# Estimation and Simulation of Network Delay Traces for VoIP in Service Overlay Network

Hong Li* and Lorne Mason
Electrical & Computer Engineering,
McGill University, Montreal, Quebec H3A 2A7
Email: hong.li2@mail.mcgill.ca*, lorne.mason@mcgill.ca

*Abstract*— **The network delay is one of the most important impairments for VoIP (Voice Over IP). This paper focuses on the bi-objective optimization problem of minimizing sampling cost and also minimizing the error in the estimation of the actual delay traces. We also analyze the real network delay measurements and the simulated network delay trace for fractional Brownian motion ( fBm ) traffic. We find that a shifted Gamma distribution models the network delay marginal distribution very well. The simulation based on the simulated network delay trace shows that our novel optimal sampling strategy can obtain a 20dB Signal to Noise Ratio(SNR) in measurement accuracy at as low as 1.3HZ sampling frequency.**

**Index Terms**: signal sampling, optimization methods, signal reconstruction.

## I. INTRODUCTION

Toll quality VoIP ( Voice Over IP ) provision is of great interest in migrating current voice service to a converging next generation network. With current best effort Internet, it is a cost efficient way to provide VoIP service over a Service Overlay Network (SON), as shown in Fig. 1. As we know, an interactive level voice call requires a maximum of 150ms of end to end network delay, which is a big challenge for long distance phone calls in peak hours. In our research, we focus on the study of the end to end network delay experienced by voice packets in a Service Overlay Network, including an optimal sampling approach for monitoring the end to end network delay and the analysis of network delay measurements.
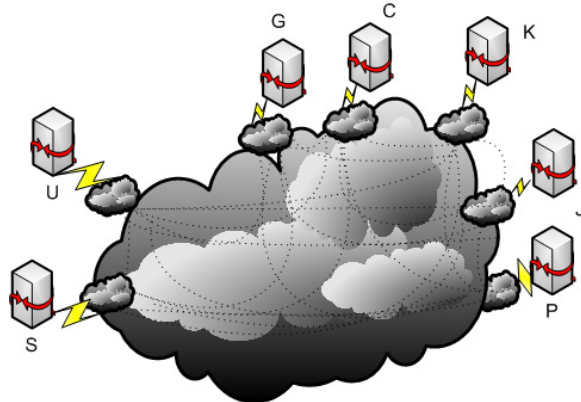
The motivation for this study is the application of monitoring in QoS routing for VoIP in SONs. We sent UDP packets periodically to the SON to measure the network delay. The monitoring process will add additional overhead on the network. In order to minimize the cost of monitoring and also minimize the estimation error of the network delay, we solve the bi-objective optimization problem to find the fair compromise between the conflicting objectives of sampling cost minimization and sampling accuracy maximization.

## II. OPTIMIZATION OF THE SAMPLING ACCURACY AND THE SAMPLING COST

### A. Definition

We denote the network delay trace as x(n). It's Fourier transform X(f) is assumed to be band limited with cut-off frequency $F_m$ and X(f) is differentiable everywhere. Let $|X(f)|$ be the amplitude spectrum of x(n) and $\Phi(f) = |X(f)|^2$ be the power spectrum of x(n), where $\Phi(f) = \Phi(-f)$ for the real valued x(n).

Signal to noise ratio (SNR) is defined as the ratio of the total energy in x(n) to the noise introduced by aliasing when we sample the network delay at frequency $2 * f$. Therefore the SNR is a function of $f$, where

$$SNR(f) = \frac{\int_{-F_m}^{F_m} \Phi(u) du}{\int_{-F_m}^{-f} 2\Phi(u) du + \int_{f}^{F_m} 2\Phi(u) du}. \quad (1)$$

The sampling cost $c(f)$ is represented by a monotonically increasing function of the sampling frequency. We use the sampling frequency as the sampling cost for simplicity. Let the sampling error $Er$ be the noise-to-signal ration $\frac{1}{SNR}$ and the sampling accuracy $Ar$ be $1 - Er$.

