Parametric Resampling Analysis of Traffic Measurements for Capacity Management[*]

Mine Çağlar
Koc University, Turkey
mcaglar@ku.edu.tr

K.R. Krishnan
Telcordia Technologies, Inc., USA
krk@research.telcordia.com

Abstract

Our aim in this paper is to apply the technique of parametric 'bootstrapping' (or 'resampling') to determine what sampling frequency of traffic measurements is necessary for the proper engineering of high-speed data networks. Recent studies have shown that Fractional Brownian Motion (FBM) is a good model for the traffic observed in high-speed data networks, capturing both self-similarity and long-range dependence. We investigate the effect of the frequency of collection of traffic measurements on the estimation of FBM traffic parameters and, in turn, on the accuracy of link engineering. We go beyond earlier studies to obtain the statistical properties of link engineering that is based on traffic estimation from sampled measurements. The key idea is to 'resample' by creating replicate data sets with the same parametric model as the original data set, from which the variability of the quantities of interest can be assessed with reference to the time-resolution of the traffic measurements, which is our basic control variable. We find that link engineering based on traffic measurements sampled at one-minute intervals produces significant error, in the range 7-26%, while engineering based on measurements sampled at one-second intervals results in an error in the range 2-4%. Hence, we conclude that for acceptable accuracy in capacity engineering for IP traffic, traffic measurements should be collected at a finer time-resolution than one-minute intervals, and that measurements at one-second intervals lead to acceptable accuracy.

1. Introduction

Traffic measurements are collected in networks both to assess current network performance and to detect traffic trends and estimate the load for use in network engineering. On the basis of the available measurements, the engineering algorithms must ensure that the provisioned capacity is sufficient to meet the anticipated demand for broadband applications, but not so excessive as to render broadband services uneconomical. In this paper, we apply the technique of 'bootstrapping' (or 'resampling') to determine what sampling frequency of traffic measurements is necessary for the proper engineering of high-speed data networks.

Many analyses of fine-grained measurements over the last decade have established the self-similar and long-range dependent nature of traffic in high-speed data networks (see e.g. [1], [11], and [12]). More recent studies (e.g. [8]) also examine the multifractal character of data traffic, where applicable. These studies show that most high-speed traffic is generated by services that use the Internet Protocol (IP), and that the percentage of World-Wide-Web traffic has grown to around 30% of network traffic. In this study, we focus on traffic aggregated from many sources, which can be modeled parsimoniously by the FBM traffic model (which is 'fractal' but not 'multifractal', reflecting the effect of a high level of aggregation). The FBM model captures self-similarity and long-range dependence in a single parameter *H*, known as the Hurst parameter, which can be estimated from the actual traffic.

In this paper, we go beyond the results in [3] in investigating how the estimation of FBM parameters of traffic (by the variance-time and wavelet methods) is affected by the frequency of collection of traffic measurements, and how this effect translates into an effect on link dimensioning. In [3], the estimates of the FBM parameters were used as 'point estimates' for the purpose of link-dimensioning calculations, with no attempt to evaluate the *statistical* properties of the dimensioning results. Even for the FBM parameters themselves, there was only limited statistical information available from theoretical analysis. While the properties of the estimators

---

of H produced by the two estimation methods are well known, the *joint* estimation of H and the peakedness parameter of the traffic model is relatively new.

In this paper, we employ statistical resampling, or bootstrap, with realistic input parameters for the FBM traffic model, in order to arrive at more conclusive results than were obtained in [3]. The key idea is to 'resample' by creating replicate data sets with the same parametric model as the original data set, from which the variability of the quantities of interest can be assessed with reference to our basic control variable, the time-resolution of traffic measurements, for both the estimation methods. For the FBM traffic parameters (a, H) themselves, Eric van den Berg [14] has carried out an analysis of their statistics, as derived from replicated data sets, for estimation by the variance-time method. What is accomplished in this paper is the derivation, by parametric resampling, of the properties of the estimator $\hat{N}$ (for both the variance-time and the wavelet estimation methods) of the number of sources that can be supported on a given link, which is a complicated function of the FBM traffic parameters. We find that link engineering based on traffic measurements sampled at one-minute intervals produces significant error, in the range 11-26% for the variance-time method and 7-8% for the wavelet method. With measurements at one-second intervals, the engineering error lies in the range 2-4% for the variance-time method and 2-3% for the wavelet method. Hence, we conclude that for acceptable accuracy in capacity engineering for IP traffic, traffic measurements should be collected at a finer time-resolution than one-minute intervals, and that measurements at one-second intervals lead to acceptable accuracy.

