On the Distribution of Round-trip Delays in TCP/IP Networks

Tamás Éltető *
Traffic Analysis and
Network Performance Laboratory
Ericsson Telecommunications Hungary
e-mail: Tamas.Elteto@eth.ericsson.se

Sándor Molnár †
High Speed Networks Laboratory
Department of Telecommunications and Telematics
Technical University of Budapest
e-mail: Molnar@ttt-atm.ttt.bme.hu

Abstract

TCP traffic round-trip delay measurements have been carried out in the Ericsson Corporate Network. Based on these data a comprehensive analysis has been performed to identify basic state changes in the network and to find the parametric distribution of the round-trip delay. The main finding of our research is that the round-trip delay can be well approximated by a truncated normal distribution. A method for parameter setting with the statistical tests for validation is also presented. Finally, we discuss the possible explanations of this result.

1 Introduction

We can observe the expanding application of the TCP/IP protocol suite in the last decade. This is mainly due to the rapid growth of the Internet. The success of Internet can partly explained by the excellent philosophy of this protocol suite [9]. The TCP provides a reliable transport layer service for many Internet applications including WWW, e-mail, news, etc. However, many drawbacks of this protocol suite have also arisen and several modifications have been developed during the previous years, e.g. TCP Tahoe, Reno and Vegas [5]. In spite of the fact that TCP/IP introduced several problems in different networks it seems to be one of the main technologies for the future. Therefore a basic understanding of the traffic produced by the TCP/IP protocols is important for teletraffic engineering of that type of networks [8, 11].

TCP provides a connection-oriented, reliable, byte stream transport layer service [9]. The application data is broken into segments of a size limited by the Maximum Transfer Unit (MTU). When TCP transmits a segment it starts a timer and waits for the acknowledgement of the reception of the segment from the other end point. If an acknowledgement is not received in time, the segment is retransmitted. The TCP has a flow control mechanism. The end points of the TCP connections have some finite amount of socket buffer space. A receiving TCP announces the current value of the available buffer space to the sender in order to force the sender to send as much data as the receiver has buffer space for. The current value of the available space is called window size. The applications can set the maximum value of the socket buffer size.

One of the key characteristics of the TCP mechanism is the Round-Trip Time (RTT). All TCP implementations attempt to estimate the current round-trip delay by observing the pattern of delay for recent segments. The philosophy of these methods is to set the timer to a value somewhat greater than the estimated round-trip delay. Since the understanding of the phenomenon which causes the current round-trip delay is essential for the design of TCP networks, we chose this characteristic as the target of our analysis.

Previous studies on IP packet delay [1, 2, 3, 10] have

*H-1037 Budapest, Laborc u. 1, Hungary
†H-1117 Budapest, Pázmány Péter sétány 1/D, Hungary

found high variability of the delay process with strong short and long term correlations [1, 10]. Self-similarity properties of round-trip delay are also observed. Moreover, periodicity is detected [3] due to the routing table updates [1]. The delay process is found to be very environment dependent and it was the main motivation of our study to find invariant characteristics supported by the physical understanding of this phenomenon.

In Section 2 we report our measurement study of TCP round-trip delays. Based on this database we performed a comprehensive analysis of the delay process. We present a method which helps us to identify the changes of network states in Section 3. We can identify stationary periods between state changes tested by trend analysis. Within each period a parametric distribution can be fitted to the data. Our basic finding is that the round-trip delay distribution can be well approximated by a truncated normal distribution under some circumstances presented in Section 4. We also present a discussion about the possible physical explanations of this phenomenon.

2 Measurements

The data was collected in the Ericsson Corporate Network between Budapest and Stockholm. The route contains a WAN X25 link and high speed LAN's. The destination addresses were at a distance of 5-7 hops. Practically, we had two LANs (10 Mbit/s Ethernet and 155 Mbit/s ATM) connected by a relatively slow link (768 Kbit/s). The buffer of the slow link could store approx. 100-150 packets while the packet size was 1500 bytes.

We used the PING program for measuring the packet round-trip time. The PING program sends echo requests to a specified address in equal time intervals. In our measurements it was set to one second. The reliability of the measurements depends on the reliability of the computer's system clock. However, in our case we found that the range of the errors was negligible that is it was several magnitudes under the range of the measurements.

