

Copyright © 2006 IEEE. Reprinted from IEEE Journal on Selected Areas in Communications, 2006; 24 (12):2299-2312

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Adelaide's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

A Comparison of Poisson and Uniform Sampling for Active Measurements

Matthew Roughan, *Member, IEEE*

Abstract—Active probes of network performance represent samples of the underlying performance of a system. Some effort has gone into considering appropriate sampling patterns for such probes, i.e., there has been significant discussion of the importance of sampling using a Poisson process to avoid biases introduced by synchronization of system and measurements. However, there are unanswered questions about whether Poisson probing has costs in terms of sampling efficiency, and there is some misinformation about what types of inferences are possible with different probe patterns. This paper provides a quantitative comparison of two different sampling methods. This paper also shows that the irregularity in probing patterns is useful not just in avoiding synchronization, but also in determining frequency-domain properties of a system. This paper provides a firm basis for practitioners or researchers for making decisions about the type of sampling they should use in a particular applications, along with methods for the analysis of their outputs.

Index Terms—Error estimation, Internet measurement, measurement planning, network performance.

I. INTRODUCTION

ACTIVE measurement of system performance is an integral part of network quality assurance (NQA). Fault detection and service level agreement (SLA) verification are just two applications. A number of companies offer active measurement services or devices (e.g., Matrix NetSystems, Keynote, Niksun, Brix Networks, etc.), and there are now a number of Internet Engineering Task Force (IETF) Requests for Comments (RFCs) [1]–[5] describing standards for such measurement. However, the basic statistical properties of such measurements are still poorly understood.

Active probing is notionally simple. We wish to understand the performance of a packet network, and so send *probe* packets into the network, and measure the performance of these probes. These packets are samples of the performance of the underlying network. They are usually performed from end-to-end across a network, this being one of the chief advantages of such an approach: no special access to the network in question is required, and so measurements can be easily made by customers or researchers.

Many samples may be needed, both to provide better sampled statistics, and to allow observations of changes in the network over time. However, there are clear scalability issues: given N end hosts, there are N^2 end-to-end paths, and if each path requires many measurements this could create a significant

network load. Hence, the pattern of probing must be carefully planned.

Naively, the probe packets would be separated by a fixed sampling interval. However, there is the possibility that periodic samples may be synchronized with a periodicity in the system under observation, either by accident, or deliberate malfeasance. In this case, the probes might not observe the true system behavior.

It is not clear how prevalent periodicities are in the modern Internet, but there are certainly theoretical grounds for their existence. Auto-synchronization of network protocols [6], [7], large volumes of streaming traffic, or other applications that periodically make requests (such as NTP) may cause periodicities. Furthermore, regular samples are easy to predict, and therefore manipulate, perhaps by deliberately scheduling all data packets away from times when probes are anticipated. In addition, the probe packets themselves alter the network, and there is the potential that they cause the network to synchronize around their own period (in addition to other impacts, for instance, see [8]). Hence, there is reasonable cause to fear synchronization of probes and network.

Moreover, some sampling problems can occur where the samples and system periodicities are not synchronized, or even where the system is a-periodic but contains high-frequency periodic components that result in *aliasing*.

Poisson sampling steps away from naive uniform sampling by sending probes at the epochs of a Poisson process (PP). A result known by the acronym PASTA (Poisson Arrivals See Time Averages) ensures that (under relatively weak conditions) the average state observed by such probes will converge to the true average of the system under observation. Furthermore, when Poisson sampling is used, the next arriving packet cannot be anticipated, and so such measurements are harder to manipulate.

Poisson sampling, therefore, avoids some of the problems of uniform sampling. However, little work has considered the cost of this type of sampling. A number of comments have been made about difficulties introduced, for instance, the problems of making some types of measurements such as delay variation. At least one IETF RFC [9] has been written in response to these issues, but in many cases it is hard for a practitioner to decide which type of measurement would be best for their application. This paper is aimed at providing quantitative, and qualitative advice towards making such decision by providing a thorough analysis of the advantages of each approach.

For instance, little work has considered implications for measuring higher order statistics of the network (not just averages). Can Poisson sampling measure properties such as the delay variation or the power spectrum of a system? This paper demonstrates that Poisson sampling can be used to perform most of the

Manuscript received October 1, 2005; revised May 1, 2006.

The author is with the School of Mathematical Sciences, University of Adelaide, Adelaide 5005, Australia (e-mail: matthew.roughan@adelaide.edu.au).

Digital Object Identifier 10.1109/JSAC.2006.884028

tasks we might reasonably wish to perform with Internet measurements, e.g., estimation of

- 1) the mean, variance, and variance of differences (e.g., the delay variation);
- 2) the frequency content via a periodogram;
- 3) the autocorrelation;

without interpolation. This paper also presents results showing the advantages and disadvantages of such sampling. In summary the following.

- Poisson sampling results in a loss of efficiency, i.e., more measurements are needed to obtain results of the same accuracy, but the efficiency loss is much smaller than the natural variation of measurement efficiency as a result of changing network load.
- Both periodic and Poisson sampling can be used to perform harmonic analysis of the network to examine it for evidence of network periodicities. Poisson sampling actually has an advantage in this regard—the Lomb–Scargle Periodogram (Section IV) based on irregular sampling can detect frequencies greater than the periodic-sampling Nyquist frequency, allowing us to perform analyses that would be otherwise impossible.
- Autocovariances of the processes of interest can be estimated, and from these we can infer the delay variation metrics that we would obtain from uniform probes.

Poisson sampling is a special case of irregular sampling methods, but other types of irregular sampling would have many of the same benefits, and similar analysis could be applied. This is important because it is hard to produce a genuinely Poisson traffic stream.

This paper is concerned with comparing the two main types of probing, so that practitioners will know which to use in a particular circumstance, and researchers will have a better foundation for research in more complicated active probing problems such as network tomography or bandwidth estimation. Following the introduction, this paper presents some background, notation, and related work in Section II. This is followed by Section III, which provides results for the accuracy and efficiency of the two probing methods. Section IV compares the two methods on the basis of how they may be used to estimate time-series properties of the observed system, such as power spectra and autocorrelation functions. Such measures have been frequently used to characterize Internet systems (for instance, see the large literature on self-similarity in network traffic). Section V discusses some of the other issues, such as the issue of measuring properties such as reordering of packets. This paper concludes in Section VI.

II. BACKGROUND AND RELATED WORK

A. Performance Measurement Technologies

There are many measurements we may collect from a network: traffic, topology, or performance measurements. We focus here on measurements of network performance, though note that other supplemental measures may be required (e.g., network topology) to make sense of this data. Network performance can mean many things, for instance: reachability, delay, loss, jitter,

reordering, and bulk throughput. We can also form more complex functions of these metrics to attain measures such as the subjective performance of an application, e.g., VoIP. Many examples of tools to perform such measurements may be found at [10].

There are a number of ways in which we may collect data about network performance.

- **Direct measurements:** It is possible for a router or switch to maintain information about its own performance. For instance, to maintain data on the number of packets or bytes in buffers, or the number of packets dropped. In principle, such information could be collected without sampling (i.e., it could be collected in enough detail to reconstruct the exact sample path). However, in practice, there are limits on the (fast) memory required to store such information, and the rate at which it is collected. Despite its potential to be one of the best sources of performance measurements available, it is often one of the worst.
- **Active probes:** The more common method used to infer network performance is the well-developed approach of active probing, for instance, see [1]–[5], [11], and [12]. In this approach, we send *probe* packets into the network with precisely controlled departure times, and measure their arrival times elsewhere in the network. Such probes require installation of probe equipment into a network, but this equipment is typically fairly cheap, and it does not require special access to a network (for instance, the boxes could be installed by a customer, or researchers). There are many possible types of probe: e.g., ICMP echo probes, TCP SYN/ACK probes, DNS probes, HTTP page downloads, as well as dedicated probe protocols, and many types of measurements: e.g., round-trip or one-way measurements. These factors have led to active probing being the most widely deployed form of IP performance measurement.
- **Passive traffic measurements:** An alternative approach is passive traffic measurements. We can infer network performance through measurements of the arrival time of a packet at multiple points [13], [14]. This approach can also provide data of very fine detail, but it also has limits. First, such measurements are limited to the locations of packet monitors. Dedicated packet monitors are not typically expensive, but involve non-negligible installation and maintenance cost, and require privileged access to the network. More importantly, passive monitoring of this type can only infer performance on paths that carry traffic. Generally, passive measurements cannot control the sampling pattern, and so cannot guarantee Poisson or periodic samples.

For a practical comparison of some of the above techniques, see [15]. In this paper, we will be mainly concerned with active probing because we have control over the sample times, where with passive measurements we are reliant on existing data traffic. However, our control over sampling times only applies to samples of network performance—the time at which each individual queue in the network is sampled will be delayed by a random amount at previous queues. We discuss this issue more completely in Section V-B.

Even simple measurements such as delays are comprised of a number of components.

- 1) *Packet processing delay* is the delay to perform tasks such as forwarding table lookup, and is very small in modern high-speed routers (e.g., $\ll 1$ ms).
- 2) *Packet transmit time* is the time from starting to send the first bit of a packet onto the wire, until the last bit is finished. It is typically small for high-speed links, e.g., $\sim 4.8 \mu\text{s}$ for a 1500 byte packet on an OC48 link.
- 3) *Propagation delay* is the delay a packet experiences on the wire, and is given by the physical distance divided by the speed of light in fibre (roughly 200,000 km/s), e.g., ~ 30 ms for an East to West Coast in North America.
- 4) *Queueing delay* is the time spent by a packet in queues, which depends on load, and can potentially be quite large, e.g., 0.2 s, even on single OC48 line cards.

The two components that are significant, and therefore of primary interest in most modern networks, are propagation delay and queueing delay. Propagation delay is determined by network topology and routing, and for the purpose of this paper, we shall consider it to be a constant (see [16] for a more realistic view), which is derivable from other network data (topology and routing information), or from long-term measurements of the network. The queueing delays may be seen as drawn from a random process, and one goal of this paper is to estimate the behavior of this process.