The outline of the paper is as follows. In Section 2, we discuss traffic measurements, and describe the two data sets that we have used in this study. In Section 3, the FBM traffic model is introduced, and the bootstrap procedure for this parametric model is then described in detail. The variance-time and wavelet methods of estimation of FBM parameters are briefly reviewed, along with the link-dimensioning algorithm, in Section 4. The resampling results are reported and analyzed in Section 5, and our conclusions are stated in Section 6.

2. Traffic Measurements

The traffic measurements collected by most switches are in the form of *traffic counts* reported at regular intervals of coarse time-resolution, typically 15 minutes. More frequent measurements are obtained only occasionally for 'special studies', since the regular collection of traffic measurements at intervals smaller than 15 minutes places large demands on switch processing, when the collection relies on the mechanism of network management queries,

as in SNMP (Simple Network Management Protocol). On the other hand, it may be reasonable to invoke the mechanism, say, once a week for a chosen hour (over which data traffic may be assumed stationary) that is judged to be critical for capacity management. Alternative measurements are discussed in [9]. Our aim in this study is to determine the influence of the time-resolution of the traffic measurements on the accuracy of capacity management that is based on those measurements. Previous results [3] indicate that a 15-minute time interval is not adequate for this purpose in the case of high-speed data traffic approximated by the FBM model. Therefore, in this paper, we focus on a comparison between measurements collected at 1-minute intervals and those collected at 1-second intervals.

It is well established now [1] [11] [12] that the traffic in high-speed data networks is self-similar and long-range dependent in nature − i.e., the traffic "looks" the same when measured over a wide range of time scales, and correlation remains significant across large time scales, respectively. As for the protocol, the traffic is mostly generated by services that use the Internet Protocol (IP). This applies, for example, to the early Bellcore traffic measurements reported by [11].

In this paper, we consider two sets of traffic measurements as our data sets. In each data set, the measurements are the traffic counts in bytes collected at regular intervals at the relevant time-resolution of observation. Data Set 1 consists of 3600 measurements on a Frame Relay link, at 1-second intervals, for a total duration of one hour. Data Set 2 consists of 24000 measurements on an ethernet, at 40-millisecond intervals, for a total duration of 16 minutes. We use these data sets to estimate the parameters of our traffic model, which is Fractional Brownian Motion (FBM), as described below. Data Sets 1 and 2 were studied in [3], and their Hurst parameters were found to be 0.67 and 0.81, respectively, as estimated from measurements at one-second intervals.

3. Traffic Model and Resampling

3.1. FBM Traffic Model

The FBM model for data traffic is completely specified by three parameters (m, a, H) where

m : mean traffic rate, say, in bytes/time-unit

a : peakedness (the ratio of variance of traffic arrivals in time-unit to mean traffic rate stated for the same time-unit)

H : Hurst parameter (a dimensionless measure of persistence of correlations in traffic, with $\frac{1}{2} \le H < 1$).

The total traffic arriving in the interval $(0, t]$ is given by

$$A(t) = mt + \sqrt{ma} X(t)$$

Table 1: Parameters (m, a, H) for data sets

Parameters	m (bytes/sec)	a (byte-sec)	H
Data Set 1	800	505	0.67
Data Set 2	3,188	12.7	0.81

where $X(t)$ is a standard FBM with Hurst parameter H [13]. The incremental traffic over the interval $(t,t+1)$ is given by

$$A(t+1) - A(t), \quad t = 1,2,3,\ldots$$

This is a parsimonious model, which captures the self-similarity and long range dependence observed in aggregate data traffic [11] [13]. The burstiness over all time scales is represented by the single parameter H.

