd. Wide area traffic: The failure of poisson modeling. *IEEE/ACM Transactions on Networking*, 3(3):226–244, June 1995.

[11] W. Richard Stevens. *TCP/IP Illustrated*, volume 1. Addison Wesley, 1994.

[12] M. S. Taqqu, V. Teverovsky, and W. Willinger. Estimators for long-range dependence: an empirical study. *Fractals*, 3(4):785–788, 1995.

[13] M. S. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *ACM/SIGCOMM Computer Communications Review*, 27:5–23, 1997.

[14] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson. Self-similarity through high-variability: Statistical analysis of ethernet lan traffic at the source level. *IEEE/ACM Transactions on Networking*, 5(1):71–86, January 1997.