

Online end-to-end quality of service monitoring for service level agreement management

Xiaoyuan Ta^{*,†} and Guoqiang Mao

School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia

SUMMARY

A major challenge in network and service level agreement (SLA) management is to provide Quality of Service (QoS) demanded by heterogeneous network applications. Online QoS monitoring plays an important role in the process by providing objective measurements that can be used for improving network design, troubleshooting and management. Online QoS monitoring becomes increasingly difficult and complex due to the rapid expansion of the Internet and the dramatic increase in the speed of network. Sampling techniques have been explored as a means to reduce the difficulty and complexity of measurement. In this paper, we investigate several major sampling techniques, i.e. systematic sampling, simple random sampling and stratified sampling. Performance analysis is conducted on these techniques. It is shown that stratified sampling with optimum allocation has the best performance. However, stratified sampling with optimum allocation requires additional statistics usually not available for real-time applications. An adaptive stratified sampling algorithm is proposed to solve the problem. Both theoretical analysis and simulation show that the proposed adaptive stratified sampling algorithm outperforms other sampling techniques and achieves a performance comparable to stratified sampling with optimum allocation. A QoS monitoring software using the aforementioned sampling techniques is designed and tested in various real networks. Copyright © 2007 John Wiley & Sons, Ltd.

Received 6 November 2006; Revised 4 May 2007; Accepted 4 May 2007

KEY WORDS: QoS; SLA; network management; packet delay; LMS; adaptive sampling

1. INTRODUCTION

With the rapid growth of Internet in scale and complexity, providing Quality of Service (QoS) in Internet becomes increasingly important for network design, traffic engineering, troubleshooting

*Correspondence to: Xiaoyuan Ta, School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia.

†E-mail: xiaoyuant@ee.usyd.edu.au

Contract/grant sponsor: Optus Ltd.; contract/grant number: BLO-2004-7260

Contract/grant sponsor: Australian Research Council Discovery Grant; contract/grant number: DP0559248

and service level agreement (SLA) management. The need for accurate network performance measurement and monitoring is increasing for both network management and SLA validation. Online QoS measurement plays an important role in the process. It is not only useful for network operators who want to keep track of the performance of their network, but also useful for the individual customer who wants an objective test on whether QoS provided by the Internet service providers (ISPs) is satisfactory. Moreover, having the ability to measure against key performance indicators facilitates the continuous quality improvement process. It helps the ISPs to locate the bottleneck of the network and to properly allocate network resources to achieve a prescribed QoS for all network users [1]. A service performance problem becomes an opportunity to structurally improve overall service quality and customer satisfaction.

There has been significant work on developing mechanisms and algorithms for online network performance monitoring and traffic measurements. This includes the work of Internet Engineering Task Force (IETF) groups such as IP Performance Metrics (IPPM) group which has proposed the framework for IP performance measurement, and other organizations such as [2–6]. Also, many papers have been published on network performance measurement and analysis, such as [7–16].

However, online QoS measurement becomes increasingly difficult and complex due to the rapid expansion of the Internet. Moreover, the dramatic increase in the speed of wide area backbones presents obstacles to data collection. The enormous amount of measurement data may significantly increase cost and resource usage [17]. Sampling-based measurement methods have been explored to reduce the amount of control data and resources required for network performance monitoring, and finally to reduce the measurement complexity and cost. The principle of sampling techniques is to investigate the characteristics of a population of elements using a representative subset. In network performance monitoring, the performance metrics (e.g. packet delay, packet loss and jitter) are computed by choosing some particular packets among the entire traffic in the network. In this paper, theoretical analysis is conducted on the performance of several major sampling techniques, i.e. systematic sampling, simple random sampling and stratified sampling. It is shown that stratified sampling with optimum allocation has the best performance. However, stratified random sampling with optimum allocation requires extra statistics from the parent traffic trace, which are not known *a priori* in real applications. To address the problem, a novel adaptive sampling strategy is proposed, which employs a least-mean-square (LMS) algorithm to predict the required statistics from past observations. The proposed strategy is able to significantly reduce the amount of measurement data required. Theoretical analysis is performed on the performance of the strategy. Simulation using real traffic trace is also performed. Both the theoretical analysis and simulation results demonstrate the accuracy of the proposed strategy. Software is designed, which implements the proposed technique.

In this paper, we focus on delay measurement because network delay is a key performance metric in SLAs and has significant impact on a number of delay-sensitive applications, such as VoIP (Voice over IP) and video traffic [13]. However, the proposed strategy can be readily extended to measuring other QoS metrics such as packet loss and jitter.

The rest of this paper is organized as follows. Section 2 outlines related work. Section 3 introduces three sampling techniques with a qualitative discussion on their advantages and disadvantages. Section 4 quantitatively analyses the performance of these sampling techniques. Section 5 describes in detail the proposed adaptive stratified sampling scheme. Section 6 introduces the traffic trace used for simulation. Section 7 presents the simulation results using real traffic traces provided by the WAND group. Section 8 introduces the QoS monitoring software designed by us, and finally Section 9 concludes this paper.

2. RELATED WORK

Sampling techniques have been widely applied to network performance measurements [1, 17–26], and efficient adaptive sampling schemes have also been developed.

Claffy *et al.* [17] investigated the performance of various sampling methods related to wide area network traffic characterization. Their simulation results revealed that the time-triggered sampling techniques do not perform as well as the packet-triggered ones. However, the performance differences between packet-based sampling techniques and their time-based counterparts are small. In [1], several adaptive sampling methods were developed and evaluated to address inaccuracies of static (or conventional) sampling. It was shown that the adaptive sampling methods achieve more accurate estimates of the mean, variance and the Hurst parameter (a measure of traffic self-similarity). In [19], two adaptive sampling methods were proposed based on linear prediction and fuzzy logic, respectively. The performance of these techniques was then compared with conventional sampling methods by conducting simulative experiments using Internet and videoconference traffic patterns. The simulation results showed that the proposed adaptive techniques are significantly more flexible in their ability to dynamically adjust to fluctuations in network behaviour, and in some cases they are able to reduce the sample count by as much as a factor of 2 while maintaining the same accuracy as conventional sampling methods. Zseby [21] investigated the deployment of sampling techniques for SLA validation focusing especially on the application of sampling to non-intrusive end-to-end measurements. Experiments are performed using systematic, random and stratified sampling. Tests with stratified sampling showed how the estimation accuracy can be improved if *a priori* information was available. Zseby [22] investigated the stratification strategy to improve estimation accuracy without increasing the sample size. It was shown that the sample size could be significantly reduced if packets were stratified according to their size. Furthermore, adaptive schemes with a dynamic adjustment of the stratification boundaries were compared to schemes with fixed boundaries. Zseby and Zander [23] investigated SLA validation for highly interactive applications such as multiplayer online games. They proposed a novel solution for passive SLA validation based on direct a sampling of the customer traffic. They also presented an elegant solution for estimating the sampling error prior to the sampling and for computing the minimum sample rate required depending on SLA parameters. Ma *et al.* [26] proposed a new adaptive sampling scheme for monitoring network performance given the knowledge of the traffic type, i.e. voice traffic. They then compared the proposed adaptive sampling scheme with systematic sampling and stratified sampling methods through simulation using voice traffic. They showed that the proposed adaptive scheme achieved the best accuracy on voice traffic.

3. SAMPLING TECHNIQUES

Traditional sampling techniques can be classified into three categories: systematic sampling, random sampling and stratified random sampling [17, 21, 22]. Figure 1 illustrates these three sampling techniques.

3.1. Systematic sampling

Systematic sampling generates sampling traffic according to a deterministic function. Generation of the sampling traffic is triggered by either time (i.e. at fixed intervals) or packet count (i.e. every k th packet). Figure 1(a) shows systematic sampling with a period of T seconds.



Figure 1. An illustration of the three categories of sampling techniques: (a) systematic sampling; (b) random sampling; and (c) stratified random sampling.