In this paper, we focus on data traffic aggregated from many independent sources. For this case, FBM is a good traffic model. As described in Section 2, we have two data sets, which represent widely different values for the parameters (m, a, H). Methods for estimating these parameters, described in [3], are summarized in Section 4. The estimated values of these parameters are given in Table 1.

3.2   Resampling Traffic Measurements

Traffic measurements are fundamental for the proper engineering of high-speed data networks. Our main focus is the *frequency* of these measurements for that purpose. Since very few data sets are available from real networks for such an extensive study, in this paper, we exploit *resampling* ideas from statistics [4] to fill the gap. Also, the frequency of measurements in existing switching systems is rarely at the one-second interval that was considered in [3]; rather, the usual frequency is 15 minutes. In this situation, one has to obtain measurements collected at fine time-resolutions as part of special studies, and these are too few to be conclusive. With resampling, once the basic traffic model is accepted as a reasonable representation for the traffic, it is possible to generate as many samples as are needed.

The key idea is to 'resample' from replicate data sets created on the same parametric model as the original data, from which the variability of the quantities of interest can be assessed. Such 'resampling' is especially useful when the variability is not known theoretically and hence cannot be computed or estimated from a single data set. This approach is also referred to as *bootstrap method* [5], which term is used interchangeably with resampling below. Historically, Efron [6] introduced the term "bootstrap" as a nonparametric procedure when he made the connection of Monte Carlo methods to standard methods of parametric inference, and drew the attention of statisticians to their potential use for nonparametric inference. This work and subsequent developments made connection with earlier resampling ideas and formed the so-called *nonparametric bootstrap*. Currently, bootstrap procedures refer *both to parametric and nonparametric* resampling, where the common idea is to replicate a given data set (see e.g. [4] for a comprehensive treatment of the subject).

Our main purpose is to evaluate the effect of the frequency of measurements on the estimated number of independent and homogeneous sources N, each described by the FBM model (m, a, H), that can be carried on a T1 link. Certainly, this can be observed on the real Data Sets 1 and 2 that we have described in Section 2, by using their respective parameters (m, a, H), but *with no variance information* on the point estimate of the number of sources. Indeed, this was done in [3]. A more reliable comparison is to involve in the analysis the mean and variance of the number of sources estimated for each sampling frequency; and these, being the properties of the estimator $\hat{N}$ of the number of sources, can only be found from bootstrap, since the theoretical mean and variance are not known. The same is true for the confidence intervals. Although, theoretically, confidence intervals are available for the triplet (m, a, H) from a single data set, these intervals do not translate into a corresponding interval for N due to the complicated relationship between N and (m, a, H).

The Bootstrap Procedure

In this paper, we use *parametric resampling* as our bootstrap procedure [4] where a parametric model (in this paper, the FBM model) is assumed. A computer simulation of FBM traffic streams with input parameters obtained from the available data set is the main idea. A nonparametric bootstrap, on the other hand, would be based on no assumptions on the model, and would work through 'resampling' measurements (with replacement) from the available data set. We choose the parametric bootstrap because of the good fit and desirable properties of the FBM model for aggregate data traffic.

We typically produce R replicates of each data set from a fast and accurate FBM traffic generator [2]. The physical relationship of a data set with its replicates is the common set of FBM parameters that characterize both the

original data set and the derived replicates; these parameters are inputs for the generation algorithm. Hence, no model validation needs to be done, as the generated samples are indeed FBM streams (up to numerical approximation, of course) for a given triplet (m, a, H). As a result of randomness, R different estimators $(\hat{m}_r, \hat{a}_r, \hat{H}_r), r = 1, 2, \ldots, R$ are obtained from this simulation. The two estimation methods used, namely variance-time and wavelet methods, are described in the next section. Using $(\hat{m}_r, \hat{a}_r, \hat{H}_r), r = 1, 2, \ldots, R$, we compute the estimated number of sources $\hat{N}_1, \hat{N}_2, \ldots, \hat{N}_R$. This set of R replicates is, in turn, used to estimate the properties of the estimator $\hat{N}$, such as its mean, bias, and standard deviation. Also, a confidence interval on N can be obtained by using the percentiles of the set $\hat{N}_1, \hat{N}_2, \ldots, \hat{N}_R$. For each original data stream, this procedure of generating R replicate streams and constructing R estimators is carried out twice, once using the parameter estimates (m, a, H) derived by sampling the original data stream at one-second granularity, and a second time using the parameter estimates derived by sampling the same stream at one-minute granularity.