In addition to statistical variations in queueing delays, there are measurement errors. Any set of performance measurements contains errors and artifacts such as delays in time-stamping a packet once it is received at a monitor. In a well-designed measurement system they should be an order of magnitude smaller than the queueing variations, and so we shall not consider these in detail here.

This paper focusses on network performance measurements, but the analysis in this paper could equally be applied to measurements of server performance. In principle, any system that can be modelled as a queueing system is susceptible to this type of analysis.

B. Statistical Notation

We shall denote a continuous-time stochastic process by $X(t)$. We denote the mean of the process at time t by $E[X(t)]$ (where $E[\cdot]$ denotes the expectation, or average), its variance by $\text{Var}(X(t))$, and its autocovariance by $\text{Cov}(X(t), X(t+s))$. We will be concerned with wide-sense stationary processes, where the mean, variance, and autocovariance are all constant with respect to t , and can consequently be written $E[X(t)] = \mu_X$, $\text{Var}(X(t)) = \sigma_X^2$, and $\text{Cov}(X(t), X(t+s)) = R_X(s)$, respectively, where we refer to s as the *lag*. We also define the autocorrelation function (ACF) of the process by $r_X(s) = R_X(s)/\sigma_X^2$. Note that the term ACF is used (in some texts) to denote $W_X(s) = E[X(t)X(t+s)]$, but we shall refer to this as the mean-zero autocovariance function (MZACF), as it is equivalent to the autocovariance $R_X(s)$ for a mean-zero process. In general, the relationship is

$$W_X(s) = R_X(s) + \mu_X^2 = \sigma_X^2 r_X(s) + \mu_X^2. \quad (1)$$

Another useful notation is an indicator function of the process, for instance

$$I(X(t) \in B) = \begin{cases} 1, & \text{if } X(t) \in B \\ 0, & \text{otherwise} \end{cases}. \quad (2)$$

Indicators show when arbitrary events occur. Note that $E[I(X(t) \in B)] = P\{X(t) \in B\}$, and we may therefore use indicators to construct other functions of the process.

1) *Sampling*: We shall assume that we can measure the process $X(t)$ at sample epochs T_1, T_2, \dots, T_N . The measurements form a discrete-time random process (or time-series) $X(T_1), X(T_2), \dots, X(T_N)$, which we will sometimes denote X_1, X_2, \dots, X_N . The two main cases considered in this paper are periodic sampling, where $T_n = nt_s$ for some intersampling interval t_s , and Poisson sampling, where the T_n form a PP (described below). When we apply periodic sampling, we shall equivalently use the term *uniform* sampling to be consistent with the related signal-processing literature.

We wish to estimate some parameter s of the stochastic process $X(t)$, and so construct an estimate $\hat{s} = f(X_1, X_2, \dots, X_N)$. We shall measure the accuracy of this estimate by the mean-squared error (MSE) $E[(\hat{s} - s)^2]$. Standard arguments show us that we can decompose this error into a variance term, and a bias term, i.e., $E[(\hat{s} - s)^2] = \text{Var}(\hat{s}) + (E[\hat{s}] - s)^2$, where the variance of the estimator $\text{Var}(\hat{s})$ is of prime interest here, as we shall consider unbiased estimators, i.e., estimators for which the bias $E[\hat{s}] - s = 0$, at least asymptotically as the number of measurements $N \rightarrow \infty$. For instance, the sample mean defined by

$$\hat{X}_N = \frac{1}{N} \sum_{i=1}^N X_i \quad (3)$$

is an unbiased estimate of the true mean of the stochastic process X_1, X_2, \dots, X_N given that this process is stationary and ergodic (where ergodicity is a technical condition we shall not discuss in detail here, but it ensures that a time average such as the sample mean above will converge to the true mean of the process). The variance of \hat{X}_N depends on the autocovariance of the sampled measurements X_i (we will consider this in more detail below).

It is simple to show that \hat{X}_N may not converge to the mean of a nonstationary process, for instance, consider the process $X(t) = a(t) + \cos(t)$, where $a(t)$ is a stationary mean-zero process. If sampled at time points $T_n = 2n\pi$, then $E[X_n] = 1$, whereas the average value $E[X(t)] = 0$. This highlights the property alluded to earlier, namely that uniform sampling, when precisely synchronized with a periodicity in the observed process may result in biased estimates. Should we worry about such a synchronization happening in reality? Certainly periodic behavior can happen in networks, but what are the chances that we inadvertently synchronize exactly to these phenomena? The problem is that active probes are just that—active. It is possible that they stimulate activity at precisely the frequency at which they arrive, thus causing the network to synchronize around these probes.

There are two methods typically used to avoid such problems. First, in signal processing, one passes a signal to be measured through a low-pass filter with cutoff threshold f_c , and then samples (uniformly) with sample frequency $f_s > 2f_c$. We shall discuss the reasoning behind this, and in particular, the problem of aliasing (of which the above is a special case) in Section IV, but note that prefiltering of the signal is not possible for the active probing methodology discussed above. The second method for avoiding the problem is to randomize the sampling intervals such that the epochs T_n form a PP, which is described below.

Note that a large part of this paper is applicable to other sampling processes T_n . This is important because:

- Even in uniform sampling there is often some jitter in the measurements (it is difficult to ensure probe packets are sent at precise intervals).
- Where Poisson sampling is used, there is some minimum time between packet probes (at least because of the finite size of packets), and this means that no probe stream is perfectly Poisson.

Frequency Domain Properties: We often wish to analyze a set of data to discover if it has periodic components. Such analysis is most easily accomplished in the frequency domain. A typical approach to obtaining such information is using the Fourier transform (FT), defined here with its inverse for a function $g(t)$ by

$$\text{FT}_g(f) = \int_{-\infty}^{\infty} g(t)e^{-i2\pi ft} dt, \quad g(t) = \int_{-\infty}^{\infty} \text{FT}_g(f)e^{i2\pi ft} df \quad (4)$$

or for uniformly sampled discrete-time data, we use the discrete Fourier transform (DFT)

$$\text{FT}_X(k) = \sum_{n=0}^{N-1} X_n e^{-i2\pi kn/N}, \quad X_n = \frac{1}{N} \sum_{k=0}^{N-1} \text{FT}_X(k) e^{i2\pi kn/N}. \quad (5)$$

To detect sinusoidal signals, we often plot the periodogram (though note there are many variations on this method, one of which we discuss below), which is proportional to the squared magnitude of the FT, for instance, in the discrete case

$$P_X(k) = \frac{1}{N} |\text{FT}_X(k)|^2 = \frac{1}{N} \left| \sum_{n=0}^{N-1} X_n e^{-i2\pi kn/N} \right|^2. \quad (6)$$

It is common to report the discrete spectra in terms of the index k , which we can convert to a frequency by taking $f = k/(t_s N)$ for uniformly sampled time series. The more data we have, the finer the frequency resolution, while the faster the sampling rate (smaller t_s), the wider a frequency band we can study (Section IV-A provides more detail of the Nyquist frequency, which specifies the frequency band we can study). Note that often frequencies are reported using $\omega = 2\pi f$ in units of radians per time interval.

Given a stochastic process $X(t)$, we typically talk about the *spectral density* rather than periodogram. Loosely, it gives the expected periodogram, though its actual values for a given random signal would be RVs. We shall abuse notation here and use P to indicate the spectral density. An important result in this context is the Wiener–Khinchine (WK) theorem, which

relates the spectral density to $W_X(s)$, by the Fourier transform, i.e., in the continuous case

$$P_X(f) = \int_{-\infty}^{\infty} W_X(t) e^{-i2\pi ft} dt. \quad (7)$$

In the discrete case, it is only true as the number of samples goes to infinity, but for finite N , we can take the approximation

$$P_X(k) \simeq \frac{1}{2N+1} \sum_{n=-N}^N W_X(t_s n) e^{-i2\pi kn/N} \quad (8)$$

and obviously we can determine W_X via the inverse transform. For both deterministic and stochastic signals, we refer to the term at $f = 0$ as the DC term.

C. The Poisson Sampling Process (PSP)

The PP is the renewal process formed by taking independent, identically distributed (i.i.d.) exponential random variables (RVs) $\{S_i\}_{i=1}^{\infty}$, and forming times $T_n = \sum_{i=1}^n S_i$. Formally, the PP is defined by the process that counts the number of renewals up to time t , i.e.,

$$N(t) = \sum_{i=1}^{\infty} I(T_n < t) \quad (9)$$

where $I(\cdot)$ is an indicator function.

The inter-renewal times S_i have an exponential distribution $P\{S_i < s\} = 1 - \exp(-\lambda s)$, where $\mu = 1/\lambda$ is the mean inter-renewal time, and for the homogeneous PP, λ is a constant called the *rate*. The mean of the PP is given by $E[N(t)] = \lambda t$, and its autocovariance function is given [17] by

$$R_N(t, s) = \begin{cases} \lambda s, & \text{if } t > s \\ \lambda t, & \text{if } t < s \end{cases}. \quad (10)$$

Given the PP is clearly nonstationary, we are not primarily interested in it, but rather in samples taken at the epochs of this process. We shall define a new process, namely, the PSP as the derivative of the PP, i.e.,

$$S(t) = \sum_{i=1}^{\infty} \delta(t - T_i) = \frac{dN}{dt} \quad (11)$$

where $\delta(t)$ is the Dirac Delta function. Expectation is a linear operator, and so

$$E[S(t)] = E\left[\frac{dN}{dt}\right] = \frac{dE[N(t)]}{dt} = \frac{d}{dt}(\lambda t) = \lambda$$

and we can likewise derive the autocovariance of $S(t)$ using the double derivative $E[S(t)S(s)] = E[(dN/dt)(dN/ds)] = (\partial/\partial t)(\partial/\partial s)E[N(t)N(s)]$, resulting in $R_S(t, s) = \lambda\delta(t-s)$. Note that the sampling process $S(t)$ is stationary, and so we write the autocovariance, and ACF as a function only of the lag $u = t - s$, e.g.,

$$r_S(u) = \delta(u). \quad (12)$$

The delta function is (by definition) zero for nonzero lags u , and so the PSP is uncorrelated in time. Another way of stating this is

to note that the PP has independent increments, a feature sometimes used in alternative definitions. This lack of correlation of the sampling process results in the key properties of interest here. Namely, that the number of samples to occur in nonoverlapping time intervals will be independent RVs. Hence, given information about the past of the process, we gain no information about its future—it is impossible to anticipate the samples.