The maximization of the SNR or the sampling accuracy conflicts with the minimization of the sampling cost or the sampling frequency. Therefore, the problem is to find the minimum sampling frequency that maximizes the sampling accuracy.

### B. Problem formulation

As defined in section II-A, sampling accuracy $Ar$ is a function of the sampling frequency. The higher the sampling frequency ( $2 * f \leq 2 * F_m$ ), the higher the sampling cost $c(f)$ and the higher the sampling accuracy $Ar(f)$, where $c(f) = f$ and $Ar(f) = 1 - \frac{1}{SNR(f)}$. Let the optimal sampling frequency be $2f^*$ and the optimal sampling accuracy be $Ar^*$.



Fig. 1. Service Overlay Network.

The problem of maximizing the sampling accuracy $Ar(f)$ while minimizing the sampling cost c(f) can be formulated as a bi-objective optimization problem given in (2).

$$\begin{aligned} max_f \quad & Ar(f) \\ min_f \quad & c(f) \\ s.t. \quad & 0 < f \le F_m \end{aligned} \tag{2}$$

### C. Solution to the bi-objective problem

*1) Related work:* The traditional method to solve multi-objective optimization problems is weighted sum method [2], which reduces the multi-objective optimization problem to a single-objective optimization problem. A Pareto front or a Pareto surface is approximated by changing the weights among objective functions. Many other methods [3][5] are also proposed to find a even Pareto front or Pareto surface. However, those methods can only give a Pareto front or Pareto surface that only satisfies Pareto efficiency.

*2) Our solution:*

*Lemma 1:* The optimal solution $f^*$ to the bi-objective optimization problem (2) is a Pareto efficient solution.

*Proof:* Supposing the sampling cost function $c(f)$ and the sampling accuracy function $Ar(f)$ are differentiable, we can write $c'(f^*) = \lim_{\Delta f \to 0} \frac{c(f^* + \Delta f) - c(f^*)}{\Delta f} < 0$. Similarly, $Ar'(f^*) = \lim_{\Delta f \to 0} \frac{Ar(f^* + \Delta f) - Ar(f^*)}{\Delta f} > 0$. Now consider two situations:
(a) If we increase $f^*$ by $\Delta f$, $\Delta f > 0$, and $0 < f^* + \Delta f \le F_m$, then $c(f^* + \Delta f) > c(f^*)$ and $Ar(f^* + \Delta f) > Ar(f^*)$, i.e. $c(f)$ becomes worse off if we make $Ar(f)$ better off by increasing the sampling frequency.
(b) Similar to (a), if we decrease $f^*$ by $\Delta f$, $\Delta f > 0$, and $0 < f^* - \Delta f \le F_m$, then $c(f^* - \Delta f) < c(f^*)$ and $Ar(f^* - \Delta f) < Ar(f^*)$, i.e. $Ar(f)$ becomes worse off if we make $c(f)$ better off by decreasing the sampling frequency. In sum, no single objective can be better off without making the other objective worse off, i.e. the optimal solution $f^*$ is a Pareto efficient solution. ∎

*Lemma 2:* The optimal solution $f^*$ to the bi-objective optimization problem (2) is invariant to equivalent representation of $Ar(f)$ and $c(f)$.

*Proof:* Let $\tilde{A}r(f)$ be $\alpha_1 Ar(f) + \beta_1$ and $\tilde{c}(f)$ be $\alpha_2 c(f) + \beta_2$. The transformation maintain the order of preference in $Ar(f)$ and $c(f)$. The original bi-objective optimization problem (2) becomes:

$$\begin{aligned} max_f \quad & \tilde{A}r(f) \text{ or } \alpha_1 Ar(f) + \beta_1 \\ min_f \quad & \tilde{c}(f) \text{ or } \alpha_2 c(f) + \beta_2 \\ s.t. \quad & 0 < f \le F_m \end{aligned} \tag{3}$$

Let a Pareto efficient solution to the problem (2) be $f^*$. It is easy to see $\tilde{f}^* = f^*$ is a Pareto efficient solution to the problem (3) since the order of preference in $\tilde{A}r(f)$ and $\tilde{c}(f)$ is the same as that in $Ar(f)$ and $c(f)$. ∎

Rewrite the bi-objective optimization problem as:

$$\begin{aligned} max_f \quad & \begin{bmatrix} Ar(f) \\ max(c(f)) - c(f) \end{bmatrix} \\ s.t. \quad & 0 < f \le F_m \end{aligned} \tag{4}$$

*Lemma 3:* The proportional fair optimal solution $f^*$ to the bi-objective optimization problem (2) maximizes the product of the two objectives in (4).