Nature of Results Available from Bootstrap Procedure

In Section 5, we compare, for example, the difference of the *means* of estimated number of sources, and find out the size of the dimensioning error that is incurred by going from fine-grained measurements to coarser-grained ones (which was our primary goal). We test the statistical hypothesis that there is a significant estimation error, which translates into error in dimensioning, and hence, in capacity management, by reducing the frequency of traffic measurements from one-second to one-minute. As by-products of the whole bootstrap procedure, we also obtain the properties of the estimators $\hat{a}$ and $\hat{H}$, in addition to those of $\hat{N}$, for a given estimation method. The properties that are already theoretically available are confirmed, and the unavailable ones are obtained empirically. Such results also offer a good basis for comparison of the two estimation methods.

Resampling, or bootstrap, is especially useful in applications such as the one considered in this paper. An alternative could be to perform Monte Carlo simulations for a range of values of (m, a, H) and make the same comparisons. However, the range of values would be rather arbitrary and might not cover some values that are estimated from new real traffic streams. In contrast, bootstrap is portable to any future applications: for each new triplet $(\hat{m}, \hat{a}, \hat{H})$, one can repeat this procedure inexpensively using computing power, and obtain results.

Since the parameters (m, a, H) for the two data sets that we use are quite different, our results appear to have general validity.

4. Estimation of FBM Parameters and Dimensioning

4.1. Estimation of FBM Parameters

In this paper, we make use of the variance-time method and the wavelet method for estimating the parameters of the FBM model from traffic measurements. The methods are described in [3], and are summarized here for reference.

Consider the traffic counts $\{x_k, k = 1, \ldots, N\}$, where x_k is the number of traffic units arriving in the k'th measurement interval, and N is the total number of samples. In both methods, the estimate of the mean arrival rate is given by

$$\hat{m} = \frac{\sum_{k=1}^{N} x_k}{N}$$

a) Estimate of a and H in Variance-Time Method

We construct the variance-time plot by considering the sequence formed by aggregating the arrivals in n *consecutive* intervals, for various values of *n*. Thus, the *n-fold aggregated sequence* $\{y_j^{(n)}\}$ is defined by

$$y_j^{(n)} = \sum_{i=(j-1)n+1}^{jn} x_i \quad , \quad j = 1, \cdots, M_n,$$

where $M_n = \left\lfloor \dfrac{N}{n} \right\rfloor, \quad n = 1, 2, \cdots N$. For an FBM source (m, a, H), the plot of $\log V(n)$ versus $\log n$, where $V(n)$ is the *true* variance of the *n-fold* aggregated process, is a straight line with slope $S = 2H$, with an ordinate-axis intercept of $\log(ma)$. We calculate the *sample* variance $\hat{V}(n)$ of the *n-fold aggregated* sequence (for values of n for which the sequence has at least four samples), and use linear regression on the set of points $\{\log \hat{V}(n), \log n\}$ to estimate both the slope S and the ordinate-axis intercept I of the curve. We then estimate H and the peakedness parameter a as follows:

$$\hat{H} = \frac{S}{2}; \qquad \hat{a} = \frac{e^I}{\hat{m}}$$

b) Estimate of *a* and *H* in Wavelet Method

We consider the joint estimation of *a* and *H* by the wavelet estimation method given in [15], using Daubechies wavelets with two vanishing moments. Let $d(j,k)$, $k = 1, \ldots, n_j$, $j = 1, \ldots, J$ denote the 'details' obtained by the discrete wavelet transform of the sequence of traffic counts $\{x_k, k = 1, \ldots, N\}$, where *J* is such that $2^{J+1} \leq N \leq 2^{J+2}$, and n_j is the number of coefficients available at octave *j*. The statistic central to the method is given by

$$\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} d^2(j,k), \;\; j = 1, \ldots, J .$$