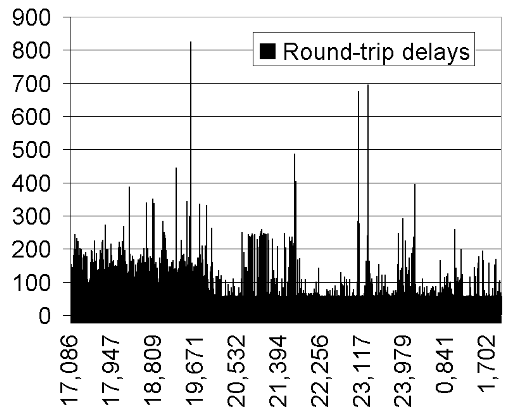

Figure 1. A measurement period of round-trip delays (ms/hours)

3 Analysis

3.1 Change-point detection analysis

We consider the data (see Fig. 1) as a series of disjoint intervals within which the delay series can be regarded as stationary. When an interval ends and a new interval starts the characteristics of the data change. In this section we will concentrate on a method of finding these change-points.

3.1.1 The two sample Kolmogorov-Smirnov test

The most important tool used by theoretical change-point detection is the (two sample) Kolmogorov-Smirnov test (see [6]). This section describes how this test works, and also tells about the special case of equal sample series length.

Let

$$X_1, X_2, \dots X_N$$

be the series of the first sample (of length N), and

$$Y_1, Y_2, \dots Y_M$$

be the series of the second and let $F_N(x)$ and $G_M(x)$ be the corresponding empirical distributions. Then the statistics is

$$D_{N,M} = \max_x |F_N(x) - G_M(x)|$$

The limit distribution theorem says that

$$\lim_{N,M\to\infty} P\left(\sqrt{\frac{NM}{N+M}} D_{N,M} < y\right) =$$

$$= K(y), \quad 0 < y < \infty,$$

where $K(y)$ is the Kolmogorov distribution.

Assume that $N = M$. In this case $D_{N,N}$ is computed practically in the following way. By ordering the merged samples in an increasing order we have

$$Z_1^{(2N)} \leq Z_2^{(2N)} \leq \ldots \leq Z_{2N}^{(2N)}.$$

Now, let

$$\delta_i = \begin{cases} 1, \text{if } Z_i^{(2N)} \text{ comes from the} \\ \qquad\qquad\qquad \text{series of } X_k \\ -1, \text{otherwise} \end{cases}$$

and

$$S_j = \delta_1 + \delta_2 + \ldots + \delta_j.$$

Here we get $D_{N,N}$ so that

$$D_{N,N} = \frac{1}{2N} \max_{1\leq j\leq 2N} |S_j|.$$

This method computes a distance between two empirical distributions using the original data. The method was described assuming an equal number of observations but this is not the general case. The value of $D_{N,N}$ related to the difference between the originating distributions. The larger its value, the larger the difference.

3.1.2 Sliding intervals for change-point detection

Our purpose is to find the edges of the intervals where the distribution remains the same [4]. In order to determine these intervals the empirical distributions of two neighboring time windows are compared and these two windows are slid together along the series. In this way, we get another series, called test-series. (The comparison was done by Kolmogorov-Smirnov test of the subsamples.) The edges of an interval are marked by values in the test-series which are much more significant than their neighborhood.

We had some difficulties in using the inverse Kolmogorov function so instead of $K^{-1}(D_{N,N})$, we used the plot of $D_{N,N}$'s. In fact, this gives the same result as the other because now we are interested in local maxima, and not in exact probabilities and any distribution function is monotone non-decreasing.

