ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/228892000

Synthesis of network delays for voice packets in Service Overlay Networks

Article · August 2007

DOI: 10.1145/1577222.1577229

CITATIONS
8

READS
20

2 authors:

Hong (Cindy) Li
McGill University

6 PUBLICATIONS **30** CITATIONS

SEE PROFILE

L.G. Mason
McGill University

88 PUBLICATIONS **1,677** CITATIONS

SEE PROFILE

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.

Available from: L.G. Mason
Retrieved on: 09 May 2016

Synthesis of network delays for voice packets in Service Overlay Networks

Hong Li
Electrical and Computer Engineering Department
McGill University
Montreal, Quebec, Canada H3A 2T5
hong.li2@mail.mcgill.ca

Lorne Mason
Electrical and Computer Engineering Department
McGill University
Montreal, Quebec, Canada H3A 2T5
lorne.mason@mcgill.ca

ABSTRACT

Toll quality Voice-over-IP (VoIP) provision over Service Overlay Networks (SONs) is still a big challenge for the current best-effort Internet. The end to end network delay, as one of the most important performance concerns, is the focus in this paper. In the paper, network delay characteristics, e.g. probability distribution and autocorrelation are studied. We propose an approach to synthesize network delay traces for QoS routing research in a full mesh Service Overlay Network(SON) given only partial information of the network delay trace. The main contributions of the paper include: 1) A parametric model of the network delay; 2) Method to synthesize network delay traces for voice packets sent over SONs. The synthesized network delay traces are shown to be close to real network delay traces. The synthetic network delay traces can be used to study quality of service provision mechanisms for VoIP applications over the Service Overlay Networks.

Categories and Subject Descriptors

I.6 [**Simulation and Modeling**]: Miscellaneous

General Terms

Performance

Keywords

Voice-over-IP, Service Overlay Network, Internet delay

1. INTRODUCTION

With the Next Generation Internet (NGI) still in its infancy, the current best-effort Internet has to face the challenges of increasing real-time multimedia communication applications, e.g. Voice-over-IP(VoIP), on-line games, video conferences, etc. We are especially interested in the VoIP application. Transmitting voice packets over heterogeneous networks at toll quality level is a big challenge for VoIP service provision over current best-effort Internet. VoIP has strict requirements on network performance impairments, i.e. network delay, delay jitter and loss.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
QSHINE '07, August 14-17, 2007, Vancouver , Canada
Copyright 2007 ACM 978-1-59593-756-8 ...$5.00.

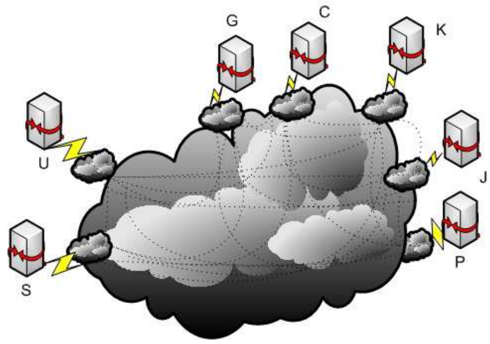


Figure 1: Service Overlay Network.

The current Internet is basically a hierarchy of interconnected Autonomous Systems (ASes), whereby no single ISP has control of the end-to-end service quality. Indeed a given VoIP call might originate within a wireless LAN offering only a best effort service (no QoS guarantees), then transit a local ISP before reaching a tier one or backbone ISP which implements MPLS/DiffServ, before going down the hierarchy of local ISP, and perhaps an ADSL loop and then reaching a standard fixed telephone. The packet delay, loss and jitter are introduced within each segment of the end-to-end path with no one being responsible for the end-to-end quality of the call.

An overall quality evaluation of VoIP service with E-model is recommended by ITU-T G.107[1], in which the end-to-end network delay is one of the most significant network impairments that determine the quality for VoIP. Therefore, we focus on studying network delay in the paper.