Then, the Hurst parameter *H* and the coefficient c_f of the spectrum of the data stream are estimated through a weighted linear regression of

$$y_j = \log_2(\mu_j) - g_j$$

over $j = j_1, \ldots, j_2$, where j_1 and j_2 are the scales relevant for long-range dependence. The constant

$$g_j = E(\log_2(\mu_j)) - j(2H-1) - \log_2(c_f C)$$

is introduced to ensure that the fundamental hypothesis of regression holds (with *C* a constant that depends on *H*). Then, the slope α of the regression line is $(2H-1)$ and its intercept β is $\log_2 B$, where $B = c_f C$. Veitch and Abry [15] devise unbiased estimators for $(2H-1)$ and $c_f C$ under mild assumptions, which effectively hold for FBM. The estimate of *H* is given by

$$\hat{H} = \frac{\alpha + 1}{2}$$

which is unbiased and consistent. The estimator of $B = c_f C$ is given by $\hat{B} = p \cdot 2^{\beta}$, where *p* is a constant to ensure unbiasedness of $\hat{B}$. We now set $\hat{c}_f = \dfrac{\hat{B}}{\hat{C}}$, where $\hat{C}$ is computed using $\hat{H}$. Using the relationship of c_f to the variance of the traffic counts for larger scales, we obtain the estimate of the peakedness parameter *a* by

$$\hat{a} = \frac{\hat{c}_f (2\pi)^{2\hat{H}-1}}{2\hat{m}\Gamma(2\hat{H}-1)\cos((2\hat{H}-1)\pi/2)(2\hat{H}-1)\hat{H}}$$

where $\hat{m}$ is the estimate of the mean of the traffic counts. This estimate is a complicated function of $\hat{c}_f$ and $\hat{H}$, and its behavior is not known.

4.2 Dimensioning Calculations for FBM Traffic Model

The problem is to determine the number of independent and identical FBM sources, each defined by the parameters (m, a, H), that can be supported on a link with capacity *L* and buffer-size *B*, for a specified bound ε on the loss-rate, and a specified bound *D* on the mean delay. The multiplexing of *n* independent, homogeneous FBM sources, each described by the parameters (m, a, H), produces an FBM source with the parameters (nm, a, H) [13]. Owing to the multiplexing gain inherent in the statistical multiplexing of independent sources, the bandwidth requirement is a non-linear function of the number of sources. Using the scaling results presented in [10] and in [7], one can determine the bandwidth required to carry the traffic of *n* independent sources while meeting the *more stringent* of the above two requirements on the loss-rate and the mean delay. This calculation helps determine the number of sources that can be supported on the given link with capacity *L* and buffer-size *B*.

For the calculations in this paper, we assume a loss rate of 0.0001 and a mean delay of 10 msec as the service criteria, with a 500-byte buffer. We assumed a peak source-rate equal to twice the mean rate. Although there is no finite 'peak rate' in the FBM traffic model, the actual traffic sources that we approximate by the FBM model do have finite peak rates. The engineering algorithm that is used in [10] does take account of the fact that the effective bandwidth for a source cannot exceed its peak rate.

5. Bootstrap Results

The results in this section are presented in two parts. In Section 5.1, we compare the two methods used for the estimation of the FBM traffic parameters, by looking at the properties of their estimators for *a* and *H*, both for 1-second and 1-minute time-resolutions of measurement. In Section 5.2, which is our main result, we turn to the variable of interest in capacity engineering, viz., the number of sources that a link can support, and consider the error that is incurred as a result of less frequent measurements, for either method of estimation of FBM parameters.

Table 2: Estimators and their properties for one-second traffic stream

Actual values of (a,H)		(505,0.67) Data Set 1		(12.7,0.81) Data Set 2	
Parameter		a	H	a	H
V-T Method	Estimator	2094	0.64	63.4	0.75
	Bias	1589	-0.03	50.7	-0.06
	Standard Deviation	3590	0.06	79.1	0.05
Wavelet Method	Estimator	620	0.68	18.6	0.81
	Bias	115	0.01	5.9	-0.005
	Standard Deviation	529	0.04	12.7	0.04