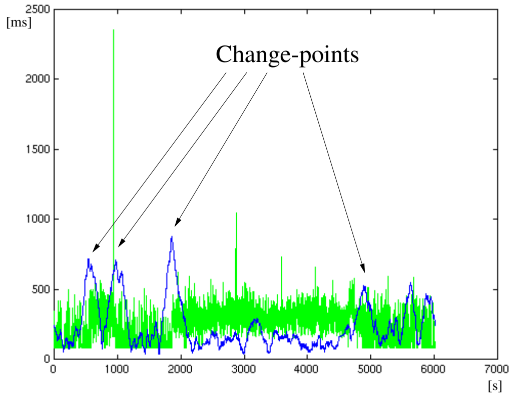

Figure 2. Result of the test (black) and the original data (grey) for a 6000 seconds period

Periods which can be thought to be appropriate for fitting parametric distributions marked by strong local maxima are shown in Figure 2. The scale in Figure 2 is valid for the measured data. In this case the length of the sampling intervals was 250 data. The resulting test-series was rescaled in order to be able to depict the two series in the same plot.

Figure 3 shows a shorter interval which has less than twice the length of the sampling window's length, while Figure 4 zooms to a smaller time scale showing a case when the window size equals the interval length.

Fortunately, we can recognize the sharp edges here. We think that this is because the distribution changes suddenly, and the distributions remain approximately constant on both sides of the edges. We observed that if a period is much longer than the sampling length one can observe the edges quite clearly. If the length of a

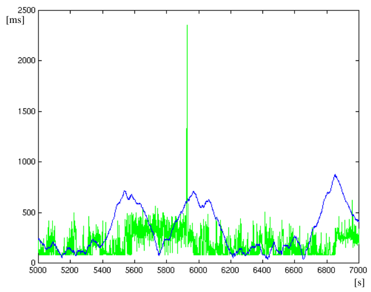

Figure 3. Result of the test and the original data for a 5000 seconds period

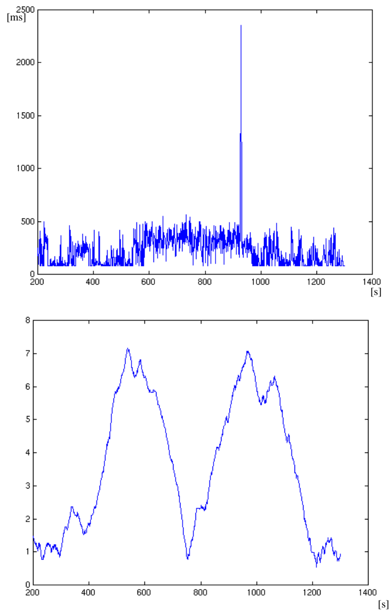

Figure 4. Result of the test and the original data for a 1000 seconds period

"calm" period is approximately equal to the length of the sampling intervals this method cannot be applied.

With the previously mentioned assumptions, a possible change-point detection method is quite straightforward as it is based on detecting local maxima.

Two change-points may bound an interval if the length of the interval is greater or equal to the length of the window size.

This algorithm was applied to find possible intervals which are appropriate for examination for a common parametric probability distribution. A trend analysis was performed based on the method of moving averages to reveal the presence of long trends in the intervals which may ruin the stationarity assumption.

3.2 Distribution analysis

In this section we present the methods which have been used to fit parametric distributions to the data of each interval.

3.2.1 The estimated quantile functions, and histograms

In order to set up a parametric model to the distributions of the chosen periods we have to find good methods for the exploratory data analysis. Parzen in [7] advises quantile functions and density quantile functions in addition to the empirical distribution function.

The quantile function

$$Q(u),\ 0 \leq u \leq 1$$

is defined as follows: for a general distribution function $F(x)$ which is continuous from right,

$$Q(u) = F^{-1}(u) = \inf\,(x:\ F(x) \geq u).$$

For $-\infty < x < \infty$ and $0 < u < 1$

$$F(x) \geq u \quad \leftrightarrow \quad Q(u) \leq x.$$

Denote by $\hat{F}(\cdot)$ the empirical distribution function. Then denote by $\hat{Q}(\cdot)$ the empirical quantile function, which is:

$$\hat{Q}(u) = \inf\,(x:\ \hat{F}(x) \geq u).$$

This is a piecewise constant function whose values are the ordered statistics:

$$\hat{Q}(u) = X_{(j)} \quad \text{for}$$

$$(j-1)/n < u \leq j/n, j = 1, \ldots, n$$

where $X_{(j)}$ is the j^{th} value of the ordered sample:

$$X_{(1)} < X_{(2)} < \ldots < X_{(n)}.$$

For $u = 0$ define $\hat{Q}(0)$ to be the sample minimum.