The use of systematic sampling always involves the risk of biasing the results. If the systematics (e.g. periodic repetition of an event) in the sampling process resemble the systematics in the observed stochastic process (e.g. occurrence of event of interest in the network), there is a high probability that the estimation will be biased. Typical examples of the systematics in the network are the periodic update of the routing table by a router, which has been shown in the literature to contribute to the periodic surge in packet delay, and the periodic exchange of information between routers due to SNMP protocol.

3.2. Random sampling

Random sampling, e.g. the Poisson sampling, employs a simple random distribution function to determine when a sample should be generated. Typically, the samples are generated according to a Poisson process or a uniform distribution. As shown in Figure 1(b), random sampling may produce a varying number of samples in a given time interval. With random sampling, an unbiased estimate of the QoS metric can be obtained [27, p. 21]. However, the entirely random nature of the sampling process may also cause the undesirable effect that the sampling intervals are not evenly distributed, and therefore the network may not be sampled for a rather long time.

3.3. Stratified random sampling

Stratified random sampling combines the fixed time interval used in systematic sampling with random sampling [19]. Figure 1(c) shows stratified random sampling with a stratum size of $4.5T$ and the elements of the sample are randomly generated in each stratum.

In stratified sampling, the elements of the parent population are first grouped into subsets (i.e. strata). The elements of the sample are then taken from each subset. Because the selections in different strata are made independently, the variances of the estimators for individual strata can be added together to obtain the variance of the estimator for the whole population. A smaller variance indicates a more accurate estimator. Since only the within-stratum variances enter into the variance of the estimator, the principle of stratification is to partition the population in such a way that the elements within a stratum are as similar as possible. The stronger the correlation between elements within a stratum, the more accurate the estimator will be. Therefore, even though strata may differ markedly from one another, a stratified sample with a desired number of elements from each stratum in the population will tend to be 'representative' of the population as a whole [28, p. 117]. Depending on how the sample size (i.e. number of elements in a sample) is distributed among strata, stratified sampling can be further classified into *proportional allocation* and *optimum allocation* [28]. In *proportional allocation*, the sample size in each stratum is allocated such that

it is proportional to the size of the parent population in that stratum. In *optimum allocation*, the sample size in each stratum is allocated such that it is proportional to the standard deviation of the variable of interest (e.g. packet delay) in that stratum. In this paper, the sampling time of the stratified sampling is divided into fixed length intervals (i.e. stratum) according to the correlation of the variable of interest (i.e. packet delay) to be measured, then sampling packets are selected according to a random process during each interval. The stronger the correlation between packet delays in an interval is, the more accurate the estimation of the mean packet delay will be.

4. PERFORMANCE ANALYSIS OF SIMPLE RANDOM SAMPLING AND STRATIFIED SAMPLING

In this section, we shall analyse the performance of different sampling techniques. As the variance of the sample mean has been widely used as a performance measure [18, 27, p. 15], the performance of these sampling techniques is analysed by comparing the variance of the sample mean of different sampling schemes under the constraint that they use the same sample size. The smaller the variance of the sample mean is, the better performance the sampling technique has. The sampling gain Δ is defined as the difference between the variances of the sample mean of two different sampling techniques [22]. Table I shows the notations used in our analysis.

Equations (1) and (2) present the two assumptions used in the analysis, and these are widely used assumptions in the area [21, 22]:

$$N_l - 1 \approx N_l, \quad l = 1, 2, \dots, L \quad (1)$$

$$\frac{n}{N} < 0.05 \quad (2)$$

Equation (1) requires that the parent population size in each stratum is a large number. Equation (2) requires that the sample size is small in comparison with the parent population size. Both assumptions can be readily satisfied.

Table I. Notations used in the analysis.

| Property | Parent population | Sample |
|--|-------------------|-------------|
| Total number of elements in a sample | N | n |
| Number of elements in the l th stratum | N_l | n_l |
| Number of strata | L | L |
| Mean value | μ | \bar{y} |
| Mean value in the l th stratum | μ_l | \bar{y}_l |
| Variance of the variable of interest | σ^2 | s^2 |
| Standard deviation of the variable of interest | σ | s |
| Variance of the variable of interest in the l th stratum | σ_l^2 | s_l^2 |
| Standard deviation of the variable of interest in the l th stratum | σ_l | s_l |
| Variable of interest (i.e. packet delay) | y | y |

For stratified sampling, it can be shown that the overall sample variance σ^2 is related to the sample variance in each stratum by

$$\begin{aligned}\sigma^2 &= \frac{1}{N-1} \sum_{i=1}^N (y_i - \mu)^2 = \frac{1}{N-1} \sum_{l=1}^L \sum_{i=1}^{N_l} (y_{li} - \mu)^2 \\ &= \frac{1}{N-1} \sum_{l=1}^L \sum_{i=1}^{N_l} [(y_{li} - \mu_l) + (\mu_l - \mu)]^2 \\ &= \frac{1}{N-1} \sum_{l=1}^L (N_l - 1) \sigma_l^2 + \frac{1}{N-1} \sum_{l=1}^L N_l (\mu_l - \mu)^2\end{aligned}\quad (3)$$

Applying the approximation in Equation (1), and multiplying both sides of Equation (3) by a common factor $1/n(1 - n/N)$, where $(1 - n/N)$ is the finite population correction (fpc) factor, it can be obtained that

$$\frac{1}{n} \left(1 - \frac{n}{N}\right) \sigma^2 = \frac{1}{n} \left(1 - \frac{n}{N}\right) \sum_{l=1}^L \frac{N_l}{N} \sigma_l^2 + \frac{1}{nN} \left(1 - \frac{n}{N}\right) \sum_{l=1}^L N_l (\mu_l - \mu)^2\quad (4)$$

Note that the approximation $N - 1 \approx N$ has been used in the above derivation. This approximation is a natural outcome of assumption 1.

The variance of the sample mean for simple random sampling is [28, p. 15]

$$\text{Var}_{\text{ran}}(\bar{y}) = \left(1 - \frac{n}{N}\right) \frac{\sigma^2}{n}\quad (5)$$

For stratified random sampling, the variance of the sample mean is given by [27, p. 91]

$$\text{Var}_{\text{st}}(\bar{y}) = \sum_{l=1}^L \left(\frac{N_l}{N}\right)^2 \left(\frac{N_l - n_l}{N_l}\right) \frac{\sigma_l^2}{n_l}\quad (6)$$

If stratified sampling with proportional allocation is used, then n_l is related to n , N and N_l by

$$n_l = n \frac{N_l}{N}\quad (7)$$

The variance of the sample mean for stratified sampling with proportional allocation becomes

$$\text{Var}_{\text{prop}}(\bar{y}) = \frac{1}{n} \left(1 - \frac{n}{N}\right) \sum_{l=1}^L \frac{N_l}{N} \sigma_l^2\quad (8)$$

Comparing Equations (4) and (5) with Equation (8), it can be shown that when the total sample size n is the same:

$$\text{Var}_{\text{ran}}(\bar{y}) = \text{Var}_{\text{prop}}(\bar{y}) + \frac{1}{n} \left(1 - \frac{n}{N}\right) \frac{1}{N} \sum_{l=1}^L N_l (\mu_l - \mu)^2\quad (9)$$

Hence, the sampling gain of the stratified sampling with proportional allocation compared with simple random sampling is

$$\Delta_{\text{prop}} = \text{Var}_{\text{ran}}(\bar{y}) - \text{Var}_{\text{prop}}(\bar{y}) \quad (10)$$

$$= \frac{1}{nN} \left(1 - \frac{n}{N}\right) \sum_{l=1}^L N_l (\mu_l - \mu)^2 \geq 0 \quad (11)$$

The sampling gain is always positive, which implies that the performance improvement can be achieved in moving from simple random sampling to stratified sampling with proportional allocation.