Table 3: Estimators and their properties for one-minute traffic stream

Actual values of (a,H)		(505,0.67) Data Set 1		(12.7,0.81) Data Set 2	
Parameter		a	H	a	H
V-T Method	Estimator	2357	0.57	176	0.65
	Bias	1852	-0.10	163	-0.16
	Standard Deviation	40114	0.13	2254	0.12
Wavelet Method	Estimator	1.5E+08	0.61	3.2E+07	0.69
	Bias	1.5E+08	-0.06	3.2E+07	-0.12
	Standard Deviation	3.9E+09	0.25	6E+08	0.23

5.1. Estimation of FBM parameters

For each set of parameters (m, a, H) obtained from the two data sets described earlier, we generate $R = 1000$ replicates of an FBM traffic stream defined by those parameters, using the method described in [2]. Each stream is an hour-long data-sequence with samples at one-second intervals, i.e., there are 3600 sample points. Then, the parameter estimates $(\hat{m}_r, \hat{a}_r, \hat{H}_r)$, $r = 1,2,\ldots,1000$ are obtained by both the variance-time and wavelet estimation methods. The relative performances of these two methods are well known for estimation of H. For example, the estimator $\hat{H}$ is unbiased for the wavelet method, whereas it is biased for V-T method. For $\hat{a}$, on the other hand, our empirical results are especially useful. They give information about the bias as well as about the standard deviation, which may, in turn, be used for constructing confidence intervals for a. Table 2 summarizes the estimation results for the fine-resolution data sets, i.e., the traffic streams generated at one-second intervals. The estimators are found as the mean of $(\hat{a}_r, \hat{H}_r)$, $r = 1,2,\ldots,1000$, for each method.

The bias is given by the difference

$$\frac{\sum \hat{H}_r}{1000} - H$$

for H, and by an analogous difference for a. The standard deviation is the sample standard deviation of $\hat{H}_1,\ldots,\hat{H}_{1000}$ for H and is obtained similarly for a. It can be seen that the wavelet estimator of H is essentially unbiased with a small standard deviation as suggested by theory. In comparison, the V-T estimator of H has a slight bias, with a somewhat larger standard deviation. The estimator of a is biased with both methods, but less

Table 4: Estimators of number of sources and their properties

Parameters (*m,a,H*)		(800,505,0.67) Data Set 1		(3188,12.7,0.81) Data Set 2	
Number of Sources N		190		53	
Sampling Interval		1-sec	1-min	1-sec	1-min
V-T Method	$\hat{N}$	195	240	51	59
	95% CI	[145,286]	[167,371]	[42,64]	[48,76]
	Bias	5 (2.6%)	50 (26%)	-2 (-3.8%)	6 (11.3%)
Wavelet Method	$\hat{N}$	195	204	52	49
	95% CI	[145,286]	[145,308]	[43,65]	[34,66]
	Bias	5 (2.6%)	14 (7.4%)	-1 (-1.9%)	-4 (-7.5%)

so with the wavelet method. Also, its standard deviation is lower with the wavelet method. This is basically because the wavelet estimator is a function of $\hat{c}_f$ defined in Section 4.1, which has good properties as an estimator.

Table 3 summarizes the estimators and their properties obtained from the 1000 data streams when aggregated at one-minute intervals. As a result, each stream now has only 60 sample points. In this case, the wavelet estimator is less biased but has larger standard deviation than the V-T estimator for *H*. When it comes to estimation of *a*, the V-T estimator performs much better, as the wavelet estimators are too big to make any sense. This is basically because, with 60 sample points there are *very few octaves* that can be used for the regression analysis in the wavelet estimation method. Indeed, for some values of *r*, we were not able to get an estimator with 0<*H*<1, for any possible choices of the octaves for regression. We had to delete these samples from our analysis.

We note also that the bias and standard deviation increase for all estimators when the traffic stream is sampled at the coarser intervals of one-minute, as expected. For long enough streams of data traffic, the wavelet method performs better in obtaining $(\hat{a}, \hat{H})$, but with short streams, the V-T method performs better ultimately, in that, at least, it produces meaningful estimators. Besides, it may not even be possible to obtain a wavelet estimator in some cases, as explained above, with sampling at one-minute intervals in an hour.