It is also simple to derive the fact that the number of samples to occur in any arbitrary interval $[t, t + s]$ will be a discrete Poisson RV, with mean λs , and the events in the interval will be uniformly distributed over that interval.

1) *PSP in the Frequency Domain*: We can also represent a PSP in the frequency domain through the WK theorem (7). Note that for the PSP $W_S(u) = \sigma_S^2 r_S(u) + \mu_S^2 = \lambda \delta(u) + \lambda^2$. We can easily take the FT of this function to obtain spectral density

$$P_S(f) = \lambda^2 \delta(f) + \lambda. \quad (13)$$

The (almost) flat nature of the spectral density is often referred to as “white.” White noise refers to a process whose spectral density contains equal power in all bands (the metaphor refers to white light despite this not containing an equal spectral content). The PSP is perhaps the simplest *sampling* process to have this property.

2) *PASTA*: This property results from the fact that the system cannot anticipate the number of arrivals in a PSP. It seems instructive to consider why this leads to PASTA, and so we shall present a sketch of the key results here, in particular, to see the limitations of PASTA.

Formally, we define an indicator function $U_B(t) = I(X(t) \in B)$ for measurable set B , and then we define the time spent in state B , and the number of arrivals that see state B over interval $[0, t)$ by

$$V(t) = \int_0^t U_B(u) du, \quad Z(t) = \sum_{i=1}^{N(t)} U_B(T_i) \quad (14)$$

respectively, where $N(t)$ is the PP defined above. Although technical conditions such as left-continuity of sample paths are required, the PASTA theorem intuitively comes from one result: $Z(t)$ can be written as a Riemann–Stieltjes integral, and this integral can then be approximated by finite sums

$$\begin{aligned} Z(t) &= \int_0^t U_B(u) dN(u) \\ &\simeq \sum_{k=0}^{n-1} U_B(kt/n) \left[N\left(\frac{(k+1)t}{n}\right) - N\left(\frac{kt}{n}\right) \right] \end{aligned} \quad (15)$$

for sufficiently large n . For any arrival process which cannot be anticipated the number of arrivals in interval $[kt/n, (k+1)t/n)$ is independent of the state of the system at time t , and so

$$E[Z(t)] \simeq \sum_{k=0}^{n-1} E[U_B(kt/n)] E\left[N\left(\frac{(k+1)t}{n}\right) - N\left(\frac{kt}{n}\right)\right].$$

Now, in the case of Poisson samples, the number of arrivals in time $[kt/n, (k+1)t/n)$ is $\lambda t/n$, and so we can write

$$E[Z(t)] \simeq \lambda \frac{t}{n} \sum_{k=0}^{n-1} E[U_B(kt/n)] \simeq \lambda \int_0^t E[U_B(u)] du$$

for sufficiently large n (a real proof requires use of dominated convergence theorems to deal with the limit as $n \rightarrow \infty$). It is then a technical matter to show that convergence of $E[Z(t)]/E[N(t)]$ to $E[V(t)]/t$ occurs as $t \rightarrow \infty$ [18], [19].

The argument above relies on arrivals being Poisson, but it is important to note that the arrivals see time averages (ASTA) property applies for other arrival processes [18], [19]. The key detail is the lack of anticipation. For instance, note that none of the above requires stationarity of the process $X(t)$. If it is, and the increments of the sampling process $N(t+s) - N(t)$ are independent of the state of the process $X(t)$, then $E[U_B(t)] = E[U_B(0)]$, and we can write

$$\begin{aligned} E[Z(t)] &\simeq E[U_B(0)] \sum_{k=0}^{n-1} E\left[N\left(\frac{(k+1)t}{n}\right) - N\left(\frac{kt}{n}\right)\right] \\ &\simeq E[U_B(0)] E[N(t)] \end{aligned} \quad (16)$$

and once again we can prove ASTA. Note that here the arrival process $N(t)$ is no longer necessarily Poisson, however, the lack of anticipation property must still hold, and so we cannot use this result for an arbitrary arrival process. The lack of anticipation requires the system does try (deliberately, or through accidental synchronization) to anticipate the samples. In fact, the weak lack of anticipation (WLA) property [19] requires only that $N(t+s) - N(t)$ and $U(t)$ be uncorrelated for $0 \leq s \leq s_0$ for some s_0 (along with technical conditions on the continuity of the functions). The WLA property is true in many systems, but it is dangerous to assume so without any evidence. A simple example where WLA does not apply is where the system and measurements are synchronized in some way, but WLA is always true for Poisson samples. Hence, Poisson sampling is considered to be safer.

The above proof sketch is instructive because it shows that the ASTA property applies to point estimates of properties of the system. ASTA does *not* apply where we have a property that depends on the process at more than one point in time, such as the spectral density, or delay variation of the process. Such measures are derived for multiple time points in the system, bringing the system correlations into the estimates.

Other Properties of the PP: The intersample times of the PSP S_i take an exponential distribution and, therefore, the PSP is easy to generate by generating the S_i , but we could also generate a PSP by exploiting the fact that the number of events to occur in any arbitrary interval $[t, t + s]$ will be a discrete Poisson RV (with mean λs) and the events in the interval will be uniformly distributed over that interval. Also, note that the sum of two PPs also forms a PP, as does a PP which is thinned by randomly removing points with probability p .

The time between n samples will be the sum of n exponential distributions, which is given by an Erlang- n distribution [17], with density function

$$P \left\{ \sum_{i=1}^n S_i \in [s, s + ds) \right\} = \frac{\lambda(\lambda s)^{n-1} e^{-\lambda s}}{(n-1)!} ds. \quad (17)$$

This property is important in this context because the Erlang- n distribution approaches a normal distribution $N(n/\lambda, n/\lambda^2)$ for large n . As n becomes large, we could consider the Erlang times to be uniformly spaced with some small random jitter. Hence, while at one level, the PSP is completely random, if we take every n th probe, and only use this data, it is close to uniform sampling with probes spaced at n/λ intervals. So periodic subsamples from a PSP can approximate uniform sampling, and hence there may be less difference between Poisson and uniform sampling than is commonly assumed.

D. Related Work

The PASTA result [18]–[21] is well known in the communications networking community, as are the problems of synchronization of probes with network periodicities (e.g., see [6] and [7]) with a result that Poisson sampling is discussed in a number of IETF RFCs [3]–[5], [22], [23].

However, there has been little work on the topic since these foundations. One recent paper [24] starts to consider the issue. The paper uses a large number of experiments to consider whether the use of Poisson sampling makes any difference in practice. The conclusions are largely negative, i.e., they found little evidence for differences between Poisson and uniform samples of delays, losses, or packet-pair dispersion. There are intrinsic difficulties to such a study primarily because we can only compare the difference between the two, not how accurate they are with respect to the underlying process. Further, experiments cannot prove a negative—they cannot show that there is never a problem with network and probe synchronization. An additional paper [21] considers the difficulties of inverting measurements based on Poisson samples.

This paper will draw on the work contained in [8], in particular, methods for estimating the accuracy of measured network performance metrics. There are also a number of works on irregular sampling in other contexts. We discuss these in more detail in Section IV below.

III. ACCURACY AND EFFICIENCY

We wish to measure properties of a stochastic process describing the network behavior, and so it is natural to ask how accurately we can do so. Note that this is not accuracy as specified in the relevant RFCs, where it means the accuracy of the devices used for measurement. Here, we mean accuracy with which we can determine the parameters of a (wide-sense) stationary stochastic process (the observed delays). It is valid to measure a nonstationary process, for instance, to perform change detection for the purpose of detecting anomalies, however, when doing so, the concept of accuracy of the measurements has less meaning.

Given we shall typically apply unbiased estimators, the variance of the estimator is equal to the MSE. Variance is usually monotonically decreasing with the number of probes N , and so an alternative way of describing the problem is using statis-

tical efficiency, which describes how efficiently each data point is used in estimating parameters. More precisely, statistical efficiency is the minimum possible variance for an unbiased estimator divided by its actual variance. In general, this depends on the particular parameter to be estimated, and the estimator used. We shall focus on a simple parameter (the average delay), and use a simple statistic to estimate this parameter (the sample mean). This is not necessarily the best linear estimate of the mean of the data, but it is known to have high relative efficiency (>98%) compared with the best linear estimate (which is unknown *a priori*) [25].

The PASTA property [20] ensures that $E[X(t)] = E[X_n]$, at least in a limiting sense. However, PASTA says nothing about the rate of convergence, and it is intuitive that the variance would increase with Poisson sampling, given we are adding variability to our measurement process, but we show this is not necessarily true.

Take stationary stochastic process $X(t)$ and further simplify (without loss of generality) by shifting the process so that it is mean zero, and the variance of the sample mean

$$\begin{aligned} \text{Var}(\hat{X}_N) &= E \left[\hat{X}_N^2 \right] = E \left[\left(\frac{1}{N} \sum_{i=1}^N X(T_i) \right)^2 \right] \\ &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N E[X(T_i)X(T_j)]. \end{aligned} \quad (18)$$

Hence, the variance of the estimator depends on correlations between samples. For uniform samples $T_j - T_i = t_s(j - i)$, and hence for a mean-zero process $E[X(T_i)X(T_j)] = R_X(t_s(j - i))$. For Poisson samples, the times $T_j - T_i$ are Erlang- $(j - i)$ RVs, and so the expectation is given by a probability integral over the autocovariance. It is commonly assumed when making measurements that correlations drop to zero if we space the measurements sufficiently apart, and so we need not include these terms in our calculations of variance. However, in the Internet, the correlations may extend further than we can naturally space measurements [8], and furthermore, when performing Poisson sampling, the intervals between probes are random, and so some intervals may be very short. Hence, we must take account of the correlations in the system under observation.