If stratified sampling with optimum allocation [28] is used, then n_l is given by

$$n_l = \frac{n N_l \sigma_l}{\sum_{k=1}^L N_k \sigma_k} \quad (12)$$

The variance of the sample mean for stratified sampling with optimum allocation can be obtained from Equations (6) and (12):

$$\text{Var}_{\text{opt}}(\bar{y}) = \frac{1}{n} \left(\sum_{l=1}^L \frac{N_l}{N} \sigma_l \right)^2 - \frac{1}{N} \sum_{l=1}^L \frac{N_l}{N} \sigma_l^2 \quad (13)$$

From Equations (8) and (13), we can derive the difference between $\text{Var}_{\text{prop}}(\bar{y})$ and $\text{Var}_{\text{opt}}(\bar{y})$:

$$\text{Var}_{\text{prop}}(\bar{y}) - \text{Var}_{\text{opt}}(\bar{y}) = \frac{1}{nN} \sum_{l=1}^L N_l (\sigma_l - \bar{\sigma}_l)^2 \geq 0 \quad (14)$$

where $\bar{\sigma}_l$ is

$$\bar{\sigma}_l = \sum_{l=1}^L \frac{N_l}{N} \sigma_l \quad (15)$$

This sampling gain is positive, which means the stratified sampling with optimum allocation has better performance than the stratified sampling with proportional allocation.

Similarly, fixing the sample size n and ignoring the fpc factor (which is close to 1 by using assumption in Equation (2)), the sampling gain of stratified sampling with optimum allocation *versus* simple random sampling is

$$\Delta_{\text{opt}} = \text{Var}_{\text{ran}}(\bar{y}) - \text{Var}_{\text{opt}}(\bar{y}) \quad (16)$$

$$= \frac{1}{nN} \left[\sum_{l=1}^L N_l (\mu_l - \mu)^2 + \sum_{l=1}^L N_l (\sigma_l - \bar{\sigma}_l)^2 \right] \quad (17)$$

$$\geq 0 \quad (18)$$

This means that stratified sampling with optimum allocation also has better performance than simple random sampling.

Based on Equations (11), (14) and (18), we can conclude that the stratified sampling with proportional allocation performs better than the simple random sampling, and stratified sampling with optimum allocation has the best performance among the three sampling techniques.

Note that earlier analysis is performed on count-based sampling techniques, i.e. sampling is triggered by packet count. Based on the observation in [17] that the difference between count-based sampling and timer-based sampling (sampling is triggered by time) is very small, these conclusions may also extend to the timer-based sampling techniques.

5. ADAPTIVE STRATIFIED SAMPLING ALGORITHM

In the last section, we have shown that stratified sampling with optimum allocation has the best performance. However, Equation (12) implies that the stratified sampling with optimum allocation requires the knowledge of the standard deviation of the parent population in the l th stratum, i.e. σ_l , in order to allocate the sample size in the l th stratum. This requirement is unrealistic for real-time applications. In this section, we develop an adaptive stratified sampling algorithm, which uses the LMS algorithm to predict the value of σ_l for sample size allocation. The proposed algorithm is then applied to packet delay sampling.

5.1. Least-mean-square algorithm

The LMS algorithm is one of the most widely used adaptive linear algorithm. The computational procedure of the LMS algorithm for one-step prediction is listed in the following [29, p. 655] for completeness:

- Compute the required output:

$$\hat{x}_k = \sum_{i=0}^{m-1} w_k(i)x_{k-1-i} = \mathbf{W}_k^T \mathbf{X}(k) \quad (19)$$

where m is the order of the predictor, $\mathbf{X}(k)$ is the input vector and \mathbf{W}_k is the prediction coefficient vector:

$$\mathbf{X}(k) = [x_{k-1}, x_{k-2}, \dots, x_{k-m}]^T \quad (20)$$

$$\mathbf{W}_k = [w_k(0), w_k(1), \dots, w_k(m-1)]^T \quad (21)$$

Initially, each weight $w_k(i)$ is set to an arbitrary fixed value.

- Compute the prediction error:

$$e_k = x_k - \hat{x}_k \quad (22)$$

- Update the coefficient vector:

$$\mathbf{W}_{k+1} = \mathbf{W}_k + 2ve_k \mathbf{X}(k) \quad (23)$$

where v is the step size.

5.2. Prediction of the sample size within a stratum

It has been shown in Section 3.3 that for stratified sampling with optimum allocation, the sample size within a stratum is

$$n_l = \frac{nN_l\sigma_l}{\sum_{k=1}^L N_k\sigma_k} \quad (24)$$

To simplify the estimation of n_l , an assumption is made that the parent population size N_l is approximately the same in each stratum, i.e.

$$\frac{N_l}{N_k} \approx 1, \quad l \neq k \quad (25)$$

This assumption is valid when the parent population size N_l is very large and the stratum size is a constant in time. This assumption has been validated using a real traffic trace. Details of the traffic trace is explained in Section 6. Figure 2 shows the ratio N_k/N_1 of the real traffic trace with a stratum size of 50, 100, 130 and 200 s, respectively, where N_k , $k = 1, 2, \dots, L$ is the total number of packets within the k th stratum of the real traffic trace and N_1 is the total number of packets within the 1st stratum of the real traffic trace. We can see that the ratio N_k/N_1 is approximately bounded in the interval [0.8, 1.2].

Using the assumption in Equation (25), Equation (24) can be simplified as

$$n_l \approx \frac{n\sigma_l}{\sum_{k=1}^L \sigma_k} = \frac{n\sigma_l}{L\bar{\sigma}_s} = \phi\sigma_l \quad (26)$$

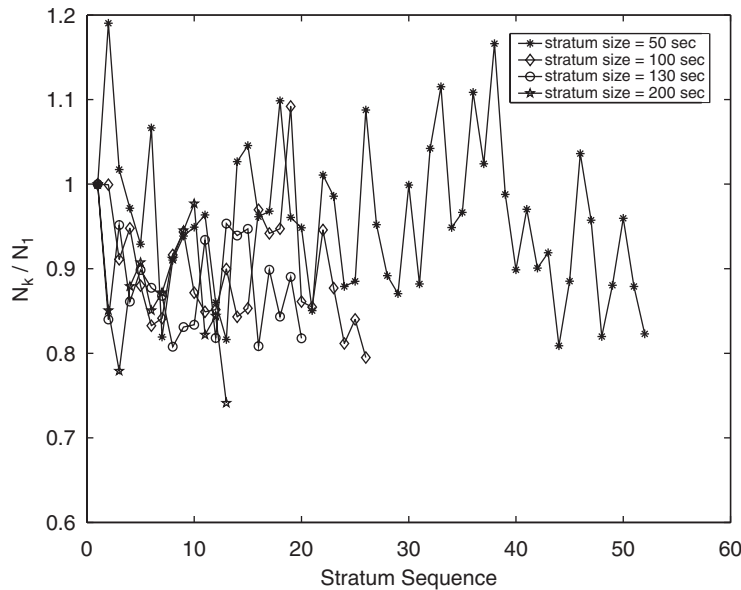


Figure 2. Ratio of packet number between different strata.

where

$$\varphi = \frac{n}{\sum_{k=1}^L \sigma_k} = \frac{n}{L\bar{\sigma}_s} \tag{27}$$

In real applications, φ can be simply treated as a proportionality constant which controls the sampling rate. φ can be chosen empirically and a larger φ will produce a higher sampling rate.

Since the standard deviation of packet delay σ_l is the true value of the parent delay trace, which cannot be obtained in real applications, it is approximated by the corresponding standard deviation of sampling packet delay s_l . Then, the LMS algorithm is employed to predict s_l from its past values. Hence, the estimator \hat{n}_l of sample size for the l th stratum is computed by

$$\hat{n}_l = \varphi \hat{s}_l \tag{28}$$

where \hat{s}_l is predicted from the past values using the LMS algorithm:

$$\hat{s}_l = \sum_{i=0}^{m-1} w_l(i) s_{l-1-i} \tag{29}$$

$$e_l = s_l - \hat{s}_l \tag{30}$$

$$w_{l+1}(i) = w_l(i) + 2ve_l s_{l-1-i}, \quad i = 0, 1, \dots, m - 1 \tag{31}$$

The predictor order can be obtained using the AICC criterion for order selection [30, p. 171].