5.2. Estimation of Number of Sources

Next, we estimate the number of sources N that can be supported on a T1 link (1544 kbits/sec), by calculating $\hat{N}_r$

from the parameter set $(\hat{m}_r, \hat{a}_r, \hat{H}_r)$, for $r = 1,2,\ldots,1000$, by both estimation methods. The results are given in Table 4.

The bias (error in capacity engineering) increases in all cases when the sampling interval is increased from seconds to minutes, taking values in the 7-26% range in the latter case; this size of engineering error can have significant cost penalties. When the two methods of estimation are compared, one cannot really distinguish them in terms of bias and the width of the confidence intervals for N for the fine analysis, i.e., one-second intervals. In particular, for the parameters of Data Set 1, both the wavelet and V-T methods produce exactly the same sequence $\hat{N}_1, \hat{N}_2, \ldots, \hat{N}_{1000}$ for the fine analysis. In the one-minute interval case, wavelet estimation performs slightly better. This is notable, as the wavelet estimators of the model parameters are, in fact, very poor with the coarse samples. Besides, for some replicates *r* (about 5-10% of the cases), we were not able to estimate (a, H) and, hence *N*, due to the small number of scales available. As far as the analysis of the fine-resolution one-second data is concerned, the estimation of the number of sources is robust to the method of estimation.

We now quantify the *statistically significant* differences in one-minute and one-second analysis. We employ the two-sample test for paired data [5], as the 1-sec. and 1-min. calculations occur in pairs for each data stream. The test statistic for the mean difference between 1-sec. and 1-min. sources is given by $|Z|$, where

$$Z = \frac{\overline{D}}{s / \sqrt{n}}$$

Table 5: Difference in $\hat{N}$ between 1-sec. and 1-min. measurements

Parameters (*m,a,H*)		(800,505,0.67) Data Set 1	(3188,12.7,0.81) Data Set 2
Number of Sources N		190	53
V-T Method	Observed difference	45	8
	Significant difference δ	43 (22.6%)	8 (15.1%)
	P-value	0.0007	< 0.0002
Wavelet Method	Observed difference	9	-3
	Significant difference δ	7 (3.7%)	-2 (3.8%)
	P-value	< 0.0002	0.0004

and $\overline{D}$ is the mean of the differences $\hat{N}_r(1\min) - \hat{N}_r(1\sec)$, $r = 1, \ldots, n$, and s is the standard deviation of these differences. Note that $n = R = 1000$ for the variance-time method, and is less than 1000 (about 5-10% less) for the wavelet method because this method cannot produce estimates for some of the replicates in the one-minute case due to too few octaves being available. Since the traffic streams are independent from one another, and n is large, Z is approximately a standard normal random variable. The statistical test results are listed in Table 5, where the P-values are reported along with the *greatest significant difference* between fine and coarse measurements. This is the largest positive difference when $\overline{D} > 0$ and it is the smallest negative difference when $\overline{D} < 0$. Explicitly, we test the null hypothesis that the fine and coarse measurements produce a mean difference δ, against the alternative hypothesis that the difference is greater (smaller) than δ when $\overline{D} > 0$ ($\overline{D} < 0$). In such a test, the P-value, being the probability of the test statistic exceeding the observed value, can be considered as the strength of the *alternative* hypothesis: the *smaller the P-value*, *the stronger* the rejection of the null hypothesis in favor of *the alternative*. In Table 5, we report those values of δ that produce a P-value less than or equal to 0.01, which is small enough to be considered significant in statistical applications. We can see that the δ values reported in Table 5 are very close to observed values, which are rounded up to the nearest integer and reported here for reference. For example, for V-T method and Data Set 1, the observed difference is (240-195) = 45 in our experiments; equivalently, it is the *difference in the biases*: (50-5) = 45. After the statistical test, we say that 1-min. measurements will significantly *overestimate* the number of sources by at least 43 units compared to 1-sec. measurements.

The bias resulting from coarse measurements alone and the corresponding engineering impact have been discussed above. We now use Table 5 to compare the fine and coarse measurements relative to one another. The wavelet method again performs slightly better in respect of the increase in bias of one-minute samples when compared to one-second samples. However, this method may not be useful in short traffic traces, in which case, the variance-time estimation should be used.