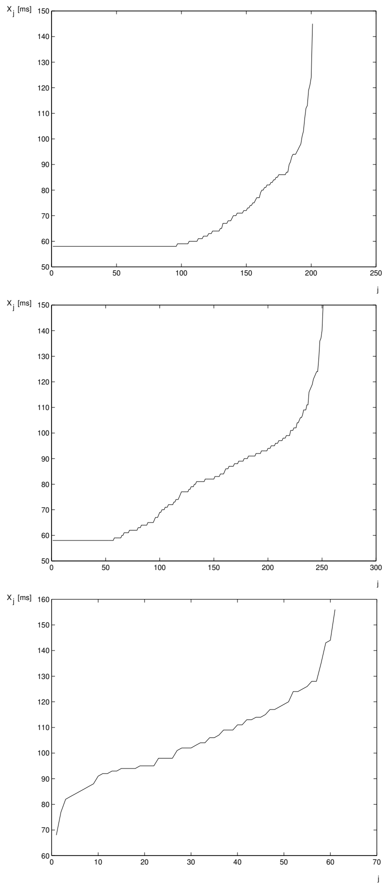

Figure 5. The quantile functions of the rare, medium, and busy periods

By investigating the quantile functions of the data in different intervals we can identify three main types of functions related to the rare, medium and busy traffic activities.

There are three different empirical quantile functions corresponding to the three periods with different traffic types (see Figure 5).

The minimal value of the a daily series was 64 ms. A series takes this value quite often during the night. It is the propagation delay, that is, it takes physically 64 milliseconds to have the packet from sending to arriving, and delivering back to us. We have three main types of histograms depicted in Fig. 6.

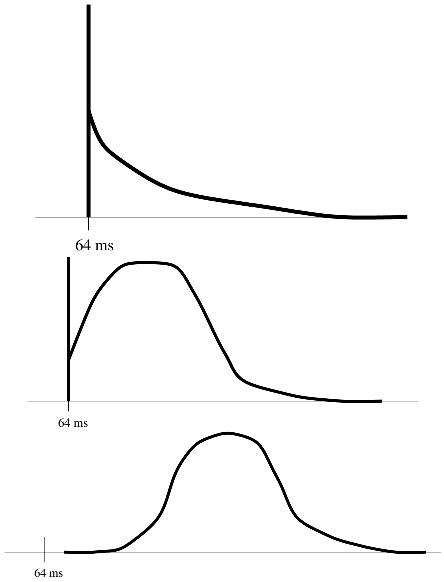

Figure 6. Roughly outlined histograms of the rare, medium, and busy periods

The main difference between these histograms is the number and location of local maxima. As Figure 6 shows, there is one local maximum in the rare period, namely the propagation delay, and there is also one local maximum in the busy period, which is approximately equal to the average delay, because the histogram is nearly symmetrical in the busy period. However, there are two local maxima in the medium period.

The symmetrical, bell-shaped apperance of the histogram in busy periods suggests normality (Figure 7). The quickest way of checking this is the quantile-quantile plot which will be detailed in Section 4.2. The Q-Q plot of a busy period is depicted in Figure 8. Here the sample average and the empirical variance were used as estimators of the mean and variance of the normal distribution.

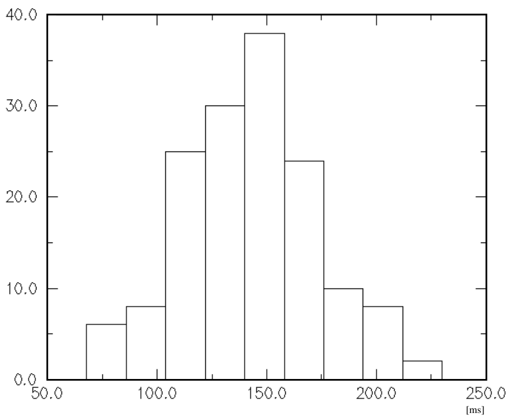

Figure 7. The histogram of the busy period

4 Modeling

In this section we present our procedure to estimate the parameters of a truncated normal distribution to the data with its validation test.