The formula above is computable, but more instructive results regarding the above variance can be derived in the limit as $N \rightarrow \infty$. Limiting formula have been derived in the form of Central Limit Theorems. For uniform samples, the standard theorem states that

$$\sqrt{N}(\hat{X}_N - \bar{X}) \rightarrow N(0, s_X^2) \quad (19)$$

where $N(0, \sigma^2)$ denotes a normal distribution with zero mean, and variance σ^2 , and s_X^2 is referred to (see [26]) as the *asymptotic variance* of the process X , and it is defined by

$$s_X^2 \equiv \lim_{N \rightarrow \infty} N \text{Var}(\hat{X}_N). \quad (20)$$

We may compute the asymptotic variance for uniform samples using the following relationship (from [27])

$$s_{\text{uniform}}^2 = \sigma_X^2 \left(1 + 2 \sum_{n=1}^{\infty} r_X(nt_s) \right) \quad (21)$$

where t_s is the time between the samples, and the sum is finite (convergence of the sum is not guaranteed, with alternative results in that case).

A similar theorem for Poisson samples is derived in [8]

$$s_{\text{Poisson}}^2 = \sigma_X^2 \left(1 + 2\lambda \int_0^\infty r_X(u) du \right) \quad (22)$$

when the integral is finite. Note that if we take $\lambda = 1/t_s$, then the average rate of probes for both methods is the same. Given $t_s \rightarrow 0$, then $(1/\lambda) \sum_{n=1}^\infty r_X(nt_s) \rightarrow \int_0^\infty r_X(u) du$, so the two methods are consistent for high sampling rates.

The main result is that for certain types of autocovariance, either Poisson or uniform sampling could be superior. For instance, where the autocovariance has oscillatory behavior, we could construct a function such that the sum over r_X at regular points is either greater or smaller than the integral over the whole autocovariance function. Given that either case is possible, what is likely? We shall consider this in the case of a simple queueing model, the M/M/1 model.

A. Accuracy of Measurements of a Simple Queueing Model

One of the simplest queueing systems is the M/M/1 queue, i.e., a queue with a Poisson arrival process (rate λ) of packets whose service times are exponential (with mean $1/\mu$). We shall initially use this queue to model the effect of the two sampling regimes, because of its simplicity and the fact that:

- The sum of two PPs is also a PP, and we can simply model a system including Poisson probe traffic by adding the rate of probe traffic to data traffic with the result still having Poisson arrivals.
- The output of a M/M/1 queue is also Poisson, allowing us to obtain results for concatenated queues.

We denote the traffic intensity $\rho = \lambda/\mu$. The queue is stable with a finite average buffer length for $\rho < 1$, and we restrict our attention to this region. The M/M/1 queue is well studied, with many text book results, e.g., see [28]. For instance, the mean and variance of the number of packets in the system are

$$\mu_Q = E[Q] = \frac{\rho}{1-\rho}, \quad \sigma_Q^2 = \text{Var}(Q) = \frac{\rho}{(1-\rho)^2}.$$

The mean and variance of the waiting times are

$$E[W] = \frac{1}{\mu} E[Q], \quad \text{Var}(W) = \frac{1}{\mu^2} (\mu_Q + \sigma_Q^2).$$

The M/M/1 autocovariance results are not as simple, because of their dependence on the transient behavior of the M/M/1 queue. We can find the autocovariance of this queue in [29] and [30]. The autocorrelation function is given by

$$r_Q(s) = \frac{(\mu - \lambda)^3}{\pi} \int_0^{2\pi} \sin^2 \theta \frac{e^{-w|s|}}{w^3} d\theta \quad (23)$$

where $w = \lambda + \mu - 2\sqrt{\lambda\mu} \cos \theta$. This is not quite the same as the autocorrelation of the delays, but it is close enough to give us the required insight for this paper, whereas the formula for waiting times are considerably more complicated, without providing any additional insight. Morse [29] gives the integral of the ACF $r_Q(s)$ over s as does [26, eq. (22)], but note the time

scaling of $1/\mu$ in [26]. Given such a scaling, the observation time is measured in units of number of (average) service times. Given this integral we can analytically compute the asymptotic variance for Poisson sampled measurements of the number of packets in the system, and this is closely approximated by

$$s_B^2(p, \lambda, \mu) \simeq \frac{\rho^2}{(1-\rho)^2} + p \frac{4\rho^3}{(1-\rho)^4} \quad (24)$$

where we note that the Poisson samples compose proportion p of the traffic sent to the queue, so that $\lambda_s = p\lambda$, where λ is the total arrival rate of packets to the system.

We do not have a closed form for the discrete sum (21), but it can be computed numerically by summing over the above autocorrelation, which we can evaluate using numerical integration. Alternatively, we can apply the approximation given in [31, eq. (3.7)] i.e.,

$$r_Q(s) \simeq \frac{1}{2} \left[e^{-A|s|} + e^{-B|s|} \right]$$

where

$$A = \frac{(1-\rho)^2}{1+\rho+\sqrt{\rho}}, \quad B = \frac{(1-\rho)^2}{1+\rho-\sqrt{\rho}}. \quad (25)$$

This form can be analytically integrated to give

$$\int_0^\infty r_Q(s) ds \simeq \frac{2(1+\rho)}{2(1-\rho)^2} \quad (26)$$

and summed to give

$$\sum_{n=1}^\infty r_N(nt_s) \simeq \frac{1}{2} \left[\frac{e^{-At_s}}{1-e^{-At_s}} + \frac{e^{-Bt_s}}{1-e^{-Bt_s}} \right]. \quad (27)$$

Fig. 1 shows a comparison of the asymptotic variances for Poisson and uniform sampling. The figures show both the approximation (lines), and the exact numerical results (markers). Fig. 1(a) shows the impact of varying the sampling interval. The figure shows that Poisson sampling has a higher asymptotic variance, which should be expected for a monotonically decreasing ACF. Note also that the figure seems to indicate that the two asymptotic variances converge for $t_s \rightarrow 0$ as we would expect. Fig. 1(a) shows the results for a moderate load $\rho = 0.75$, but the results above show a dependence on ρ , which we explore in Fig. 1(b) and (c), by showing (for a fixed sample rate) the results for a range of values of ρ . For $t_s = 100$, we note significant variation between the two asymptotic variances, but that the differences decrease for heavy and light loads.

However, the most interesting thing to note from Fig. 1(b) and (c) is the dramatic dependence of the asymptotic variance on the system load ρ . As noted in [8], the total traffic rate has a very significant impact on measurement accuracy. This was argued in [8] to be a much more general effect than we have demonstrated here (for the M/M/1 queue). It is important here because it shows that there is *much* more variation in measurement accuracy between measurements taken at different loads, than we might see using different sampling methodologies. The variations for different sampling methods should be seen as a minor effect in comparison.

Fig. 1 also shows that for large sampling rates the asymptotic variance increases. This results from the fact that the samples

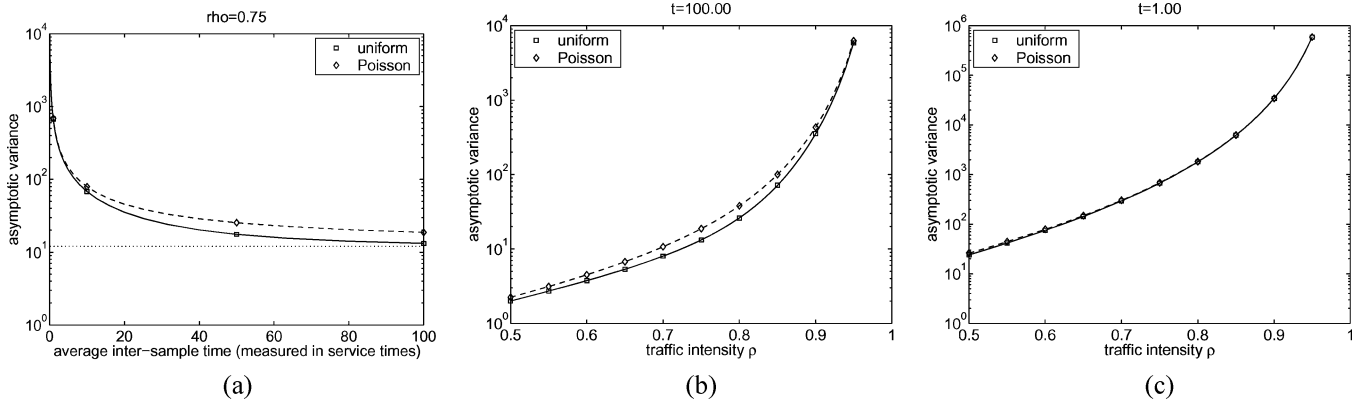


Fig. 1. Asymptotic variance for Poisson and uniform sampling of the M/M/1 queue. The lines show results derived from approximations (26) and (27), whereas the markers show results derived from direct numerical integration. Note time is given in units of average service times. (a) Asymptotic variance for $\rho = 0.75$ and a range of sampling rates. (b) Asymptotic variance for range of traffic intensities ρ with $t_s = 100$. (c) Asymptotic variance for range of traffic intensities ρ with $t_s = 1$.

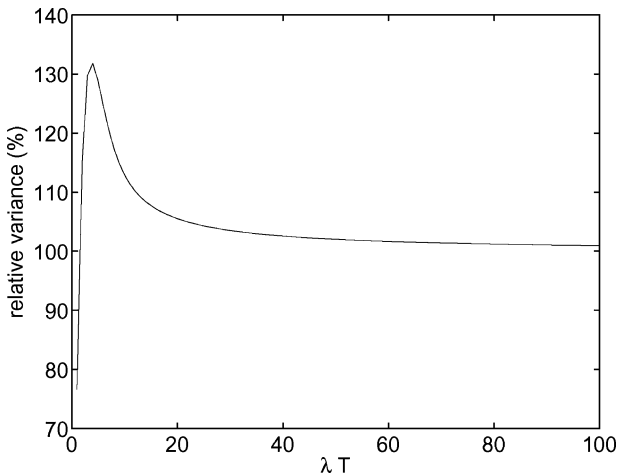


Fig. 2. The relative variance of the sample mean for i.i.d. RVs generated by uniform versus Poisson samples.

are more closely spaced, and so they are more correlated, and less additional information is added per sample.

B. Fixed Time-Interval Sampling

One issue that arises in using Poisson sampling is the fact that the time taken to obtain a fixed number N of samples varies. Alternatively, given a fixed time interval T , the number of samples will be random. The results above tell us the asymptotic variance for a fixed number of samples. We should also consider the case of a fixed time interval.