5.3. An analysis on the estimation error

The estimation error in \hat{n}_l may increase the variance of the sample mean, which in turn leads to a decrease in the estimation accuracy of the adaptive sampling method. From Equation (6), the variance of the sample mean when using the predicted stratum sample size \hat{n}_l (instead of n_l calculated from Equation (12)) is

$$\text{Var}_{\text{act}}(\bar{y}) = \sum_{l=1}^L \left(\frac{N_l}{N} \right)^2 \frac{\sigma_l^2}{\hat{n}_l} - \sum_{l=1}^L \left(\frac{N_l}{N} \right)^2 \frac{\sigma_l^2}{N_l} \tag{32}$$

From Equations (13) and (32), and ignoring the fpc factor, we can derive the relative error between $\text{Var}_{\text{act}}(\bar{y})$ and $\text{Var}_{\text{opt}}(\bar{y})$ ($\text{Var}_{\text{opt}}(\bar{y})$ is calculated using the ‘true value’ of n_l):

$$\frac{\text{Var}_{\text{act}}(\bar{y}) - \text{Var}_{\text{opt}}(\bar{y})}{\text{Var}_{\text{opt}}(\bar{y})} = \frac{1}{n} \sum_{l=1}^L \frac{(\hat{n}_l - n_l)^2}{\hat{n}_l} = \frac{1}{n} \sum_{l=1}^L n_l \frac{(\phi_l - 1)^2}{\phi_l} \tag{33}$$

where $\phi_l = \hat{n}_l/n_l$. Equation (33) relates the performance of the proposed adaptive stratified sampling scheme (compared with the stratified sampling with optimum allocation) to the prediction error in \hat{n}_l .

From Equations (26) and (28), we can obtain that

$$\phi_l = \frac{\hat{n}_l}{n_l} = \frac{\varphi \hat{s}_l}{\varphi \sigma_l} = \frac{\hat{s}_l}{\sigma_l} \tag{34}$$

Therefore, the prediction error in \hat{n}_l comes from two sources:

- the estimation error in using s_l as an approximation of σ_l ;
- the prediction error in using the LMS algorithm to predict s_l .

For the first item, it can be shown that for a sample with ω elements the statistic $\eta = (\omega - 1)s^2/\sigma^2$ has a χ^2 distribution with $(\omega - 1)$ degrees of freedom [31, p. 216], where s^2 is the sample variance. Let λ_1 and λ_2 denote the two critical values of the χ^2 distribution, which are determined by the confidence level $1 - \alpha$, then the confidence interval of the variance of the parent population σ^2 is [31, p. 216]

$$\frac{(\omega - 1)s^2}{\lambda_2} \leq \sigma^2 \leq \frac{(\omega - 1)s^2}{\lambda_1} \tag{35}$$

Therefore, for the l th stratum, we have

$$\sqrt{\frac{\lambda_1}{\hat{n}_l - 1}} \leq \frac{s_l}{\sigma_l} \leq \sqrt{\frac{\lambda_2}{\hat{n}_l - 1}} \tag{36}$$

For the second item, from Equation (30), we have

$$s_l = \hat{s}_l + e_l \tag{37}$$

From Equations (34), (36) and (37), we can obtain the confidence bounds of the ratio \hat{n}_l/n_l with a given confidence level $1 - \alpha$:

$$\sqrt{\frac{\lambda_1}{\hat{n}_l - 1}} - \frac{e_l}{\hat{s}_l} \leq \frac{\hat{n}_l}{n_l} \leq \sqrt{\frac{\lambda_2}{\hat{n}_l - 1}} - \frac{e_l}{\hat{s}_l} \tag{38}$$

or

$$\sqrt{\frac{\lambda_1}{\hat{n}_l - 1}} - \frac{e_l}{\hat{s}_l} \leq \phi_l \leq \sqrt{\frac{\lambda_2}{\hat{n}_l - 1}} - \frac{e_l}{\hat{s}_l} \tag{39}$$

where λ_1 and λ_2 is determined by $1 - \alpha$.

Now we have obtained the lower bound and the upper bound for the ratio $\phi_l = \hat{n}_l/n_l$, which is expressed in the form of \hat{n}_l and the prediction error e_l . We are ready to examine the relative error in $\text{Var}_{\text{act}}(\bar{y})$, as given in Equation (33). Let Λ denote the function $\Lambda(\phi) = (\phi - 1)^2/\phi = \phi + 1/\phi - 2$, Figure 3 shows the variation of $\Lambda(\phi)$ with ϕ . Given this function, if we know the upper bound and the lower bound of ϕ , which is determined by Equation (39), the maximum value of $\Lambda(\phi)$ in this limited range of ϕ can also be determined. Figure 3 shows that this maximum value of $\Lambda(\phi)$ will occur at either the lower bound of ϕ or the upper bound of ϕ . Let ϕ_{max} denote the value at which Λ reaches its maximum value. Then Equation (33) can be further simplified as

$$\frac{\text{Var}_{\text{act}}(\bar{y}) - \text{Var}_{\text{opt}}(\bar{y})}{\text{Var}_{\text{opt}}(\bar{y})} = \frac{1}{n} \sum_{l=1}^L n_l \frac{(\phi_l - 1)^2}{\phi_l} \tag{40}$$

$$\leq \frac{1}{n} \sum_{l=1}^L n_l \frac{(\phi_{\text{max}} - 1)^2}{\phi_{\text{max}}} \tag{41}$$

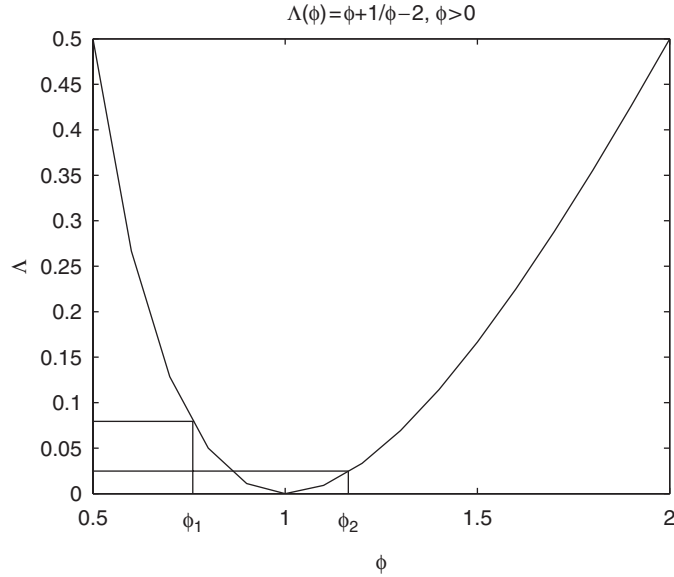


Figure 3. Variation of $\Lambda(\phi)$ as a function of ϕ .

$$= \frac{(\phi_{\max} - 1)^2}{\phi_{\max}} \frac{1}{n} \sum_{l=1}^L n_l \quad (42)$$

$$= \frac{(\phi_{\max} - 1)^2}{\phi_{\max}} \quad (43)$$

When $\phi_{\max} = 0.9$, the relative error between $\text{Var}_{\text{act}}(\bar{y})$ and $\text{Var}_{\text{opt}}(\bar{y})$ is smaller than 0.0111; when $\phi_{\max} = 1.2$, the relative error between $\text{Var}_{\text{act}}(\bar{y})$ and $\text{Var}_{\text{opt}}(\bar{y})$ is smaller than 0.0333. We can see that the impact of the estimation error on the measurement accuracy is marginal when \hat{n}_l is reasonably accurate.

In Section 7, we have implemented simulations for the adaptive stratified sampling. The ratio $\phi_l = \hat{n}_l/n_l$ in the simulations was shown to be bounded in $[0.77, 1.34]$. $\Lambda(\phi)$ reaches its maximum value at $\phi = 1.34$. Hence $\phi_{\max} = 1.34$, and the relative error between $\text{Var}_{\text{act}}(\bar{y})$ and $\text{Var}_{\text{opt}}(\bar{y})$ is smaller than 0.086. This means that the proposed adaptive stratified sampling scheme achieves almost the same performance as stratified sampling with optimum allocation, which represents the best sampling performance that can be achieved.