4.1 Estimating parameters of truncated normal distribution

The shape of the distribution in the busy period motivated the use of the normal distribution family. Any normal distribution has no limit below, however, we cannot measure round-trip delay under the physical lower limit. Possible solution to this "contradiction" is to suppose truncated normal distribution.

Assume a normally distributed random variable

$$Y \sim N(\mu, \sigma^2).$$

We generate a new random variable X from Y by setting

$$X = \begin{cases} c, & Y \leq c \\ Y, & Y > c \end{cases},$$

with a fixed c.

Somehow, we have to estimate the parameters μ, σ and c. Obviously c would be the lowest measured value, and we obtain μ and σ in the following way.

Let the quantile be α such that

$$\alpha = \Phi(\frac{c-\mu}{\sigma}), \tag{1}$$

here the function $\Phi(\cdot)$ denotes the standard normal distribution function.

Again we assume that the originating distribution is $\mathrm{Normal}(\mu, \sigma^2)$.

$$\begin{aligned} &\mathrm{E}(X|X>c) = \\ &= \int_c^{+\infty} t \frac{1}{\sqrt{2\pi}\sigma(1-\alpha)} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt = \\ &= \frac{1}{1-\alpha} \int_{\frac{c-\mu}{\sigma}}^{+\infty} (\sigma u + \mu) \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du = \\ &= \frac{\sigma}{1-\alpha} \int_{\frac{c-\mu}{\sigma}}^{+\infty} u \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du \; + \\ &+ \frac{\mu}{1-\alpha} \int_{\frac{c-\mu}{\sigma}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du = \\ &= \frac{\sigma}{1-\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{(c-\mu)^2}{2\sigma^2}} \; + \\ &+ \frac{\mu}{1-\alpha} \left(1 - \Phi(\frac{c-\mu}{\sigma})\right). \end{aligned} \tag{2}$$

If we define

$$\begin{aligned} a &= \frac{1}{\sqrt{2\pi}(1-\alpha)} e^{-\frac{(c-\mu)^2}{2\sigma^2}} = \\ &= \frac{1}{\sqrt{2\pi}(1-\alpha)} e^{-\frac{\left(\Phi^{-1}(\alpha)\right)^2}{2}}, \end{aligned}$$

as a function of α.

By this definition and from (2) we get (3):

$$\mathrm{E}(X|X>c) = \sigma a + \mu, \tag{3}$$

$$\mu = c - \sigma \Phi^{-1}(\alpha). \tag{4}$$

From (3) and (4) we obtain the following estimation for σ

$$\hat{\sigma} = \frac{\mathrm{E}(X|X>c) - c}{a - \Phi^{-1}(\alpha)}, \tag{5}$$

and substituting σ in (4) we get $\hat{\mu}$.

Finally, we estimate α and $\mathrm{E}(X|X > c)$ such that

$$\hat{\alpha} = \sum_{i=1}^{N} \chi(X_i < c), \tag{6}$$

$$\mathrm{E}(X|X > c) \approx$$

$$\approx \frac{1}{|\{1 \le i \le N : X_i > c\}|} \sum_{i=1}^{N} X_i \chi(X_i > c). \tag{7}$$

Then we are ready to estimate $\hat{\sigma}$, and $\hat{\mu}$ by (3), and (4).

4.2 Validating

4.2.1 Quantile-Quantile plot

Q-Q plot is a clear graphic but not so exact a way of matching the empirical distributions of two samples. We have a function $q_{(X_1,\ldots),(Y_1,\ldots)}(\cdot)$ such that

$$q_{(X_1,\ldots),(Y_1,\ldots)} : \mathbf{R} \to \mathbf{R}$$

and if the graph of this function fits the graph of the $u(x) = x$ function well, then the two distributions are quite similar. However, if there is curvature in the $q(\alpha)$ function we are able to determine where the location of the difference can be. Let us define the $q(\cdot)$ function as follows.