Given a fixed time interval, the number of samples N_T will be distributed as a Poisson RV, i.e.,

$$P\{N_T = n\} = \frac{(\lambda T)^n \exp(-\lambda T)}{n!} \quad (28)$$

which has mean λT and variance λT . For large values of λT , we may approximate N_T by a normal distribution, i.e., $N_T \sim N(\lambda T, \lambda T)$. The variance of N_T is an additional source of variance in the results.

To analyze this, let us consider the simple case of uncorrelated measurements. In this case, the sample mean, conditional on N_T is has variance σ_X^2/N_T , for $N_T > 0$. Note that for true

Poisson sampling there is a non-negligible chance of zero samples occurring, so we condition on at least one sample, resulting in a distribution

$$P\{N_T = n\} = \frac{(\lambda T)^n \exp(-\lambda T)}{(1 - \exp(-\lambda T)) n!}. \quad (29)$$

Take expectation over the conditional variance, and use the notation \hat{X}_T to denote the sample mean of Poisson samples over a fixed time interval, we can see that

$$\text{Var}(\hat{X}_T) = E[\text{Var}(\hat{X}_{N_T})] = \sigma_X^2 \frac{e^{-\lambda T}}{(1 - e^{-\lambda T})} \sum_{n=1}^{\infty} \frac{(\lambda T)^n}{n \cdot n!}.$$

Fig. 2 shows the relative variance of the sample mean for i.i.d. RVs generated by uniform versus Poisson samples, i.e., the figure shows $\text{Var}(\hat{X}_T)/\text{Var}(\hat{X}_N)$, where $N = \lambda T$. The figures shows that for small λT , the variance in the Poisson samples increases the variance of the estimator by up to 32%, but that this increase quickly becomes negligible. After only around 50 samples ($\lambda T = 50$), the loss in efficiency of the two becomes around 2%, and given the number of samples collected in many experiments, this can be considered to be a negligible source of error.

The drop below 100% for $\lambda T = 1$ arises because in this case, uniform sampling results in exactly one sample, but for Poisson samples, we must condition on at least one sample being made otherwise we cannot estimate the delay at all. Hence, there is a notional advantage to Poisson sampling for low λT , but it arises due to the conditioning, and so should not be given much weight.

C. Impact of Active Probes

The results above neglect the impact of the active probes themselves (results are reported for traffic intensities, irrespective of what proportion of this traffic is probe traffic, or genuine data traffic). The work of [8] showed that the impact of active probes themselves was substantial (in that it increased load, and therefore correlation scale, and hence the asymptotic variance). We will not repeat this discussion here, due to limitations in space.

However, there is one respect in which this work differs from [8]. When Poisson probes are combined with Poisson traffic, the resulting input traffic is still Poisson. However, when uniform

probes are combined with Poisson traffic, the result now deviates from Poisson, which will distort the behavior of the queue. This type of effect should be minor for low sampling rates.

D. Long-Range Dependence

No-one would claim that the M/M/1 queue is a good model for Internet performance. A more realistic model would include features such as long-range dependence (LRD). As noted in [8], LRD changes the quantitative results described above, but qualitatively similar phenomena are observed. See [8] for some informative simulations. Most importantly, where we have LRD, the autocovariance function is not integrable, and so the CLT used above in (21) and (22) will not hold. In fact, we need to use a generalized CLT, which will result in much slower rates of convergence.

IV. TIME-SERIES PROPERTIES

Many time-series algorithms are designed with uniformly sampled data in mind. For instance, a problem of interest in the Internet is the detection of synchronization effects. These have had an impact of sampling methods, but moreover, synchronized oscillations often result in reductions in performance. A case in point is TCP congestion-control synchronization [6], but it has been noted in other contexts such as routing [7]. However, there are few papers concerning detection of such oscillations in practice (for one example see [32]), perhaps in part because detection of such frequency components using irregularly sampled data seemed to be a hard problem. This paper demonstrates that this should not be a major issue, i.e., that we can detect periodic components of signals even where the data is irregularly sampled.

The problem of detection of periodic components of uniformly sampled data has been extensively studied. There are a number of approaches, but the simplest (conceptually) is to take the DFT, and examine the corresponding periodogram for peaks that would correspond to a periodic signal. There are a number of issues related to this method, many of which we shall omit in this paper (e.g., those concerning numerical properties of the periodogram). However, there is one critical issue, that of the Nyquist frequency, which we shall discuss further in Section IV-A.

One method to apply time-series algorithms to Poisson data would be to resample the data onto a uniform grid through interpolation, but in doing so, we run the risk of introducing unknown interpolation artifacts. However, in fields such as astronomy where irregularly observations are unavoidable, considerable work has been devoted to analysis of such time series. In particular, we can determine periodicities present in such data using the Lomb–Scargle Periodogram (LSP). This is a natural generalization of the standard periodogram defined in (6). It can be seen as analogous if we expand the complex exponential in the periodogram (6) in $\omega = 2\pi f$ to give

$$P_X(\omega) = \frac{1}{N} \left[\left(\sum_{k=1}^N X_k \cos(\omega T_k) \right)^2 + \left(\sum_{k=1}^N X_k \sin(\omega T_k) \right)^2 \right].$$

The LSP [33]–[35] is instead defined by

$$P_X^{(LS)}(\omega) = \frac{1}{2} \left[\frac{\left(\sum_{k=1}^N (X(T_k) - \bar{X}) \cos(\omega(T_k - \tau)) \right)^2}{\sum_{k=1}^N \cos^2(\omega(T_k - \tau))} + \frac{\left(\sum_{k=1}^N (X(T_k) - \bar{X}) \sin(\omega(T_k - \tau)) \right)^2}{\sum_{k=1}^N \sin^2(\omega(T_k - \tau))} \right]$$

where τ is a frequency-dependent time shift included to make the periodogram time-shift invariant and defined by

$$\tan(2\omega\tau) = \frac{\sum_{k=1}^N \sin(2\omega T_k)}{\sum_{k=1}^N \cos(2\omega T_k)}.$$

The LSP $P_X^{(LS)}$ gives us information about the frequency content of the signal at frequency ω . There are several methods for deriving, or justifying the LSP. For instance, as a periodogram that guarantees a flat spectral density when applied to irregularly sampled white noise [34], or which results from performing a least-squares fit of sinusoids to the irregularly sampled data (the standard periodogram can be thought of in the same way for regularly sampled data) [33]. The statistics of the LSP are known, and so we can determine simple hypothesis tests for the presence, or absence of particular frequencies in the measured signal.

As describe above, the algorithm is not particularly efficient. However, fast $O(N \log N)$ algorithms for the computations of the LSP exist [33], [36], much as the FFT exists for the DFT.

Though it has been explicitly applied to Poisson sampled data [37], the LSP is not just applicable to Poisson sampled data. For instance, passive sampling of packet delays will also result in an irregularly sampled times series, because sample times cannot be controlled, and we could apply the LSP here.

In this paper, we shall again use the M/M/1 queue as an example. In particular, consider the ACF of the M/M/1 queue, given in (23). The WK theorem (7) allows us to compute the spectra of the M/M/1 queue from this relationship by taking the FT of $W_Q(s)$. To gain some insight into the shape of the spectra, we can use the FT of approximation (25) to analytically derive an approximation for the spectra

$$\text{FT} \{r_Q(t)\} \simeq \frac{A}{A^2 + 4\pi^2 f^2} + \frac{B}{B^2 + 4\pi^2 f^2}. \quad (30)$$

The discrete spectra turn out to be a little different. Taking the DFT of $r_W(nt_s)$, we get two terms of the form

$$\begin{aligned} & \sum_{n=-N}^N e^{-A|nt_s|} e^{-in\omega} \\ &= 1 + \sum_{n=0}^N e^{-nAt_s} e^{-in\omega} + \sum_{n=0}^N e^{-nAt_s} e^{in\omega} \\ &= 1 + \sum_{n=0}^N \left(e^{-(At_s - i\omega)} \right)^n + \sum_{n=0}^N \left(e^{-(At_s + i\omega)} \right)^n \\ &\rightarrow 1 + \frac{e^{-(At_s - i\omega)}}{1 - e^{-(At_s - i\omega)}} + \frac{e^{-(At_s + i\omega)}}{1 - e^{-(At_s + i\omega)}} \end{aligned} \quad (31)$$

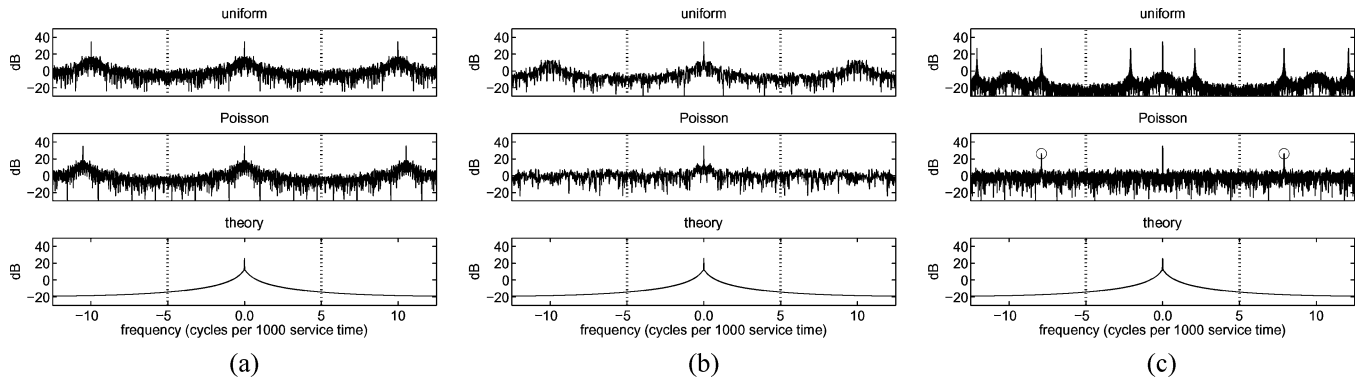


Fig. 3. Periodograms of the M/M/1 queue with $\rho = 0.95$. In each figure, the top plot shows the periodogram using uniform sampling, the middle shows Poisson sampling, and the lower plot shows the theoretical spectral density of the queueing process. The Nyquist frequency is shown by vertical dashed lines. (a) Standard periodogram. (b) LSP. (c) LSP of M/M/1 plus an additional cosine with frequency $f = 1.5/(2Et_s)$ illustrating aliasing for uniform sampling. The cosine signal can be unambiguously detected using the LSP and Poisson sampling (see circles) even though it lies above the Nyquist frequency.

in the limit as $N \rightarrow \infty$. In addition to this term, we add a similar one with A replaced by B , and the mean queue results in an additional $\mu_Q^2 \delta(k)$ term.