6. Conclusion

We have applied the 'parametric bootstrap' approach to carry out a thorough analysis of the effect of sampling interval of traffic measurements in high-speed data networks on capacity engineering, going beyond the results of earlier analysis [3]. The fitted model is a parametric one, the so-called FBM traffic model. We find that the error made in link engineering is statistically significant (in the range 7-26%) when the measurements are taken at one-minute, instead of one-second, intervals. This range of error could produce unacceptable cost penalties. This qualitative result is robust in terms of the two well-known methods of estimation used: the variance-time and the wavelet methods.

This study also confirms that wavelet estimation is more accurate in determining the Hurst and peakedness parameters in FBM traffic model when the data stream is sufficiently long to provide a significant number of octaves for regression analysis. However, when the data stream has few sample points, as with coarse measurements over a limited duration, the variance-time method is the more reliable method of estimation of the FBM parameters. This fact is important if the value of the performance measure that is being estimated as a function of those parameters can change abruptly when those parameters are perturbed.

The results in this paper suggest that, for capacity engineering of acceptable accuracy for IP traffic, traffic measurements should be collected at a finer time-resolution than one-minute intervals. Our results also indicate that measurements at one-second intervals lead to acceptable accuracy.

References

[1] J. Beran, R. Sherman, M. Taqqu, and W. Willinger, Long-Range Dependence in Variable-Bit-Rate Video Traffic, *IEEE Transactions on Communications*, 43, pp. 1566-1579, 1995.

[2] M. Caglar, Simulation of Fractional Brownian Motion with Micropulses, *Advances in Performance Analysis*, Vol. 3, 2000.

[3] M. Caglar, K. R. Krishnan, and I. Saniee, Estimation of Traffic Parameters in High-Speed Data Networks, 16th International Teletraffic Congress, Edinburgh, U.K., pp. 867-876, June 1999.

[4] A. C. Davison and Hinkley, Bootstrap Methods and their Application, Cambridge University Press, 1997.

[5] J.L. Devore, Probability and Statistics for Engineering and the Sciences, Fourth Edition, Wadsworth, Inc., 1995.

[6] B. Efron, Bootstrap Methods: another look at the jacknife, Annals of Statistics 7, 1-26, 1979.

[7] A. Erramilli and J. L. Wang, A connection admission control algorithm for self-similar traffic, Globecom'98, Sydney, November 1998.

[8] A. Feldmann, A.C. Gilbert, W. Willinger and T. G. Kurtz, The changing nature of network traffic: Scaling phenomena, ACM SIGCOMM Computer Communication Review, April 1998.

[9] J.L. Jerkins, A.L. Neidhardt, J.L. Wang, A. Erramilli, Operations measurements for engineering support of high-speed networks with self-similar traffic, Proc. 16th International Teletraffic Congress, Edinburgh, UK, pp. 895-906, 1999.

[10] K. R. Krishnan, A. L. Neidhardt, and A. Erramilli, Scaling Analysis in Traffic Management of Self-Similar Processes, Proc. 15th International Teletraffic Congress, Washington, D.C., pp. 1087-1096, 1997.

[11] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, On the Self-Similar Nature of Ethernet Traffic (Extended Version), IEEE/ACM Trans. On Networking, 2, pp. 1-15, 1994.

[12] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, Self-Similarity in High-Speed Packet Traffic: Analysis and Modeling of Ethernet Traffic Measurements, Statistical Science, 10, pp. 67-85, 1995.

[13] I. Norros, A Storage Model with Self-Similar Input, Queueing Systems, 16, pp. 387-396, 1994.

[14] Eric van den Berg, Parameter Estimation For Long Range Dependent Data: Single Fine vs. Multiple Coarse Measurements, ITC Specialist Seminar on IP Traffic Modeling, Measurements, and Management, September 2000.

[15] D. Veitch and P. Abry, A Wavelet Based Joint Estimator of the Parameters of Long-Range Dependence, submitted to special issue of IEEE Trans. Information Theory, on Multiscale Statistical Signal Analysis and its Applications, 45, pp. 878-897, April 1999.