Consider $Q_X(u)$ the estimated quantile function from the sample

$$X_1, X_2, \ldots X_m$$

and $Q_Y(u)$ from the sample

$$Y_1, Y_2, \ldots Y_n.$$

Now let $q(\cdot)$ be the following

$$q(Q_X(u)) = Q_Y(u),$$

that is

$$q(x) = Q_Y(F_X^*(x)), \tag{8}$$

where

$$F_X^*(x) = \frac{1}{m} \sum_{i=1}^{m} \chi(X_i \le x).$$

There are some advised restrictions on plotting $q(\alpha)$. For the sake of good visibility, one should evaluate less or an equal number of α's than the minimum of m and n, and use rather identical increments of α.

In order to show that the assumptions about the quantile functions and the histograms of the periods hold, we plot the quantile-quantile plot of the measured data and the fitted distributions.

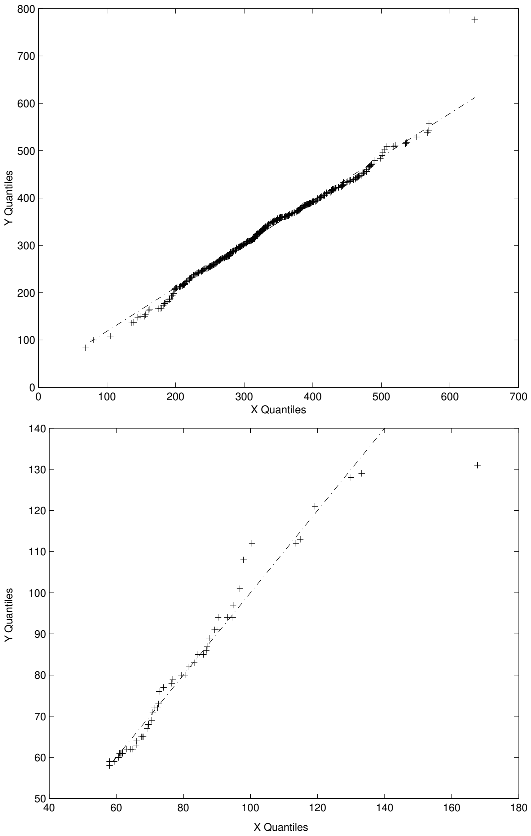

Figure 8. Q-Q plot with an estimated normal and truncated normal distribution

4.2.2 Kolmogorov-Smirnov test

The one-sample Kolmogorov-Smirnov test with normal distribution is a very sensitive and reliable method of

Table 1. Results of Kolmogorov-Smirnov test

Descriptive Statistics

	N	Average	Std. Dev.	Min	Max
X	543	332.94	91.12	83.00	1042.00

One-Sample Kolmogorov-Smirnov Test

		X
N		543
Normal Param.'s	Mean	332.94
	Std. Dev.	91.12
Most Extreme Diff.'s	Abs.	.046
	Pos.	.046
	Neg.	-.036
Kolmogorov-Smirnov Z		1.075
Asymp. Significance	(2-tailed)	.198

Table 2. Results of Kolmogorov-Smirnov test of the modified series

Descriptive Statistics

	N	Average	Std. Dev.	Min	Max
X	539	329.35	80.32	83.00	565.00

One-Sample Kolmogorov-Smirnov Test

		X
N		539
Normal Param.'s	Mean	329.35
	Std. Dev.	80.32
Most Extreme Diff.'s	Abs.	.022
	Pos.	.022
	Neg.	-.021
Kolmogorov-Smirnov Z		.508
Asymp. Significance	(2-tailed)	.959

testing whether our data has normal distribution or not. The method of the test is almost the same as it was outlined in Section 3.1. However, there are some differences in this case when we test a specific distribution against the data.

There are many statistical software packages available which contain this test. We used SPSS 8.0. We had to choose such a period where we had busy traffic activities in the network because only stationary intervals during this time shows resemblance to the normal distribution.

Thus we chose an interval for the test. We repeated the test twice on the interval. For the first time we did not modify the data. The results are given in Table 1.