*Proof:* As proven in [6][7], the proportional fair allocation can be obtained by $max \sum_s ln(x_s)$ over the set of feasible allocations, where $x_s$ is the share allocated to user $s$.
Then the proportional fair solution to the problem (4) or (2) is given by:

$$\begin{aligned} max_f \quad & ln(v_1) + ln(v_2) \\ v_1 \quad & = Ar(f) \\ v_2 \quad & = max(c(f)) - c(f) \\ s.t. \quad & 0 < f \le F_m \end{aligned} \tag{5}$$

And it is easy to see the solution to (5) is also pareto efficient. ∎

If we formulate the bi-objective optimization problem as a bargaining problem between two players, with the set U of payoff under agreement being the points interior to and on the boundary of $Ar(f)$ and the payoff d under disagreement being (0,0), we can obtain the same result with Nash bargaining solution [8] as that of (5).

## III. MODELING OF THE NETWORK DELAY

To estimate the network delay trace x(n), we need to know the Cumulative Distribution Function (CDF) of x(n),which will be shown in the next section.

One-way network delay is composed of three parts: propagation delay, transmission delay and queuing delay. We consider the propagation delay as the minimum delay that a voice packet experiences, which is constant assuming the path does not vary during our measurement, and the transmission delay is very small to be negligible. The queuing delay is represented as a random variable that follows a certain distribution. By fitting various distributions to the real measurements, which is obtained from a SON of a network operator, we find that shifted Gamma distribution is a best approximation to the empirical distribution of the real data compared to other distributions.

As shown in Fig.2, Gamma distribution fits best to the random part of measured network delays, and Weibull distribution also gives good fit. The Chi-square test on goodness-of-fit also shows that Gamma distribution is a good fit. Therefore, we use Gamma distribution to model queuing delay and the total network delay is modeled by a shifted Gamma distribution. This model can be used to simulate new network delay traces.

## IV. SIMULATION

We generated a stationary network delay trace $x(n)$ with fBm traffic, as shown in the top figure of Fig. 4. The parameters of the fBm traffic from Bellcore data OctExtTL
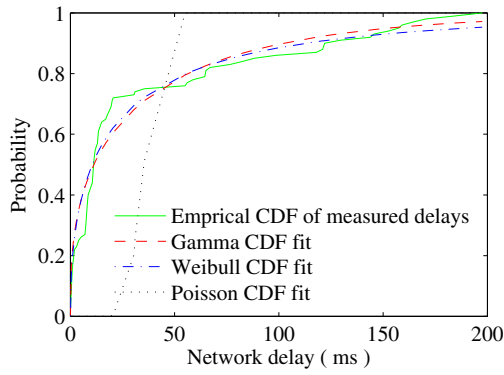
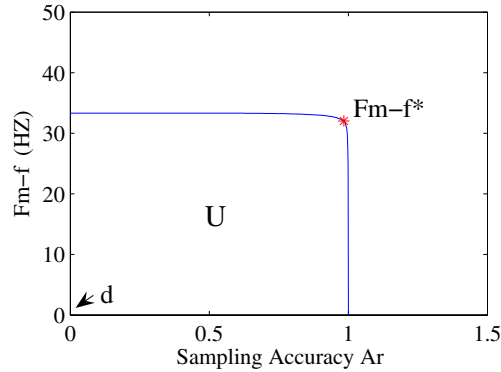Fig. 2. Distribution fitting of measured samples.



Fig. 3. Trade off between sampling frequency and sampling accuracy.