Fig. 3 shows the standard and LSPs for uniform and Poisson sampling of a simulated M/M/1 queueing process with $t_s = 100$ unit service times, $\rho = 0.95$, and $N = 100,000$ arrivals (note the first 5000 are discarded to avoid transients). The total number of samples is therefore near 1000 in each case. We deliberately choose a high-load example to illustrate the spectrum more clearly. Cases with low load are qualitatively similar, though the periodograms become flatter (apart from the delta at $f = 0$) as the strength of the correlations is reduced. The figure also shows the theoretical spectral density as described above.

Fig. 3(a) shows the periodogram for uniform sampling (top plot), which matches what we would expect from the theoretical spectral density (bottom plot) within the Nyquist band (vertical dashed lines), but thereafter repeats periodically due to aliasing (see below). Surprisingly, the standard periodogram for Poisson sampling (middle plot) also appears to match the true spectral density within the Nyquist band. We will discuss this in Section IV-B, but note that it is not a generic property for Poisson sampling, but one peculiar to the M/M/1 queue.

Fig. 3(b) shows the LSPs for uniform and Poisson sampling. We will discuss these further in Section IV-A but note that the LSP for Poisson sampling (middle plot) differs noticeably from the theoretical spectral density. The main difference is that the noise floor for the LSP is around 0 dB, masking the true spectrum of the process for frequencies greater than one cycle per 1000 service times. On the other hand, it avoids the aliasing effects seen in the other plots.

A. Aliasing

When data is uniformly sampled at frequency f_s (that is, the samples are separated by times $t_s = 1/f_s$), we lose data (concerning the signal between samples), and this in turn may lead to an ambiguity. Signals with different frequencies may result in identical samples, e.g., Fig. 4 shows an example of two sinusoids which could generate the same set of sample data points (shown in the figure as circles). This problem is generally referred to as aliasing. Aliasing can be avoided if the highest frequency present in the signal has frequency $< f_s/2$, the so called

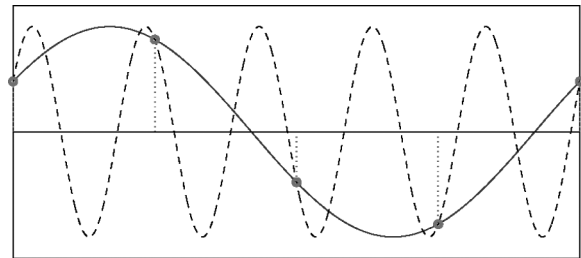


Fig. 4. Aliasing in the time and frequency domains. The figure shows two sinusoids that generate the same set of sample points (circles).

Nyquist frequency, with the result that a signal $f(t)$ can unambiguously be reconstructed from its samples if sampled at twice the rate of the highest frequency present.

Fig. 3 illustrates the problem of aliasing in our context. The top plots show the periodograms for uniform samples. We can see that the central peak around $f = 0$ is aliased to create additional peaks outside the Nyquist band (shown by vertical dashed lines). This is what one typically expects to see in the frequency domain—periodic repetitions of the spectra of the process at intervals of the Nyquist band.

In fact, the bias introduced when we uniformly sample a system synchronized with its periodicities is a special case of aliasing. Frequency $f = n/t_s$ is aliased to the DC term in the spectrum, hence distorting the estimate of the mean.

In order to ensure that aliasing not occur, signals are often filtered prior to sampling, typically with a low-pass analogue filter with cutoff below the Nyquist frequency. Given such filtering, we can unambiguously detect periodic components of the signal, but not over an arbitrary range—signals above the cutoff are effectively removed from the signal, and so are invisible. However, there are a number of applications where the sampling methodology makes prefiltering the signal inherently hard. As in many astronomical observations, Internet measurement samples are obtained by experiments, and we have no opportunity to low-pass these samples prior to performing the experiments. This results are a problem for Internet measurements. How can we prevent aliasing?

As it turns out, irregular sampling has its key benefit here. It has been shown [38] that the analogue of the Nyquist frequency is considerably higher for irregularly sampled data than for uniformly sampled data. The naive intuition is that the increased

Nyquist frequency arises because with irregular sampling, some samples will be closer together than the average sampling distance, thereby removing ambiguity. However, it is not true that the new Nyquist frequency is $1/2 \min(T_{i+1} - T_i)$. In fact, [38] shows considerably higher frequencies may be resolved (it uses 122 data points of astronomical data, and successfully finds periodic behavior with frequency of the order of 100 times the uniform sampling Nyquist limit).

Fig. 3(b) illustrates this property. Note that when Poisson sampling is used (middle plot), the aliasing of the central peak to higher frequencies is eliminated. Fig. 3(c) further illustrates aliasing. In this case, we add a cosine wave with frequency $1.5 \times$ the uniform sampling Nyquist frequency to the queueing process before sampling. When uniform sampling is used (top plot), we see multiple aliased peaks resulting from the cosine function. When Poisson sampling is applied (middle plot), the two genuine peaks appear at the correct frequency. The cosine has frequency outside the uniform sampling Nyquist band, but can still be unambiguously detected using the LSP.

To correctly understand the result, we must correctly understand the impact of sampling in the Fourier domain, often referred to as the *window function*. We may think of sampling a continuous function as the process of multiplying this function by a sample process, i.e.,

$$(X_n)(t) = X(t)S(t) = X(t) \sum_{k=-\infty}^{\infty} \delta(t - T_k). \quad (32)$$

The FT of the product of two functions is the convolution of the respective FTs, and so

$$FT\{(X_n)\} = FT\{X(t)\} * FT\{S(t)\} \quad (33)$$

where $*$ denotes the convolution operation. In the case of uniform sampling

$$FT\left\{\sum_{k=-\infty}^{\infty} \delta(t - kt_s)\right\} = \sum_{k=-\infty}^{\infty} t_s \delta(\omega - k/t_s) \quad (34)$$

which, when convolved with the true spectrum, results in repetitions of the spectrum at intervals of $1/t_s$. These repetitions do not overlap if the spectral density of the spectra is isolated into a band $1/t_s$ in width. Note also that the FT of a real process is even, and so there is only half this band available for nonredundant information.

On the other hand, the spectral density of the PSP is given in (13). We can see that it is a constant, plus a δ DC term. Hence, when we convolve with this spectral density, the δ term results in the correct spectra (without repetitions), while the constant results in even spreading of the spectral content of the signal across all bands, albeit at a lower level than the signal of interest. Hence, this explanation provides an intuitive understanding of the origin of the noise in the LSP for the Poisson sampling cases. It also suggest that we can find periodic components of arbitrary frequency in a signal (i.e., there is no Nyquist limit), but note that this property only applies in the limit as the number of samples goes to infinity. However, even for finite samples, we can certainly detect frequencies well above $1/2t_s$.

Fig. 3(b) shows the increase in background noise that occurs in the LSP. Although methods exist to reduce this, it is still an issue if we are attempting to characterize the spectral density of

the process. On the other hand, if we wish to detect the presence (or absence) of a limited number of sinusoids in the time series, then it is quite practical to do so, and standard statistical tests exist [37]. Hence, this approach is quite viable should we wish to, for instance, detect network synchronization effects.

B. Autocorrelations

The standard estimator of autocovariance for regularly sampled data is well known. Obviously, given such samples we do not get access to the autocovariance of the continuous process $X(t)$ from which we are sampling, but we can obtain asymptotically unbiased estimates at lags that form integer multiples of the sampling interval t_s using the estimator

$$\hat{R}_X(nt_s) = \frac{1}{N} \sum_{k=1}^{N-n} [X_k - \hat{X}_N][X_{n+k} - \hat{X}_N] \quad (35)$$

where \hat{X}_N is the sample mean of the process.

This estimator is predicated on uniform sampling. Consider its behavior when Poisson sampling is applied. We shall do so for a mean-zero sampled process $X(T_k)$, in order to make the analysis simple. For such a process, denote our estimate (not necessarily of R_X) by

$$\hat{Q}(n) = \frac{1}{N} \sum_{k=1}^{N-n} X(T_k)X(T_{n+k}) \quad (36)$$

then we can rearrange to find

$$\begin{aligned} E[\hat{Q}(n)] &= \frac{1}{N} \sum_{k=1}^{N-n} E[R(T_{n+k} - T_k)] \\ &= \frac{1}{N} \sum_{k=1}^{N-n} \int_0^{\infty} R(s) P\{T_{n+k} - T_k = s\} ds \\ &= \frac{1}{N} \sum_{k=1}^{N-n} \int_0^{\infty} R(s) \frac{\lambda(\lambda s)^{n-1} e^{-\lambda s}}{(n-1)!} ds \end{aligned} \quad (37)$$

given that $T_{n+k} - T_k$ is the sum of n exponential RVs and therefore distributed as an Erlang- n RV. Note that the terms inside the summation do not depend on k , and so

$$E[\hat{Q}(n)] = \frac{N-n}{N} \int_0^{\infty} R(s) \frac{\lambda(\lambda s)^{n-1} e^{-\lambda s}}{(n-1)!} ds. \quad (38)$$

Clearly, $\hat{Q}(n)$ may not be a good estimator of $R_X(nt_s)$.