The siginficance of .198 indicates that if we perform a hypothesis test at a significance of .2 - the level of .8 - then we have to reject the null hypothesis that our data has normal distribution. However, this can be misleading, therefore one should not discard the results of the Q-Q plot analysis. There may be a little perturbance in the data of which has strong negative effects especially in the case of our sensitive test.

For the second time, we ignored the data larger than 600 ms in the series and repeated the test. There were 4 pieces of data segments dropped out of 543 that is less than 1%. This time the results are shown in Table 2.

It can be seen that the significance increased strongly. This shows that only a small number of "bad" values can modify the results in an undesired manner. In this case, we can accept the null hypothesis of normality at almost any level. This fact does not prove clearly that the originating distribution is a normal one but it indicates it very strongly.

4.3 Engineering considerations

TCP traffic strongly differs from those traffic types - in a wider sense - that we could observe so far. (For example road traffic, arrivals to a shop or to a telephone exchange, customers' demands.)

It can be inferred that the governing distribution changes suddenly, because of the properties of the TCP connection.

When a new connection is established, it explores the available capacities. This causes considerably increasing traffic - in slow start state the connection increases its traffic exponentially. This is the reason why change-point detection can be used to examine the series and search for the beginning of connections.

Important attribute is the normality of the queue length. In our case the packet arrivals are not independent of the past, furthermore they may have very strong correlations. To obtain suggestions about how we can get normal-like queue length, a slightly different approach than simply examining the packet arrivals has been used.

The size of the congestion window is approximately the number of packets in the network that are unacknowledged. This implies that the number of packets in the bottleneck queue plus the number of packets in the "wire" elsewhere gives us the congestion window.

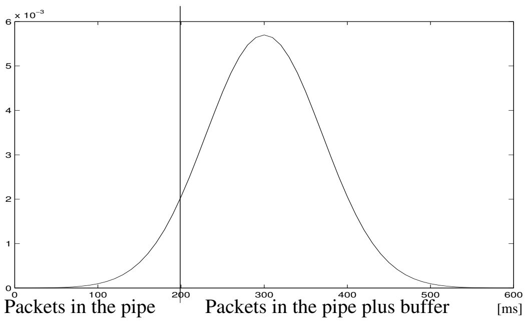

Figure 9. Possible explanation of truncated normality

In fact, the number of data packets and their acknowledgements together equals the size of the congestion window. Consider only the number of packets of a single TCP connection traveling to the receiver. Assume that the values of the process in each second are independent from each other. Take other TCP connections and assume also that these are independent from each other. Assume finally that the "wire" contains a fixed number of packets. Then we can use the central limit theorem on the sum of independent, identically distributed random variables. Thus the number of packets of the different connections in the queue and the wire together may be independent identically distributed random variables, and their sum is asymptotically normally distributed. In this case, queueing occurs when the wire is filled with packets, and so we get the appropriate truncated normal distribution.

Notice that the addition of a single connection to the queue length at the bottleneck is related to the congestion window, which is controlled by the sender TCP, and therefore somehow determinative.

4.4 Remarks about the expected goodness of fit

In the previous sections a method to find intervals which are suitable to distribution fitting is presented. The larger part of a daily series is not examined for parametric distributions because the characteristics (the average, empirical variance etc.) changed rather quickly, and the window size should have been set so small that the results had not been significant enough.

The autocorrelation function of the "calm" intervals was also estimated. As one can see in Figure 10 the estimation shows independence at longer lags, but suggests dependence within four or five lags, that is, the packet delays are correlated within five seconds. We have doubts concerning this dependence for two main reasons:

- *Theoretically* there must be no recognizable dependence between samples of queue length on a rate of several seconds, because the size of the congestion window which is approximately equal to the number of the packets of the connection in the queue changes very quickly, and they may depend on 100-200 ms having passed.
- *Practically* if the autocorrelation function of a smaller period of data was estimated, one should observe the same correlation pattern and it is obvious, for example in Figure 11, that this is not the case.

The estimated autocorrelation function of a series where the parameters have small changes and the series has small trend may exhibit similar phenomena. We think that the problem is that we cannot find *clear* periods. This raises further doubts about the goodness of fit here.