Network traffic modeling and synthesis have been studied extensively. However, network delay modeling and synthesis are not studied a lot, although network delay characteristics have been studied by some research[5][6]. In the paper we try to synthesize delay traces for research on QoS mechanisms which employ trace based simulations. Similar work SS-SVM [3] is proposed as a non-parametric method that generates network delays. However, the synthetic delay trace generated by SS-SVM is only statistically similar to the real trace in terms of the marginal probability distribution of real delays, and there is no explicit comparison between the synthetic delay trace and the real delay trace in the paper [3]. Our work is different from SS-SVM in that we developed a parametric model of the end-to-end network delay in SONs by fitting various distributions to network delay traces measured by sending UDP packets periodically in a SON, and we estimated the parameters of the model with the statistics of the network delay trace. We also synthesized new delay trace for voice packets in the SON with the network delay model. A comparison of the synthesized network delay trace with the real network delay trace is shown at the end of the paper.

2. MODELING THE PACKET DELAY IN SERVICE OVERLAY NETWORKS

Service Overlay Networks, as shown in Fig. 1, have been employed for applications such as VoIP and multi-media communications as a cost effective approach to provide value-added services over the current best-effort Internet. Recently a variety of SON architectures for VoIP have been reported[2][4].

The most important impairments for a VoIP phone call in the SON is the network delay. We use network delay measurements of a network service provider. They are collected by sending UDP probing packets to the network service provider's SON periodically per 30ms among seven globally located monitors. A total of forty-two SON paths and seven service overlay gateways are involved, which form a small full mesh SON network. We study the autocorrelation and the probability distribution of the network delays experienced by the UDP probing packets, and construct a parametric model for the packet level network delay which is given in the section 2.1.

2.1 Probability distribution

One-way network delay is composed of three parts: propagation delay, transmission delay and queueing delay. We consider the propagation delay, which is constant assuming the path does not vary during the whole measurement period, as the minimum delay that a voice packet will experience. The transmission delay considered as negligible since it is very small for a high speed link. and we consider the queueing delay is a random variable which follows a certain distribution. By fitting various distributions to real measurements on various links at various time, we find that shifted gamma distribution is a good approximation. As shown in fig.2 visually, gamma distribution fits best to the 100 queuing delay samples (it is also true for large quantity of delay samples). The Weibull distribution also gives good fit. The Chi-square test on goodness of fit also shows that gamma distribution is a good fit. We use the gamma distribution to model queueing delay and the total network delay is modeled by a shifted gamma distribution. The Probability Distribution Function (PDF) of a random variable X with shifted gamma distribution is: $f(x) = \frac{(\frac{x-\theta}{\beta})^{(\gamma-1)} \cdot exp(-\frac{x-\theta}{\beta})}{\beta \cdot \Gamma(\gamma)}; x \geq \theta; \theta, \gamma, \beta > 0$, where γ and β are the shape parameter and scale parameter respectively, and θ represents the minimum network delay experienced by a voice packet. Therefore the mean network delay is $\mu = \theta + \gamma\beta$, the standard deviation of network delay is $\sigma = \sqrt{\gamma}\beta$, and the skewness is $k = \frac{2}{\sqrt{\gamma}}$. It is easy to estimate parameters (θ, γ, β) given samples of network delay traces with maximum likelihood or MAP method. However, it is difficult to infer the parameters with less than three statistics of the network delay trace.Section 3 gives methods on the inference of the parameters.

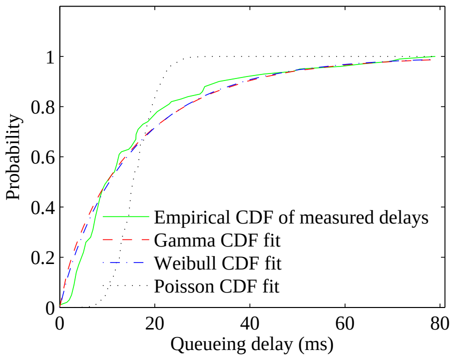


Figure 2: Cumulative Distribution Function fit of 100 queueing delay samples.