It is interesting to consider the case of Poisson sampling of the M/M/1 queue. We shall do so by placing the approximation to $R(s)$, e.g., (25), for the M/M/1 queue into (38), such that we get two integrals of the form

$$\begin{aligned} &\left(\frac{N-n}{N}\right) \frac{\lambda \sigma^2}{2} \int_0^{\infty} \frac{(\lambda s)^{n-1} e^{-(\lambda+A)s}}{(n-1)!} ds \\ &= \left(\frac{N-n}{N}\right) \frac{\sigma^2}{2} \frac{1}{(1+A/\lambda)^n} \\ &\rightarrow \frac{\sigma^2}{2} \exp(-An/\lambda) \end{aligned} \quad (39)$$

for large n and N . Hence, we can see that for the M/M/1 queue, the Poisson sampled ACF approaches the true ACF in the limit

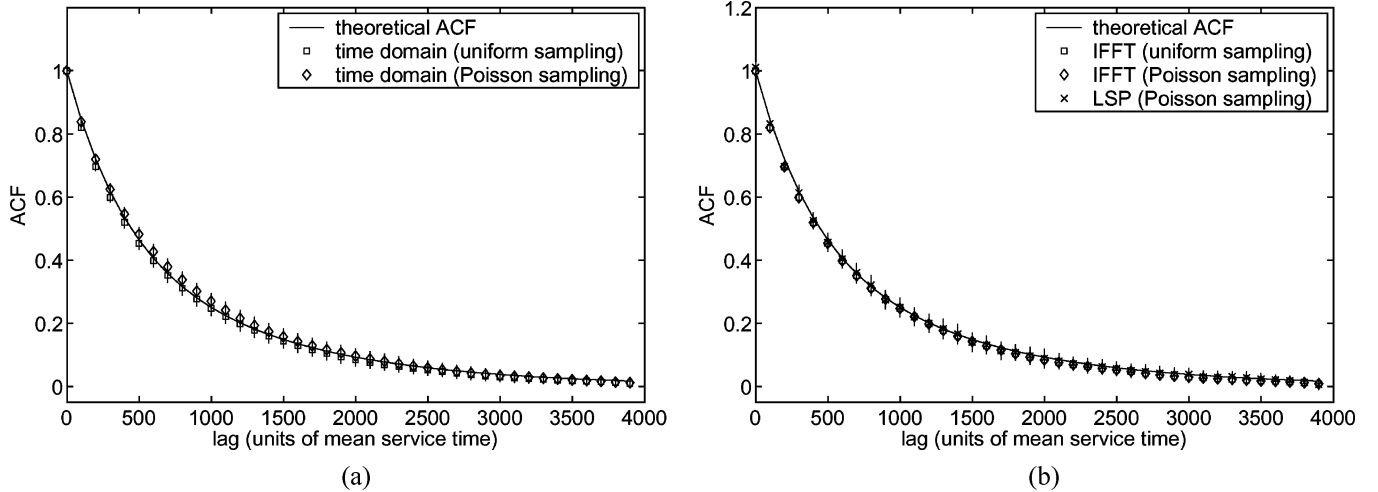


Fig. 5. The estimated and theoretical ACF for $\rho = 0.95$ and $t_s = 100$. Vertical lines show 95th percentile confidence intervals. (a) Time-domain estimates using the estimator (35). (b) Frequency-domain estimates, obtained by taking the periodogram and applying the inverse DFT.

for a large number of samples, and large lags. See Fig. 5 to see how close the approximation is in a real case. In conjunction with the WK theorem (7), we can now see why the standard periodogram reported in Fig. 3(a) (middle plot) gives almost correct results for Poisson sampling. Given the ACF is estimated reasonably using Poisson sampling, we should expect the spectra to be equally correct.

The results show two things. First, PASTA as a result is not necessarily true for higher order statistics of a process. This is important to realize when we consider measurement of statistics such as delay variation. In the M/M/1 case, the distortion is minimal, but could perhaps be larger for other processes, though in general, we might expect estimates for very large lags to be reasonable, given higher order Erlang distributions tend towards a deterministic distribution.

To obtain a true estimate of the ACF, we could use interpolation of our original data set, but as before interpolation may introduce unknown artifacts. Alternatively, we might bin the time differences $t_i - t_j$ and compute an approximation to the autocovariance over such intervals. However, in this approach, information is lost.

On the other hand, the WK theorem, which relates the spectral density to $W_X(s)$, gives a method for computing the ACF by applying an inverse FT to the periodogram [37]. Starting from a LSP plotted with the standard range and resolution for uniform sampling, we obtain estimates at discrete lags $\hat{W}_X(nt_s)$ (as we would with the time-domain estimator). We can plot a LSP at higher frequencies, and with finer resolution than the standard periodogram, resulting in finer resolution when we produce the resulting ACF, however, note that the same information is represented in the periodogram, and so this represents a form of implicit interpolation [34].

There is one important detail to correct when applying the inverse FT. As noted above, the FT of the sampled data is a convolution of the signal of interest with the FT of the window function. When we compute the periodogram, we take the magnitude of this function, and the result is a distortion of the spectra. The distortion is passed back to the time domain by the inverse DFT and so the measured ACF will be

$$\hat{r}(nt_s) = \hat{r}_X(nt_s)r_g(nt_s) \quad (40)$$

where r_g is introduced by the sampling. This effect must be removed to obtain $\hat{r}_X(nt_s)$, by dividing through by $r_g(nt_s)$.

Fig. 5 shows ACF results for various approaches to estimation for a set of 30 simulated samples of the M/M/1 queue. The results use somewhat more samples (of the order of 10 000), as we might expect when trying to estimate higher order moments of a process. Fig. 5(a) shows time-domain estimates based on uniform and Poisson samples, including 95th percentile confidence intervals. We can see that over the range plotted they are both reasonably close to the true ACF, though note that (as described above) this is not a guaranteed property of Poisson samples, but rather a property that arises when sampling the M/M/1 queue. Fig. 5(b) shows frequency-domain estimates derived by estimating the periodogram, and inverting. Three techniques are displayed, the first two being based on the standard periodogram of both uniform and Poisson samples to illustrate that these produce almost identical results to their time-domain equivalents. The last is based on a LSP, and shows that we could estimate the ACF using Poisson sampling.

A similar method (inverting spectra) can also be used to derive cross correlations for two sets of measurements [37].

C. Delay Variation

Delay variation (as defined by the IETF [23]) is intended to capture a notion of the variability of delay measurements. A possible application is the adaptive design of playout buffers, which must buffer enough data so that even if some packets are delayed a steady stream of data can be provided to an application. RFC 3393 [23] specifies delay variation be measured by examining differences in delays, i.e., we consider the process $Y_i = X_i - X_{i-1}$. Such a process gives us some idea of the variation of the delays. The RFC describes a number of useful statistics of such measurements—we shall consider its variance here, i.e., $\text{Var}(Y_i)$.

In the case of uniform sampling, the interpretation of this metric is typically related to an application which sends traffic at the same intervals as we sample. However, it should be obvious from our discussion of correlations in queueing that $\text{Var}(Y_i)$ depends on the sampling interval. In fact, we can derive the vari-

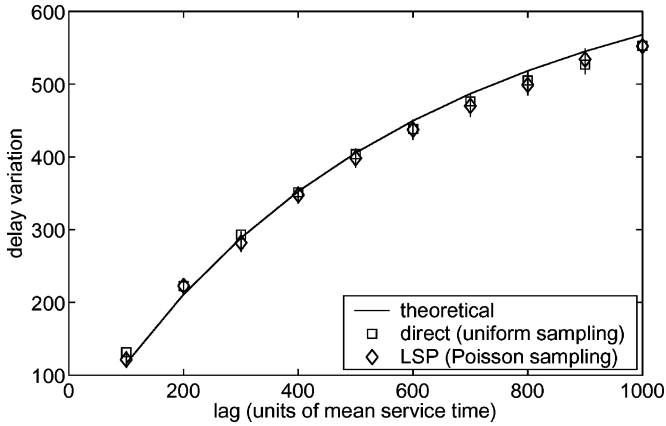


Fig. 6. A comparison of the direct computation of delay variation from uniform samples and estimates obtained from inverting the LSP for Poisson samples for the M/M/1 queue with $\rho = 0.95$, and $t_s = 100$. Vertical lines show 95% confidence intervals which overlap for both methods.

ance explicitly in terms of the autocovariance of $X(t)$. First note that for a stationary process $X(t)$

$$E[Y_i] = E[X_i] - E[X_{i-1}] = 0. \quad (41)$$

Now, given $E[Y_i] = 0$

$$\begin{aligned} \text{Var}(Y_i) &= E[Y_i^2] = E[(X_i - X_{i-1})^2] \\ &= E[X_i^2 - 2X_iX_{i-1} + X_{i-1}^2] \\ &= 2E[X_i^2] - 2E[X_iX_{i-1}]. \end{aligned} \quad (42)$$

Let us further simplify (without loss of generality) to consider X_i to be mean zero, and we see that

$$\text{Var}(Y_i) = 2\sigma_X^2 - 2E[R(T_i - T_{i-1})]. \quad (43)$$

For uniform sampling, this reduces to

$$\text{Var}(Y_i) = 2[\sigma_X^2 - R(t_s)]. \quad (44)$$

So clearly, for uniform sampling, the delay variation depends on the sampling interval t_s . It is therefore believed that to measure a delay-variation of relevance to a particular application, that the measurements must be made at the same intervals as the application's traffic.

For Poisson sampled measurements $X_i = X(T_i)$, and

$$E[R(T_i - T_{i-1})] = \lambda \int_0^{\infty} R(s)e^{-s\lambda} ds \quad (45)$$

resulting in different metrics for Poisson and uniform samples.

However, we can invert the LSP (with Poisson samples) to derive an ACF for the delay measurements, and from this, we could easily substitute values directly into (44) in order to obtain an estimate of the delay variation for an arbitrary sampling interval (whereas uniform sampling will provide this only for a fixed sampling interval and integer multiples thereof). Fig. 6 shows a comparison of the direct computation from uniform samples with estimates obtained from inverting the LSP for Poisson samples obtained above. We can see that both provide a reasonable

approximation to the true delay variation, and that the direct and LSP approaches are statistically indistinguishable.

V. OTHER ISSUES

There are many other issues that impact the choice of sampling used for performance measurements.