Fig. 4. Network delay experienced by a probe.

are as follows: Hurst parameter H=0.86, mean input rate m=12.3kbit/sec, and variance coefficient a=68.6kbit· sec [1]. We fit the generated network delay to various distributions and find that shifted Gamma distribution gives the best fit compared to other distributions. Then we apply the optimal sampling method in section II to sample the generated network delay trace with fBm traffic, as shown in the top of Fig. 4. The sampling accuracy function Ar(f) is a concave function as shown in Fig. 3.

Therefore, the theoretical optimal sampling frequency is given by the point f* on the curve that maximize the product of the two coordinates. The resulting optimal sampling frequency given by (5) is f*=1.3HZ which reduces the original sampling frequency by about 30 times. The reconstructed signal is shown in the bottom right of the Fig. 4. The corresponding SNR in the reconstructed signal is about 20dB.

## V. CONCLUSION

We have formulated the optimal sampling problem as a bi-objective optimization problem of minimizing the sampling frequency while maximizing the sampling accuracy which is desired in an Internet monitoring scheme. We show that the proportional fair and pareto optimal solution to the bi-objective optimization problem is given by the point that maximize the product of the objectives, which is equivalent to a Nash
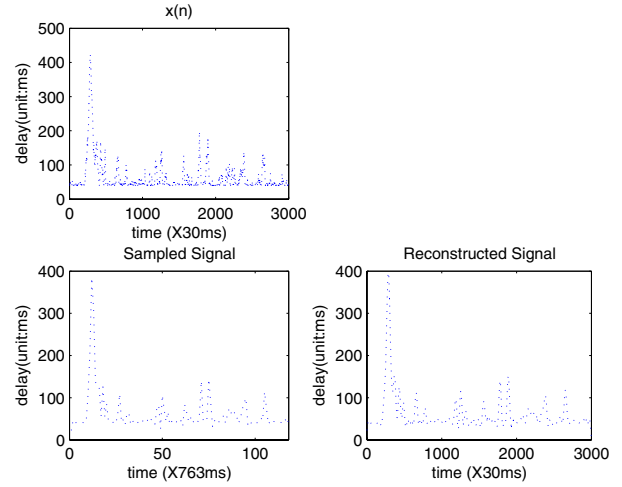
bargaining solution. Our method guarantees accurate estimation and reconstruction of the actual delay trace at lowest sampling cost. The results show that significant savings in monitoring cost (as quantified by the sampling traffic volume) is possible. In addition, the statistical analysis made on real measurements has shown that a shifted Gamma distribution adequately captures the marginal distribution of the delay traces, for subsequent data simulation.

## REFERENCES

[1] I. Norros,"On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks," *IEEE Journal on selected areas in communications*, Vol. 13, No. 6, August 1995.
[2] I.Y. Kim and O.L. de Weck, "Adaptive weighted-sum method for bi-objective optimization: Pareto front generation," *Struct Multidisc Optim*, Vol. 29, pp. 149158, 2005.
[3] J. Rakowska, R. T. Haftka and L. T. Watson, "Tracing the Efficient Curve for Multi-Objective Control-Structure Optimization," *Computing Systems in Engineering*, Vol. 2, No. 6, pp. 461-471, 1991.
[4] IY Kim and OL de Weck, "Adaptive weighted-sum method for bi-objective optimization: Pareto front generation," *Structural and Multi-disciplinary Optimization*, 2005.
[5] I. Das and J. E. Dennis, "Normal-Boundary Intersection: An Alternate Approach for Generating Pareto-optimal Points in Multicriteria Optimization Problems," *SIAM J. on Optimization*, Vol. 8, pp. 631–657, 1998.
[6] H. Yaiche, R.R. Mazumdar, and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks", *IEEE/ACM Trans. on Networking*, Vol. 8, No. 5, pp. 667-678, Oct. 2000.
[7] R. Mazumdar, L.G. Mason, and C. Douligeris, "Fairness in network optimal flow control: Optimality of product form," *IEEE Transactions on Communication*, Vol. 39, pp. 775782, 1991.
[8] M.J. Osborne and A. Rubinstein, *A course in game theory*, July 1994.