6. GENERATION OF THE PARENT PACKET DELAY TRACE

In order to compare the performance of different sampling techniques, experiments are necessary. In this paper, all experiments are performed using a one-way delay trace as the parent packet

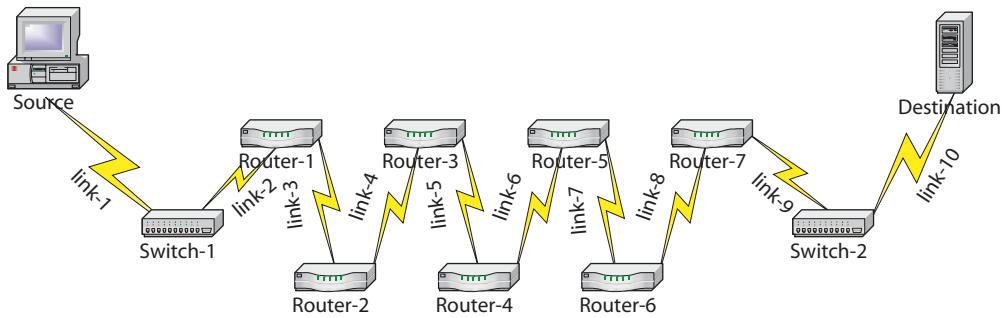


Figure 4. Network topology used in Opnet Modeler.

Table II. Selection of network nodes and background traffic utilizations of links.

| Nodes | Description | Background traffic utilization |
|---------------------|-----------------------------------|--------------------------------|
| Switch-1, 2 | 3 Com's SuperStack II Switch 3800 | N/A |
| Router-1, 2, ..., 7 | CISCO 12008 | N/A |
| Link-1, 10 | 100 Mbps Link | 0% |
| Link-2, 3, 8, 9 | 100 Mbps Link | 50% |
| Link-4, 7 | 100 Mbps Link | 70% |
| Link-5, 6 | 100 Mbps Link | 55% |

Table III. A summary of the statistics for packet delay, packet size and inter-arrival time of the parent packet delay trace.

| Property | Minimum | Maximum | Mean | Variance |
|-------------------------|---------|----------|--------|----------|
| Packet delay (ms) | 41.092 | 141.305 | 86.024 | 8529 |
| Packet size (bytes) | 64 | 1518 | 440.5 | 302 080 |
| Inter-arrival time (ms) | 0.006 | 203.3280 | 4.5181 | 74.4127 |

delay trace. This delay trace is generated by importing a real traffic trace into Opnet Modeler. This real traffic trace ('20010613-060000-e1.gz') was collected by the WAND research group at the University of Waikato Computer Science Department. It was captured between 6.00 a.m. and 8.54 a.m. on 13 June 2001 on a 100 Mbps Ethernet link. IP headers in the traffic trace are GPS synchronized and have a time accuracy of 1 μ s. More information on the traffic trace and the measurement infrastructure can be found on the research group's website [32].

The network topology used in the Opnet Modeler is shown in Figure 4. The selection of network nodes (e.g. switch, router, link) and background traffic utilizations of the links are shown in Table II. The first 2600-s part of the entire trace is then imported into the Opnet Modeler. After the simulation, we obtain a one-way packet delay traffic trace of a duration of 2600 s with 577 718 packets. For the purpose of our study, we treat the 2600-s delay traffic trace as the parent packet delay trace. Table III shows a summary of the statistics for the packet delay, packet size and inter-arrival time of the parent packet delay trace.

The minimum sample size required to obtain an estimate satisfying a given accuracy criterion is given in [33, p. 27], which is

$$n \geq \frac{z_{\alpha/2}^2 \sigma^2}{r^2 \mu^2} \quad (44)$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ quantile of the Gaussian distribution and is determined by the confidence level $1 - \alpha$; r is the estimate accuracy (i.e. the bounds of the relative error between the actual value and its estimate). In this paper, $\mu = 86.024$ ms and $\sigma^2 = 8529$. If an accuracy of $r = \pm 5\%$ and a confidence level of $100(1 - \alpha)\% = 95\%$ are required in estimating the mean packet delay, then $z_{\alpha/2} = 1.96$. Hence, the minimum sample size is 1739. Since the parent packet delay trace has a time duration of 2600 s and the sampling frequency is chosen to be 1 packet/s, the total sample size is approximately 2600, which satisfies the accuracy requirement. Moreover, the ratio between the sample size (2600) and the parent population size (577 718) is 0.45%, which satisfies with the assumption in Equation (2). If an accuracy of $r = \pm 1\%$ and a confidence level of $100(1 - \alpha)\% = 95\%$ are used, the minimum sample size would be 43 464, which means that a much higher sampling rate would be required. This paragraph gives a guideline on choosing the sampling frequency. In real applications, the values of μ and σ^2 can be estimated empirically from past measurements.

7. SIMULATION RESULTS

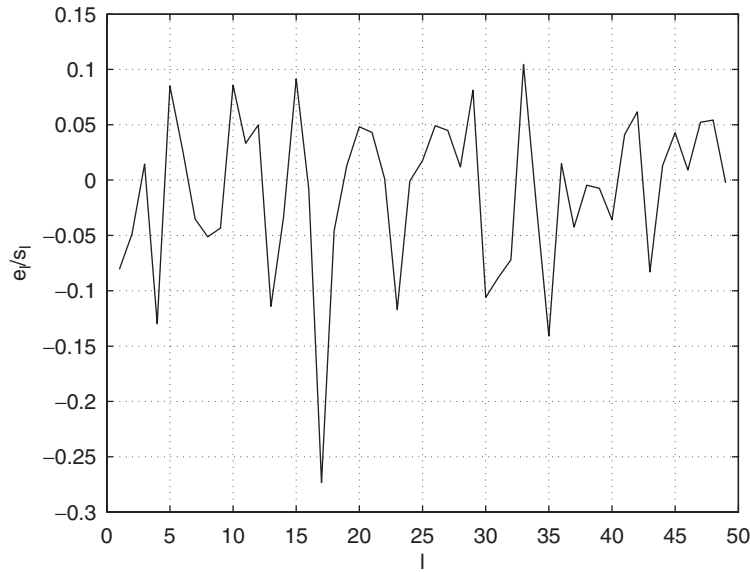
In this section, the performance of different sampling schemes, i.e. systematic sampling, the Poisson sampling, stratified sampling with optimum allocation and the proposed adaptive stratified sampling, is compared using Monte-Carlo simulations [34]. The total sample size for each sampling technique is chosen to be the same. The parent packet delay trace used for simulation is the one-way delay trace presented in Section 6.

For stratified sampling with optimum allocation, the parameters required to calculate the sample size n_l are all true values obtained from the parent packet delay trace. It is used as a benchmark, which represents the best sampling performance that can be achieved. The sample packet delay traces are selected directly from the parent delay trace. The sampling goal is to estimate the mean packet delay μ and the variance of packet delay σ^2 of the parent packet delay trace.

Several C programs were developed for obtaining the sample packet delay traces and calculating the estimated mean packet delay $\hat{\mu}$ and the estimated variance of packet delay $\hat{\sigma}^2 = s^2$ from the sample packet delay traces, where $\hat{\mu}$ is the mean packet delay of the sample delay trace and s^2 is the variance of packet delay of the sample delay trace. For simulation, each kind of sampling technique (e.g. systematic sampling, systematic sampling) is repeated a number of times, and the random seed in the C programs is updated in each repetition. Let M denote the number of repetitions (termed sampling rounds in this paper). After M sampling rounds, we obtain M different sample delay traces. The estimated mean delay $\hat{\mu}$ and estimated variance of delay s^2 are calculated for each sample delay trace in the M sampling rounds. Then, we can obtain M estimated mean delay, i.e. $\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_M$ and M estimated variances of packet delay, i.e. $s_1^2, s_2^2, \dots, s_M^2$. The absolute error of the estimated mean delay, i.e. $|\hat{\mu}_i - \mu|$, and the absolute error of the estimated variance, i.e. $|s_i^2 - \sigma^2|$, are also calculated for each M sampling round, where the true values μ and σ^2 are obtained from the parent packet delay trace in Section 6 and shown in Table III.