A. Sampling Intervals

A common concern with random probes is that there may be large gaps between some adjacent probes, and that we might miss a particular event, or that an opportunist might somehow exploit these gaps. Uniform sampling has a maximum time between probes t_s , and so provides guarantee of the maximum duration event that can escape notice. However, the lack of anticipation property that forms the basis of PASTA prevents anyone from exploiting the gaps between measurements, as they cannot tell in advance where these will be. Furthermore, it is worth considering that the probability of an interprobe time 100 times the mean time for a PSP is less than 10^{-14} . Thus we will not observe such large gaps between probes.

In the case where the mean interprobe times are already large, then we may wish to limit the possible gaps a little more stringently, and it may be appropriate to place a cap on the interprobe interval. ASTA will still apply as long as the lack of anticipation property holds. A probe sequence with truncated intervals will be more easily anticipated by an assailant, but is unlikely to result in synchronization, and so we may find this solution suitable in some cases. Furthermore, the LSP does not depend on Poisson sampling.

Similar issues arise when we consider the fact that probe packets have a minimum interprobe times either because of the probe generation mechanism, or the packet transmission time for the probe packets.

B. Component Performance

Another point of importance is the fact that a network is really made up of many queues. PASTA applies to a system as a whole, and so a single queueing model such as used here may work for modeling end-to-end measurements of delay through a series of queues. However, the sample sequences of probes at each link in the network will not be Poisson. They will be distorted by queueing with cross-traffic at previous links. The PASTA property will, therefore, not hold at individual components of an end-to-end path. Does this matter? Not if we only consider end-to-end measurements. However, by performing probes across a set of end-to-end paths, we can form an inverse problem whose solution gives the performance of the network at individual components (for example, see [39] and the papers therein). PASTA will not hold for these individual components, and so there is an apparent reliance on the more general ASTA results.

C. Intractable Measurements

Some measurements are intrinsically hard to make using Poisson probes. For instance, reordering metrics are critically dependent on the spacing between probe packets. Packets far apart in time are very unlikely to be reordered. We could model this via delays in multiple queues, and attempt to infer reordering from Poisson samples, however, the estimate would be highly dependent on how accurate a model was used. Hence,

for metrics such as this, uniform probing seems to make more sense. On the other hand, the same affect noted above—the distortion of probes away from their ideal probe stream at downstream links—also applies to uniform samples. Hence, if the reordering occurs late along a probe path, the samples will not be uniformly spaced, and the reordering metric will not measure the same thing as it would were the reordering happening early in the path. This seems to be a difficulty in the notion of defining a reordering metric, rather than the sampling method used.

VI. CONCLUSION

This paper has quantitatively compared differences between Poisson and uniform sampling for active performance measurements, and more generally presented techniques for understanding the properties of any sampling approach.

The key findings of this paper are the following.

- Poisson sampling results in a reduction in efficiency, and hence more Poisson probes would be required to provide estimates of the same statistical accuracy. However, the reduction in efficiency is dwarfed by the natural variation as a result of different system loads, and so this seems not to be the major consideration.
- Irregular probes can be used to find time-series properties of a system, such as its periodogram or ACF. Such techniques could be used to detect periodicities in the system, and warn of potential bias due to synchronization (if pure Poisson sampling is not used).

It is clear that these findings do not rule out either sampling approach. Rather a practitioner should choose the most appropriate measurements for their application, but this should be a choice informed by the above results.

REFERENCES

- [1] V. Paxson, G. Almes, J. Mahdavi, and M. Mathis, "Framework for IP performance metrics," IETF, RFC 2330, 1998.
- [2] J. Mahdavi and V. Paxson, "IPPM metrics for measuring connectivity," IETF, RFC 2678, 1999.
- [3] G. Almes, S. Kalidindi, and M. Zekauskas, "A one-way delay metric for IPPM," IETF, RFC 2679, 1999.
- [4] —, "A one-way packet loss metric for IPPM," IETF, RFC 2680, 1999.
- [5] —, "A round-trip delay metric for IPPM," IETF, RFC 2681, 1999.
- [6] S. Floyd and V. Jacobson, "On traffic phase effects in packet-switched gateways," *Internetworking: Research and Experience*, vol. 3, no. 3, pp. 115–156, 1992.
- [7] —, "The synchronization of periodic routing messages," *IEEE/ACM Trans. Netw.*, vol. 2, no. 2, pp. 122–136, 1994.
- [8] M. Roughan, "Fundamental bounds on the accuracy of network performance measurements," in *Proc. ACM SIGMETRICS*, Banff, Canada, 2005, pp. 253–264.
- [9] V. Raisanen, G. Grotefeld, and A. Morton, "Network performance measurement with periodic streams," IETF, RFC 3432, 2002.
- [10] "CAIDA performance measurement tools taxonomy." [Online]. Available: <http://www.caida.org/tools/taxonomy/performance.xml>
- [11] V. Paxson, "Measurements and analysis of end-to-end Internet dynamics," Ph.D. dissertation, Univ. California at Berkeley, Berkeley, NJ, 1997.
- [12] V. Paxson, J. Mahdavi, A. Adams, and M. Mathis, "An architecture for large-scale Internet measurement," *IEEE Commun. Mag.*, vol. 36, no. 8, pp. 48–54, Aug. 1998.
- [13] N. Brownlee, "Packet matching for NeTraMet distributions." Mar. 2000. [Online]. Available: <http://www2.auckland.ac.nz/net/Internet/rfmm/Meetings/47->
- [14] Y. Zhang, L. Breslau, V. Paxson, and S. Shenker, "On the characteristics and origins of Internet flow rates," in *Proc. ACM SIGCOMM*, Pittsburgh, PA, Aug. 2002, pp. 309–322.
- [15] P. Barford and J. Sommers, "Comparison of probe- and router-based methods for measuring packet loss," *IEEE Internet Comput. (Special Issue on Measuring the Internet)*, vol. 8, no. 5, pp. 50–56, Sep.–Oct. 2004. [Online]. Available: <http://www.cs.wisc.edu/~pbb/publications.html>
- [16] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker, "On the constancy of Internet path properties," in *Proc. ACM SIGCOMM Internet Measure. Workshop*, Nov. 2001, pp. 197–211.
- [17] H. Stark and J. W. Woods, *Probability and Random Processes With Applications to Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [18] B. Melamed and W. Whitt, "On arrivals that see time averages," *Oper. Res.*, vol. 38, no. 1, pp. 156–172, 1990.
- [19] B. Melamed and D. Yao, *Advances in Queueing: Theory, Methods and Open Problems*. Boca Raton, FL: CRC Press, 1995, ch. The ASTA Property, pp. 195–224.
- [20] R. Wolff, "Poisson arrivals see time averages," *Oper. Res.*, vol. 30, pp. 223–231, 1982.
- [21] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot, "The role of PASTA in network measurement," in *Proc. ACM SIGCOMM*, Pisa, Italy, Sep. 2006, pp. 231–242.
- [22] R. Koodli and R. Ravikanth, "One-way loss pattern sample metrics," IETF, RFC 3357, 2000.
- [23] C. Demichelis and P. Chimento, "IP packet delay variation metric for IP performance metrics (IPPM)," IETF, RFC 3393, 2002.
- [24] M. Tariq, A. Dhamdhere, C. Dovrolis, and M. Ammar, "Poisson versus periodic path probing (or, does PASTA matter?)," in *Proc. USENIX/ACM Internet Measure. Conf.*, Oct. 2005, pp. 119–124.
- [25] J. Beran, *Statistics for Long-Memory Processes*. London, U.K.: Chapman & Hall, 1994.
- [26] W. Whitt, "Planning queueing simulations," *Manage. Sci.*, vol. 35, no. 11, pp. 1341–1366, 1989.
- [27] G. S. Fishman, "Estimating sample size in computing simulation experiments," *Manage. Sci.*, vol. 18, pp. 21–38, Sep. 1971.
- [28] L. Kleinrock, *Queueing Systems, Vol. II: Computer Applications*. New York: Wiley, 1975.
- [29] P. M. Morse, "Stochastic properties of waiting lines," *J. Oper. Res. Soc. America*, vol. 3, no. 3, pp. 255–261, 1955.
- [30] G. S. Fishman and P. J. Kiviat, "The analysis of simulation-generated time series," *Manage. Sci.*, vol. 13, pp. 525–557, Mar. 1967.
- [31] J. Abate and W. Whitt, "The correlation functions of RBM and M/M/1," *Stochastic Models*, vol. 4, no. 2, pp. 315–359, 1988.
- [32] A. Basu and J. G. Riecke, "Stability issues in OSPF routing," in *Proc. ACM SIGCOMM*, Aug. 2001, pp. 225–236.
- [33] N. R. Lomb, "Least-squares frequency analysis of unequally spaced data," *Astrophys. Space Sci.*, vol. 39, pp. 447–462, 1976.
- [34] J. D. Scargle, "Studies in astronomical time series analysis II. Statistical aspects of spectral analysis of unevenly sampled data," *Astrophys. J.*, vol. 263, pp. 835–853, 1982.
- [35] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1988–1992.
- [36] W. H. Press and G. B. Rybicki, "Fast algorithm for spectral analysis of unevenly sampled data," *Astrophys. J.*, vol. 338, pp. 277–280, 1989.
- [37] J. Scargle, "Studies in astronomical time series analysis. III. Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly spaced data," *Astrophys. J.*, vol. 343, no. 2, pp. 874–887, 1989.
- [38] L. Eyer and P. Bartholdi, "Variable stars: Which Nyquist frequency," *Astron. Astrophys. Suppl. Ser.*, vol. 135, pp. 1–3, 1999.
- [39] N. Duffield, V. Paxson, and D. Towsley, "MINC: Multicast-based inference of network-internal characteristics." [Online]. Available: <http://www.research.att.com/projects/minc/>



Matthew Roughan (M'97) received the Ph.D. degree in applied mathematics from the University of Adelaide, Adelaide, Australia, in 1994.

He joined the School of Applied Mathematics, University of Adelaide in February 2004, where he is interested in the area of design, and installation of Internet measurement equipment, and the analysis and modeling of Internet measurement data. He previously worked in this capacity for four years for AT&T Research, and for Ericsson, Australia (via the Universities of Melbourne, and the Royal Melbourne Institute for Technology). He has also worked at the Cooperative Research Centre for Sensor Signal and Information Processing.