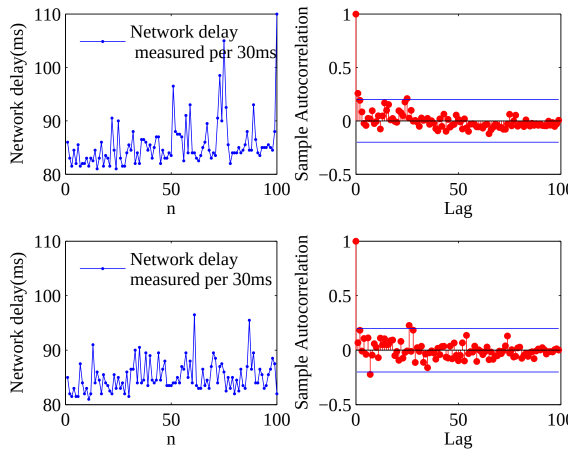


Figure 3: Network delay measurements showing weak autocorrelation and no autocorrelation

2.2 Autocorrelation

The autocorrelation coefficient of lag k of a time series Y_i, i=1,Ě,N is given in the formula:

$$r_k = \frac{\sum_{i=1}^{N-k}(Y_i - \bar{Y})(Y_{i+k} - \bar{Y})}{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}$$

where $\bar{Y}$ is the average of Y_i,i=1,...,N. We differentiate time series according to r_1.The time series has strong autocorrelation if $|r_1| > 0.75$; it has medium autocorrelation if $0.75 > |r_1| > 0.5$; it has weak autocorrelation if $0.5 > |r_1| > 0.2$; it is random if $|r_1| < 0.2$. Delay spikes occur when traffic arrives in short or long bursts. We find short burst in traffic will only lead to random delay spikes. By plotting the autocorrelation function of the network delay measurements on different links at different time, we find that the autocorrelation is usually negligible when no delay spikes or when delay spikes occur randomly, as shown in Fig. 3.

Assuming the time series in Fig. 3 are stationary random processes, we can use an identical and independent shifted gamma distribution model it when the time series has no spike or random spikes.

If there is strong or medium autocorrelation in the random process, we need to modify the model to add autocorrelation in. A possible model is to use AutoRegressive (AR) model with a residual that follows the gamma distribution.

3. PARAMETER ESTIMATION FOR NETWORK DELAY TRACES

We have only one statistics of the network delay trace, i.e. the sample mean of network delay. In order to reconstruct a synthetic which is statistically similar to the original network delay trace from the sample mean of network delay, we first need to estimate the parameters. We can then reconstruct a statistically similar network delay trace with the same mean network delay as the real measured network delay trace.

Rewrite the problem in a formal way as: reconstruct a stationary and ergodic random sequence X_i, i=1,...,M, where X_i is independently and identically distributed with a shifted gamma distribution with parameter (θ, γ, β) given statistics of network delay. Possible statistics that one can collect are the sample mean of N samples $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$, the unbiased standard deviation of N samples $S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})}$, and the minimum network delay D of the path. It seems impossible to estimate three parameters from a single statistic, however, we can study the correlation between the delay statistical parameters, and infer other statistics from the partial information, i.e. the sample mean, and the estimate the parameters (θ, γ, β) from the statistics.

3.1 Parameter estimation for modeling network delay traces

We now solve the problem under different conditions on how many statistics we have: (1) We have statistics on the sample mean $\bar{x}$, the sample standard deviation S and the minimum network delay D;(2)We have only $\bar{x}$ and D; (3) We have only $\bar{x}$. The method proposed here will be used to synthesize network delay in the next section.

3.1.1 The sample mean $\bar{x}$,the standard deviation S and the minimum network delay D are known

Suppose we have known the sample mean $\bar{x}$ and sample standard deviation S, we can estimate the real mean and real standard deviation with $\hat{\mu} = \bar{x}$ and $\hat{\sigma} = S$, which gives

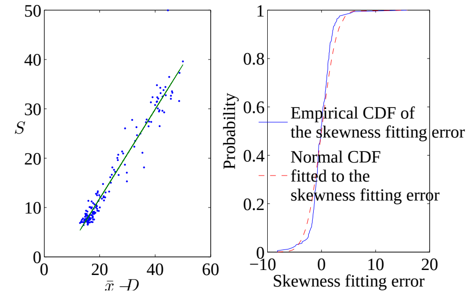
$$\hat{\mu} = \theta + \gamma\beta = \bar{x} \tag{1}$$
$$\hat{\sigma} = \sqrt{\gamma}\beta = S \tag{2}$$

And if we can have the measured minimum delay D, we can estimate θ. Then we can synthesize the original random sequence.