Table IV. Simulation results of sampling tests using different sampling methods (true values are: $\mu = 86.824$ ms, $\sigma^2 = 8529$).

| Sampling method | M | AMean (ms) | AVar | MSE |
|------------------------------------|-----|------------|------|-----|
| Systematic | 222 | 74.803 | 5996 | 145 |
| Poisson | 222 | 64.998 | 4710 | 478 |
| Stratified with optimum allocation | 222 | 88.023 | 8959 | 5 |
| Adaptive stratified | 222 | 84.895 | 8081 | 10 |


 Figure 5. Relative error in predicting s_l in Equation (30).

To compare the performance of different sampling methods, several metrics are used, which are:

- Average value of the sample mean (AMean): the average value of the sample means of the M sample delay traces. $AMean = (1/M) \sum_{i=1}^M \hat{\mu}_i$, where M is the number of sampling rounds, $\hat{\mu}_i$ is the mean value of the i th sample delay trace in the M sample delay traces. The smaller the difference between $AMean$ and μ is, the better the performance is.
- Average sample variance (AVar): the average value of the sample variances of the M sample delay traces. $AVar = (1/M) \sum_{i=1}^M s_i^2$, where s_i^2 is the variance of the i th sample delay trace. The smaller the difference between $AVar$ and σ^2 is, the better the performance is.
- Mean square error (MSE) of the sample mean $\hat{\mu}_i$: $MSE = (1/M) \sum_{i=1}^M (\hat{\mu}_i - \mu)^2$. The smaller the MSE is, the higher the accuracy is.
- Absolute error of estimated mean (AEMean): $|\hat{\mu}_i - \mu|$, the smaller $|\hat{\mu}_i - \mu|$ is, the lower the variance of the sample mean $Var(\hat{\mu})$ is.

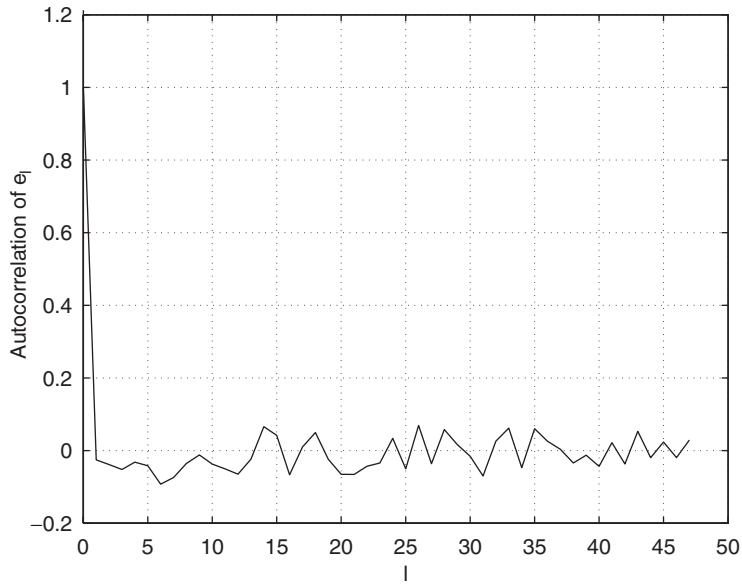


Figure 6. Autocorrelation of prediction error e_l in Equation (30).

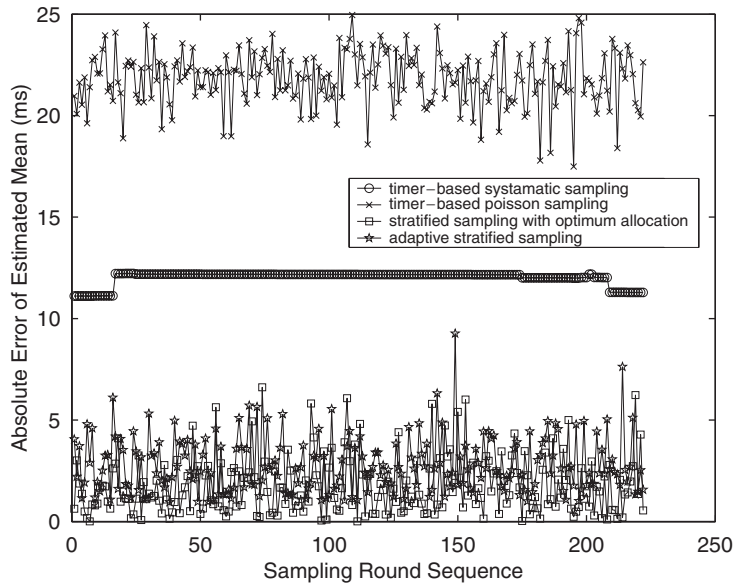


Figure 7. Comparison of absolute error of estimated mean for different sampling methods. Stratum size: 50 s, sampling rounds: 222.

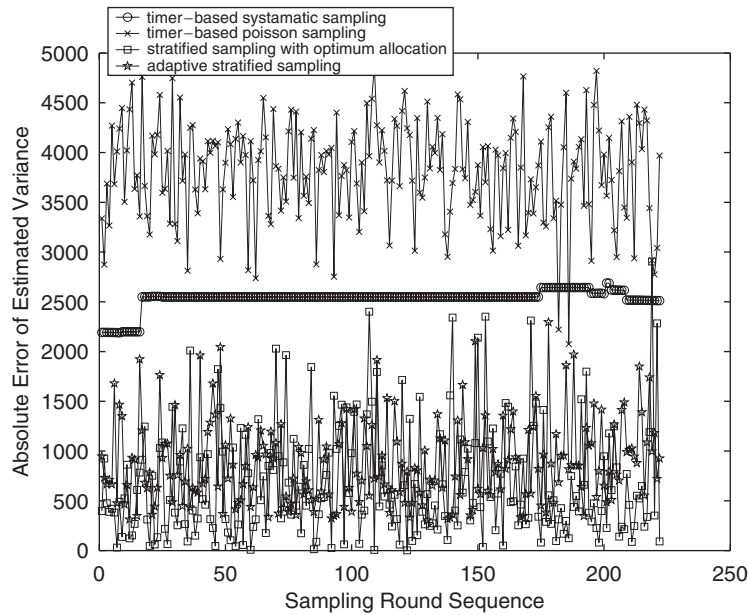


Figure 8. Comparison of absolute error of estimated variance for different sampling methods. Stratum size: 50 s, sampling rounds: 222.

- Absolute error of estimated variance (AEVar): $|s_i^2 - \sigma^2|$, the smaller $|s_i^2 - \sigma^2|$ is, the better can s^2 estimate the true variance σ^2 .

The first two metrics are used to give an intuitive gauge of the sampling accuracy by comparing them directly with the corresponding true values of the variables of interest as shown in Table IV. The rest three metrics are popularly employed metrics for comparing the performance of different sampling methods [19, 21, 22, 24, 25].

Then simulations for these four sampling methods are performed. Each simulation is repeated for 222 times (i.e. $M = 222$), where $M = 222$ is a randomly chosen large number. For systematic sampling, the sampling interval is specified as 1 s. For Poisson sampling, the mean sampling interval is 1 s. For the proposed adaptive sampling, the prediction parameters and stratum size are chosen empirically. The predictor order is 4; the initial weights are: $w_l(0) = 0.257$, $w_l(1) = 0.210$, $w_l(2) = 0.209$ and $w_l(3) = 0.260$; the step size is: $v = 0.02$; the stratum size is: 50 s. These values of the predictor order, initial weights and stratum size are obtained empirically using a different traffic trace ('20010612-060000-e1.gz'), which was captured between 6.00 a.m. and 8.54 a.m. on 12 June 2001 on a 100 Mbps Ethernet link [32], from that used in Section 6. The stratum size is determined empirically by examining the packet delay correlation in the trace such that the packet delay correlation in the same stratum is large enough.