4.5 Conclusion and future work

In this paper we presented a TCP round-trip delay measurement and analysis study. Based on the two sam-

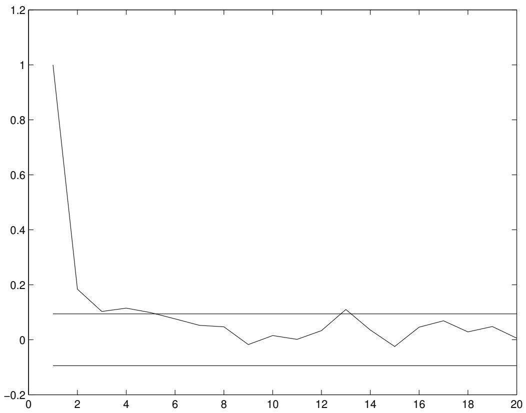

Figure 10. The estimated autocorrelation function of a calm period

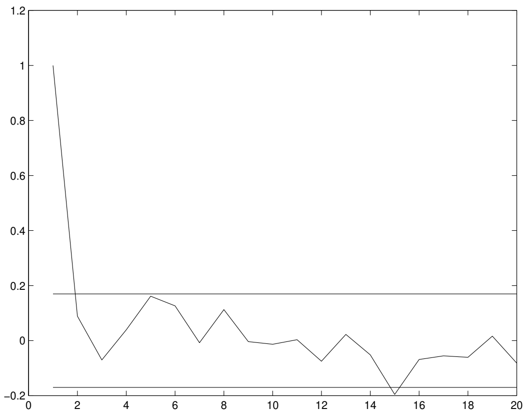

Figure 11. The estimated autocorrelation function of a short interval of a calm period

ple Kolmogorov-Smirnov test, a method has been developed to identify state changes in the delay process. We found that between these state changes the process can be assumed to be stationary with the same delay distribution. The stationary assumption was confirmed by a trend analysis test. We found that this distribution under some circumstances can be very well approximated by a truncated normal distribution.

We also presented a parameter setting method and a validation study of our results.

Our opinion is that this distribution arises as a result of the central limit theorem applied to the sum of TCP flows which can cause normally distributed queueing delays in network buffers. A deeper analysis for different traffic types, network environments and the effect of congestion avoidance mechanisms (e.g. Random Early Detection, Fair Queueing, etc.) on round-trip delay are the topics of our future research.

References

[1] J. Andrén, M. Hilding, D. Veitch, "Understanding End-to-End Internet Traffic Dynamics", *Globecom'98*, Sydney, Vol.2 pp.1118-1122, Nov 1998.

[2] J. Bolot, "Characterizing End-to-End Packet Delay and Loss in the Internet", *Journal of High-Speed Networks*, vol 2, no. 3, pp. 305-323, Dec. 1993.

[3] M. S. Borella, S. Uludag, G. B. Brewster and I. Sidhu, "Self-Similarity of Internet Packet Delay", *ICC'97*, pp. 513-517, Jun. 1997.

[4] M. Csörgő and L. Horváth, "Invariance Principles for Changepoint Problems", *Journal of Multivariate Analysis*, 27, pp. 151-168.

[5] K. Fall, and S. Floyd, "Simulation-based Comparisons of Tahoe, Reno, and SACK TCP", *Computer Communications Review*, July 1996.

[6] M. Kendall, A. Stewart, *The Advanced Theory of Statistics*, Charles Griffin London, 1966.

[7] E. Parzen, "Nonparametric Statistical Data Modeling", *Journal of the American Statistical Association*, March 1979, 105-126.

[8] V. Paxson, S. Floyd, "Why We Don't Know How to Simulate the Internet?", *Winter Simulation Conference*, 1997.

[9] W. Richard Stevens, *TCP/IP Illustrated*, Volume 1, Addison-Wesley, 1995.

[10] D. Sanghi, A. K. Agrawala, O. Gudmundsson, B. N. Jain, "Experimental Assessment of End-to-end Behaviour on Internet", *INFOCOM'93*, 1993.

[11] K. Thompson, G. J. Miller, R. Wilder, "Wide-Area Internet Traffic Patterns and Characteristics", *IEEE Network*, November-December, 1997.