$$\hat{\theta} = D \tag{3}$$

3.1.2 Only the sample mean $\bar{x}$ and the minimum delay D are know

In this case, we miss the sample standard deviation S, which means we have to infer it from $\bar{x}$ and D. Let the estimator of S be $\hat{S}$, the estimator of sample mean be $\bar{X}$ and the estimator of minimum delay be $\hat{D}$. The best estimator of sample standard deviation given sample mean is $\hat{S} = E\{S|\bar{X}\}$,which gives minimum mean square error. However, the conditional distribution is hard to calculate. An alternative way that is easier to implement is linear



(a) (b)

Figure 4: Linear estimation of the sample standard deviation S (ms) with $(\bar{x} - D)$ (ms), (a) Linear fit of skewness (each point in the figure is $\frac{S}{\bar{x}-D}$ of 100 samples) (b)CDF fit of skewness fitting error

estimation. One can find the best homogeneous or non-homogeneous linear estimator of a random variable Y from another random variable X provided that the covariance of X and Y and the variance of Y are given. However, we don't have these statistics available. Therefore, we choose to use the linear regression method to estimate Y from X. We study the correlation between sample standard deviation S and $(\bar{x} - D)$, and fit their relation to a line as shown in Fig. 4(c).

Indeed, the slope of linear fit shown in Fig.4(a) gives the skewness of the shifted gamma distribution which is $k = \frac{1}{\sqrt{\gamma}}$. And as shown in Fig.4(b), the error of linear fit is a Gaussian distributed random variable. The linear fit of skewness shows that the shape of the gamma distribution does not change much along time. It verifies the assumption that the random process is approximately ergodic in distribution. In cases where there are some outliers, is due to the fact that the delay trace is not stationary and ergodic during the whole measurement period. With the assumption of stationarity and ergodicity, we can estimate the sample standard deviation S from $(\bar{x} - D)$ if we have the estimate of the skewness $\hat{k}$.

$$\hat{\sigma} = \hat{k}(\bar{x} - D) \tag{4}$$
$$\frac{1}{\sqrt{\hat{\gamma}}} = \hat{k} \tag{5}$$

One thing to note is that the skewness $\hat{k}$ varies on different links as shown in Fig. 5. We also find the slope α of the linear fit of each path is highly correlated with the mean delay of that path.

If we do not assume stationarity and ergodicity, i.e. the relationship between S and $\bar{x} - D$ is not always linear, we use the equation (6) to estimate $\hat{\sigma}$, in which a parameter α can be tuned to be mostly constant and to vary with $\bar{x} - D$ for outliers shown in Fig. 5.

$$\hat{\sigma} = \alpha * (\bar{x} - D) + \omega \tag{6}$$

ω is a very small Gaussian noise. With $\hat{\sigma}$, $\bar{x}$ and D, we can synthesize network delay on a link according to the network delay model defined in section 2.1.

3.1.3 Only the sample mean $\bar{x}$ is known

In the case that we have only the sample mean delay $\bar{x}$, it is even harder to estimate the three parameters. By the

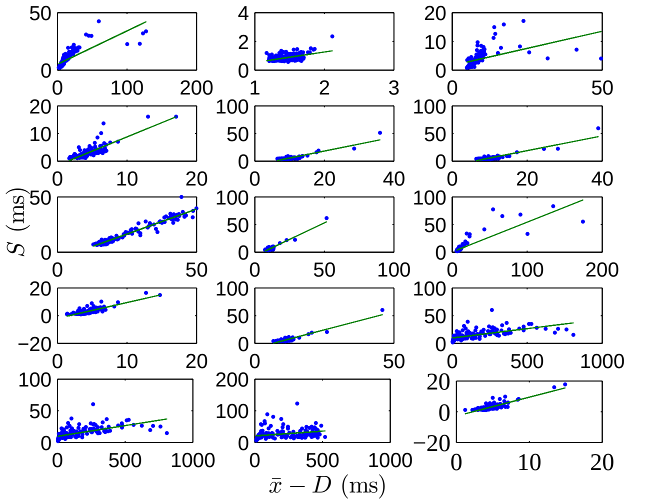


Figure 5: Linear fit for the skewness on 15 different links ($\bar{x} - D$ is the average queueing delay per 100 network delay samples, S is the standard deviation per 100 network delay samples)

derivation of the previous two cases, we will just estimate D in this case, and apply the same method in the last case to estimate all parameters. The minimum delay and length of a path are linearly related. However, the correlation between the minimum delay and geographical distance is not very high since paths do not necessarily go through the shortest route between end points. It is difficult and even infeasible to measure real link distances in terms of meters of cables due to legacy issues. An alternative approach is to estimate minimum delay by using the minimum of sample mean delays in the night and dawn to estimate it. Then we can estimate other parameters using the methods previously discussed.