Figure 5 shows the relative error (i.e. e_l/s_l) in predicting the standard deviation of sampling packet delay in each stratum and Figure 6 shows the autocorrelation of e_l . We can see that the relative error e_l/s_l is marginal and the error e_l is approximately independent, which indicates a good performance of the prediction algorithm. The total sample size n is specified as 2600

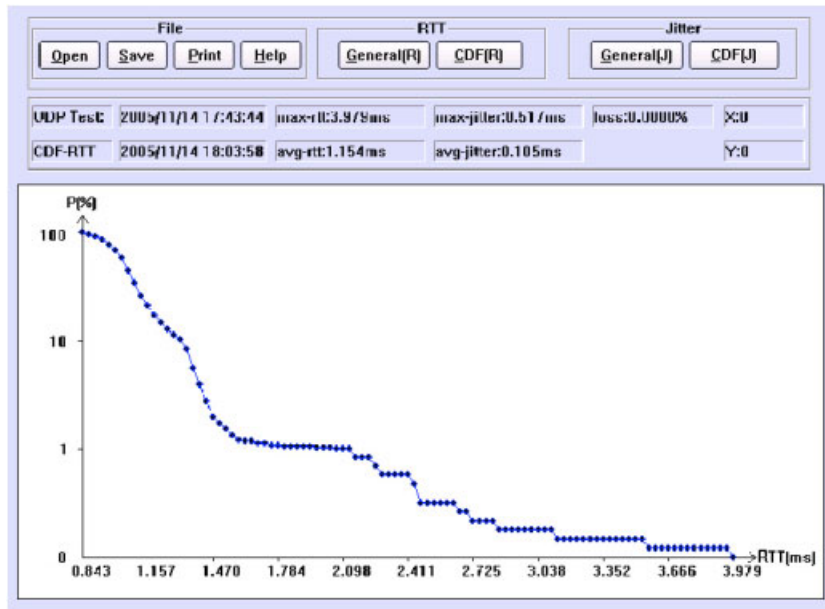


Figure 9. RTT measurements using UDP protocol between a server located at the University of Sydney and a client located at NICTA at ATP, Sydney. Both the client and the server are connected to a high-speed LAN, then to a WAN. The average RTT measured is 1.154 ms. The horizontal axis shows the packet delay and the vertical axis shows the probability that the measured packet delay exceeds a specified packet delay value.

in order to make sure it has the same sample size as the systematic sampling and the Poisson sampling ($n = \text{sampling duration} / \text{sampling rate} = 2600 / 1 = 2600$). For stratified sampling with optimum allocation, the stratum size is also specified as 50 s and the total sample size is 2600.

Table IV shows the simulation results. It can be seen that the stratified sampling with optimum allocation achieves the best performance. The proposed adaptive sampling scheme produces approximately the same performance as the stratified sampling with optimum allocation; it performs much better than systematic sampling and Poisson sampling. Figure 7 shows the absolute error of the estimated mean in different simulation rounds (i.e. 222 rounds) and Figure 8 shows the absolute error of the estimated variance for the 222 sampling rounds. Both figures show that the proposed adaptive stratified sampling gives better estimation accuracy than the systematic sampling and the Poisson sampling; and the performance of the proposed adaptive scheme is close to the performance of stratified sampling with optimum allocation.

8. SOFTWARE DESIGN

In this section, we shall introduce a QoS monitoring software designed by us based on the earlier analysis.

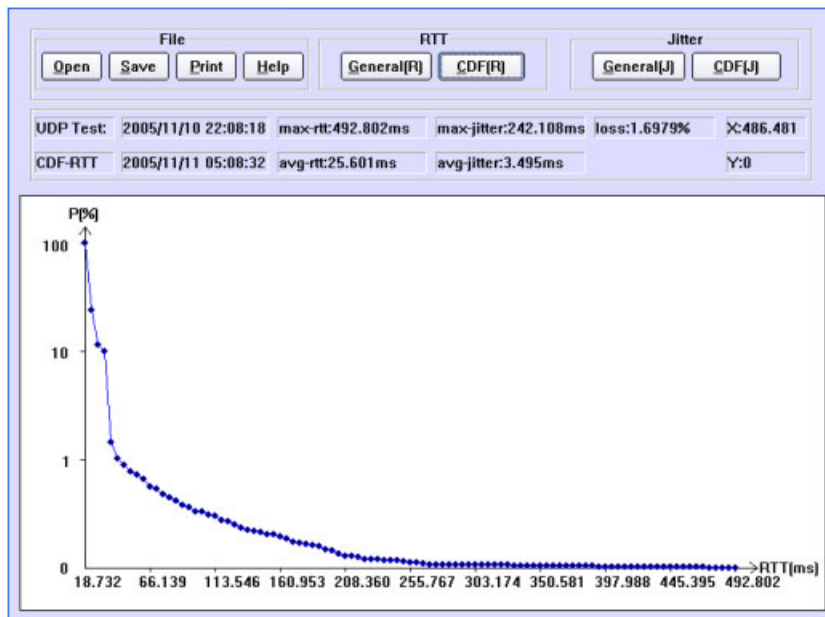


Figure 10. RTT measurements using UDP protocol between a server located at the University of Sydney and a client located at Carlton, Sydney. The server is connected to a WAN through a high-speed LAN. The client is connected to a WAN through a wireless LAN, then ADSL. The average RTT measured is 25.601 ms. The horizontal axis shows the packet delay and the vertical axis shows the probability that the measured packet delay exceeds a specified packet delay value.

The software is developed using C++ language and is designed to operate in widely used versions of the Windows operating system (i.e. Windows 2000 and XP). The software uses a client-server architecture. The client initiates the measurement request, and the server responds to the client's request, performs some simple computation and returns the required information back to the client. One server can respond to multiple clients. The measurements can be taken using TCP, UDP or ICMP protocol. The software supports systematic sampling, the Poisson sampling and the proposed adaptive sampling strategies, and also it supports RTT (Round-trip) delay, one-way jitter and one-way loss measurements. Moreover, it also provides a simple idea about the network availability by diagnosing and recording the time when the network is available and when the network is unreachable. Users can specify sampling parameters, e.g. 'IP precedence', packet size, packet size distribution (constant or random size), sampling methods, sampling frequency. The flexibility in choosing sampling parameters allows the program to be used in different network environments.

Figures 9–11 show RTT measurements performed in a number of different environments. In particular, the RTT measurement results shown in Figure 11 between a computer located at the University of Sydney and a computer located at NICTA in Canberra are impressively good. Note that in simulation environments, the true values for the performance metrics of interest (e.g. packet delay, jitter) are known, whereas in real environments, the true values cannot be obtained.

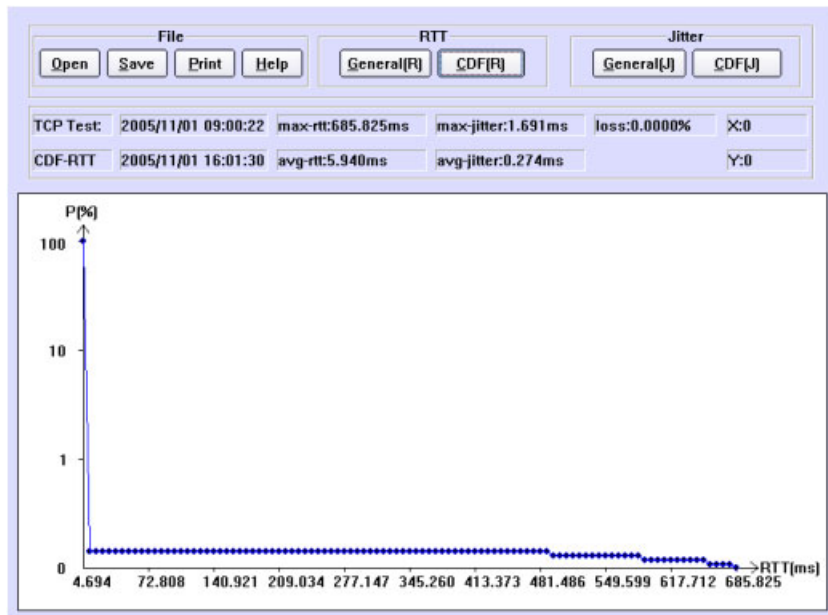


Figure 11. RTT measurements using TCP protocol between a server located at the University of Sydney and a client located at NICTA in Canberra. Both the client and the server are connected to a high-speed LAN, then to a WAN. The average RTT measured is 5.940 ms. The horizontal axis shows the packet delay and the vertical axis shows the probability that the measured packet delay exceeds a specified packet delay value.