4. NETWORK DELAY SYNTHESIS

We want to simulate VoIP phone calls over a full mesh Service Overlay Network (SON). The Service Overlay Gateways are distributed through the world. We only have partial network delay measurements from a network service provider. We have mean delay measurements per 100 samples during a week for a full mesh SON, and network delay samples with a sampling interval of 30ms between a single source and all other nodes in the SON. Therefore, we have to infer the network delay samples for the whole Service Overlay Network from the partial information of the network.

Given only the sample mean delay, we must estimate the slope of the linear fit as shown in Fig. 4 and estimate the minimum delay of the path with the method in section 3.1.2 and 3.1.3 respectively. Then we can estimate parameters of the shifted gamma distribution model with methods in section 3.1.1. Then we can synthesize network delay traces with the model described in section 2. Fig. 6 shows that the synthetic network delay trace is similar to the real measured trace.

The synthetic delay trace can be used to evaluate QoS routing mechanisms for VoIP in Service Overlay Networks. It can give similar quality score compared to that given by the real trace.

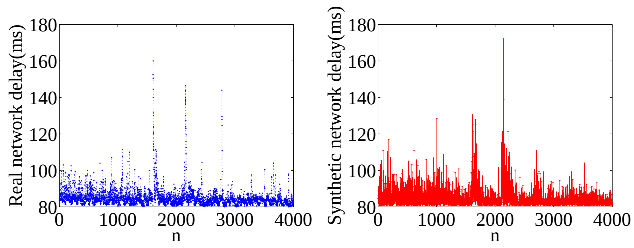


Figure 6: Real delays vs. synthetic delays(4000 samples).

5. CONCLUSION

In the paper, a shifted gamma distribution model is fitted to the network delay measurements and a network delay synthesis method is proposed. The method is able to synthesize network delay traces, including random delay spikes, given only the sample mean of network delay . It is validated by comparing the network delay traces synthesized with the model to the real network delay traces. The synthesized delay traces can capture the changes of network delays which include peaks in the office hours and dips in the evening and the dawn. The synthesis method is an easy way to get network delay traces for evaluating QoS provision mechanisms for VoIP phone calls in our research. It saves substantial cost of gathering traces, especially for large SONs with hundreds of nodes. Therefore, it saves the substantially high cost of network delay measurements for research based on network delay matrices.

The method proposed in the paper can synthesize real network delay traces given partial statistics (i.e. only the sample mean) of the network delay. Future work will be on developing more sophisticated model for strongly autocorrelated delay traces and evaluating the error between synthetic delay trace and real delay trace.

6. ACKNOWLEDGMENTS

The authors would like to thank NSERC for supporting this research.

7. REFERENCES

[1] The e-model, a computational model for use in transmission planning. *ITU-T Recommendation G.107*, March 2003.

[2] S. Baset and H. Schulzrinne. An analysis of the skype peer-to-peer internet telephony protocol. *Technical Report CUCS-039-04,Columbia University*, 2004.

[3] J. Hernandez, I. Phillips, and J. Moguerza. An ss-svm approach to generate synthetic network delays. In *Proceedings 12th International Conference ASMTA*, 2005.

[4] V. Hilt, A. Hari, and M. Hofmann. An efficient and robust overlay routing scheme for voip. In *ICICS 2005*, pages 508–512, December 2005.

[5] A. Markopouloua, F. Tobagib, and M. Karam. Loss and delay measurements of internet backbones. *Computer Communications*, 29:1590–1604, 2006.

[6] A. Mukherjee. On the dynamics and significance of low frequency components of internet load. *Internetworking: Research and Experience 5*, 4:163–205, 1994.