Therefore, it is impossible to make the same performance comparisons for the real data obtained here as that for the simulation data.

9. CONCLUSION

In this paper, we investigated several major sampling techniques, i.e. systematic sampling, random sampling and stratified sampling (with proportional allocation and optimum allocation) that can be used for end-to-end QoS measurement. Theoretical analysis was conducted, which showed that the stratified sampling with optimum allocation has the best performance. However, the implementation of stratified sampling with optimum allocation requires knowledge of the standard deviation of the variable of interest in each stratum, which is not available for real-time applications. Therefore, we proposed a novel adaptive stratified sampling scheme to solve the problem. The proposed sampling scheme is based on stratified sampling with optimum allocation. It employs a LMS algorithm to predict the standard deviation of the variable of interest in a stratum, which is required to compute the sample size for that stratum. Both analytical study and simulation were conducted on the performance of the proposed scheme, which showed that the proposed adaptive sampling scheme outperforms systematic sampling and the Poisson sampling and achieves a performance close to that of stratified sampling with optimum allocation, which is used as a benchmark representing the best performance.

ACKNOWLEDGEMENTS

This research was supported by Optus Ltd. through contract BLO-2004-7260 and was additionally supported by the Australian Research Council Discovery Grant, grant number: DP0559248.

REFERENCES

1. Drobisz J, Christensen KJ. Adaptive sampling methods to determine network traffic statistics including the Hurst parameter. *Proceedings of IEEE Annual Conference on Local Computer Networks*, Boston, MA, U.S.A., 1998; 238–247.
2. Online resource. Advanced Network & Services, Inc. <http://www.Advanced.org> (Date of Access 02 May 2007).
3. Online resource, CAIDA Measurement Tool Taxonomy. <http://www.caida.org> (Date of Access 02 May 2007).
4. Online resource, Internet2 Working Groups. <http://www.internet2.edu/wg/> (Date of Access 02 May 2007).
5. Online resource, National Laboratory for Applied Network Research (NLNR). <http://www.moat.nlanr.net> (Date of Access 02 May 2007).
6. Online resource, Networking Research at the Pittsburgh Supercomputing Center. <http://www.psc.edu/networking/welcome.html> (Date of Access 02 May 2007).
7. Shin S-C, Shin S-Y, Han S-Y. Network performance monitoring system for SLA: implementation and practices. *Sixth International Conference on Advanced Communication Technology*, Phoenix Park, Republic of Korea, vol. 2. 2004; 661–664.
8. Carlos M-J, Shrivastava S, Crowcroft J, Gevros P. On the monitoring of contractual service level agreement. *Proceedings of the First International Workshop on Electronic Contracting*, San Diego, CA, U.S.A., 2004; 1–8.
9. Bolot J-C. Characterizing end-to-end packet delay and loss in the Internet. *Journal of High-Speed Networks* 1993; 2(3):305–323.
10. Borella MS, Swider D, Uludag S, Brewster GB. Internet packet loss: measurement and implications for end-to-end QoS. *Proceedings of the International Conference on Parallel Processing*, Minneapolis, MN, U.S.A., 1998; 3–12.
11. Claffy K. Internet measurement and data analysis: topology, workload, performance and routing statistics. *NAE Workshop*, 1999.
12. Kato JY, Shimizu A, Goto S. Active measurement and analysis of delay time in the Internet. *Proceedings of the International Conference on Parallel Processing*, Aizu-Wakamutsu, Japan, 1999; 254–259.
13. Chio B-Y, Moon S, Zhang Z-L, Papagiannaki K, Diot C. Analysis of point-to-point packet delay in an operational network. *IEEE INFOCOM*, Hongkong, China, 2004; 3:1797–1807.
14. Asgari A, Trimintzios P, Irons M, Pavlou G, Egan R, Berghe SV den. A scalable real-time monitoring system for supporting traffic engineering. *Proceedings of the IEEE Workshop on IP Operations and Management*, Dallas, U.S.A., 2002; 202–297.
15. Wang PY, Yemini Y, Florissi D, Florissi P, Zinky J. Experimental QoS performances of multimedia applications. *Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, Tel Aviv, Israel, vol. 2. 2000; 970–979.
16. Chen TM, Liu SS, Procanik MJ, Wang DC, Casey DD. INQUIRE: a software approach to monitoring QoS in ATM networks. *IEEE Network* 1998; 12:32–37.
17. Claffy KC, Polyzos GC, Braun H-W. Application of sampling methodologies to network traffic characterization. *ACM SIGCOMM Computer Communication Review* 1993; 23(4):194–203.
18. Cozzani I, Giordano S. Traffic sampling methods for end-to-end QoS evaluation in large heterogeneous networks. *Computer Networks and ISDN Systems* 1998; 30(16–18):1697–1706.
19. Hernandez E, Chidester M, George A. Adaptive sampling for network management. *Journal of Network and Systems Management* 2001; 9(4):409–434.
20. Choi B-Y, Park J, Zhang Z-L. Adaptive random sampling for load change detection. *Joint International Conference on Measurement and Modeling of Computer Systems*, California, U.S.A., 2002; 272–273.
21. Zseby T. Deployment of sampling methods for SLA validation with non-intrusive measurements. *Proceedings of Passive and Active Measurement Workshop*, Fort Collins, CO, U.S.A., 2002.
22. Zseby T. Stratification strategies for sampling-based non-intrusive measurements of one-way delay. *The Passive and Active Measurement Workshop*, San Diego, CA, U.S.A., 2003; 171–179.
23. Zseby T, Zander S. Sampling schemes for validating service level agreements. *CAIA Technical Report 040706A*, 2004.

24. Wang J, Yang J, Xie G, Zhou M, Li Z. Measurement-based modeling with adaptive sampling. *Proceedings of the 12th Asian Test Symposium*, Xi'an, China, 2003; 340–345.
25. Ma W, Yan J, Huang C. Adaptive sampling methods for network performance metrics measurement and evaluation in MPLS-based IP networks. *IEEE CCECE* 2003; 2:1005–1008.
26. Ma W, Huang C, Yan J. Adaptive sampling for network performance measurement under voice traffic. *IEEE International Conference on Communications* 2004; 2:1129–1134.
27. Cochran WG. *Sampling Techniques* (2nd edn). Wiley: New York, 1964.
28. Thompson SK. *Sampling* (2nd edn). Wiley: New York, 2002.
29. Ifeachor EC, Jervis BW. *Digital Signal Processing—A Practical Approach* (2nd edn). Pearson Education Limited: London, 2002.
30. Brockwell P, Davis R. *Introduction to Time Series and Forecasting* (2nd edn). Springer: New York, 2002.
31. Walpole RE, Myers RH. *Probability and Statistics for Engineers and Scientists* (2nd edn). Macmillan: New York, 1978.
32. Online resource, Auckland-VI trace data. <http://pma.nlanr.net/Traces/Traces/long/auck/6/> (Date of Access 02 May 2007).
33. Ta X. A Quality of service monitoring system for service level agreement verification. *Master Dissertation*, The University of Sydney, 2006.
34. Bar-Shalom Y, Li XR, Kirubarajan T. *Estimation with Applications to Tracking and Navigation*. Wiley: New York, 2001.

AUTHORS' BIOGRAPHIES



Xiaoyuan Ta received the Bachelor degree in Physics in Peking University in 2003, Master degree in telecommunications in the University of Sydney in 2005. Now, he is a PhD student in telecommunication in the University of Sydney, Australia. His research interests include wireless sensor networks, wireless localization techniques and network QoS.



Guoqiang Mao received the Bachelor degree in electrical engineering, Master degree in engineering and PhD in telecommunications engineering in 1995, 1998 and 2002, respectively. After graduation, he became a Senior Research Engineer at Intelligent Pixel Incorporation, Australia. He joined the School of Electrical and Information Engineering, the University of Sydney in December 2002 where he is a Senior Lecturer now. His research interests include wireless sensor networks, wireless localization techniques, network QoS, telecommunications traffic measurement, analysis and modelling, and network performance analysis.