

Empirical Testing of the Infinite Source Poisson Data Traffic Model

Charles-Antoine Guérin, H. Nyberg, O. Perrin, S. Resnick, H. Rootzén, C. Stărică

▶ **To cite this version:**

Charles-Antoine Guérin, H. Nyberg, O. Perrin, S. Resnick, H. Rootzén, et al.. Empirical Testing of the Infinite Source Poisson Data Traffic Model. Stochastic Models, 2003, 19 (2), pp.151-200. 10.1081/STM-120020386 . hal-00083219

HAL Id: hal-00083219
https://hal.science/hal-00083219

Submitted on 10 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Empirical Testing of the Infinite Source Poisson Data Traffic Model[#]

C. A. Guerin,[1,*] H. Nyberg,[2] O. Perrin,[3] S. Resnick,[4] H. Rootzén,[5] and C. Stărică[5]

[1]Institut Fresnel, Faculte des Sciences de Saint-Jerome, Marseille cedex, France
[2]Ericsson AB, Ericsson Research, Wireless Access Networks, Kista, Sweden
[3]GREMAQ-University Toulouse I, Manufacture des Tabacs, Alle de Brienne, Toulouse, France
[4]School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, USA
[5]Department of Mathematics, Chalmers University of Technology, Gothenburg, Sweden

ABSTRACT

The infinite source Poisson model is a fluid queue approximation of network data transmission that assumes that sources begin constant rate transmissions of data at Poisson time points for random lengths of time. This model has been a popular one as analysts attempt to provide explanations for observed features in telecommunications data such as self-similarity, long range dependence and heavy tails. We survey some features of this model in cases where transmission length distributions have (a) tails so heavy that means are infinite, (b) heavy tails with finite mean and infinite variance and (c) finite variance. We survey the self-similarity properties of various descriptor

[#]Research supported by the Gothenburg Stochastic Centre, by the EU TMR network ERB-FMRX-CT96-0095 on "Computational and statistical methods for the analysis of spatial data" and by the Knut and Aline Wallenburg Foundation. Sidney Resnick's research was also partially supported by NSF grant DMS-97-04982 and NSA Grant MDA904-98-1-0041 at Cornell University.

[*]Correspondence: C. A. Guerin, Institut Fresnel, Case 162, Faculte des Sciences de Saint-Jerome, F-13397 Marseille Cedex 20, France; E-mail: charles-antoine.guerin@fresnel.fr.

processes in this model and then present analyses of four data sets which show that certain features of the model are consistent with the data while others are contradicted. The data sets are 1) the Boston University 1995 study of web sessions, 2) the UC Berkeley home IP HTTP data collected in November 1996, 3) traces collected in end of 1997 at a Customer Service Switch in Munich, and 4) detailed data from a corporate Ericsson WWW server from October 1998.

Key Words: Data transmission modelling; Internet traffic; Heavy tails; Regular variation; Pareto tails; Self-similarity; Scaling.

1991 Mathematics Subject: Primary 62N09; Secondary 60K250.

1. INTRODUCTION

Statistical simulation is of basic importance for the choice of buffer sizes, protocols, network configurations and other aspects of the design of complex telecommunication systems. Simulations, and network analysis in general, must be founded on models which capture important features of the traffic in a realistic and flexible way. Yet the models have to be simple enough to allow for understanding, theoretical analysis, and easy fitting to many kinds of observed and synthetic data. A simple model, here termed "the infinite source Poisson model" and sometimes called the M/G/∞ input model, which has the potential to satisfy these requirements for IP; HTTP, FTP, SMTP and other protocols for file transfers is surveyed and tested in this paper.

Our aim is to explore the statistical properties and limitations of this model, so its potential usefulness can be fully exploited. We examine its fit to a number of traffic measurements in order to understand which aspects of the model agree with reality, where the model is robust to deviations, and in what respects it may require extension and redefinition. A core issue is the relation between the micro level infinite source Poisson model and limiting aggregated models outlined below.

Data sets similar to the ones analyzed in this paper are rapidly being accumulated by the digital networking industry and by academic researchers. A further aim is to survey some statistical methods which we have found helpful, and which may be of use to the engineers and scientists who are coping with these data sets.

Erramili and Willinger[27] applied experimental queueing analysis to study how classical models without long range dependence can seriously underestimate delays. Random permutations of the arrival process were used by Andersen and Nielsen[8] to give further understanding these questions and of the interplay between correlations and queuing behavior.

The background is the identification of self similarity in various types of teletraffic flow rates at resolutions above a certain critical threshhold. This has created widespread interest in the possible origins and effects of the self similarity. Willinger et. al.[58 60,79,85 87] discussed self similarity of packet counts per unit time in LANS and WANS and a parallel discussion of self similarity of bytes per unit time in WWW traffic was conducted by Crovella et al.[17 19,21] Crovella, Kim and Park[20] conducted a large simulation study to assess the causes and effects of self similarity in situations that

involved slowdown nodes, buffers, varying rates and varying tail parameters. Erramili and Willinger[27] applied experimental queueing analysis to study how classical models without longrange dependence can seriously underestimate delays. Random permutations of the arrival process were used by Andersen and Nielsen[8] to further the understanding of these questions and of the interplay between correlations and queing behavior. Resnick and Samorodnitsky[69] constructed an example of a single exponential server fed by a long range dependent input which had queue lengths and waiting times which were heavy tailed. Mathematical studies of the connection between on off inputs with heavy tailed on periods appeared in Refs.[39,47 50,63,79]. The infinite source Poisson model was studied in Refs.[63,72].

Attempts to explain observed self similarity in network traffic have largely focused on heavy tailed transmission times of sources sending data to one or more servers. The common assumptions is that transmission times have iid random lengths with common distribution F. Often F has heavy tails in which case it is assumed F has a Pareto, or more generally, regularly varying tail so that

$$\bar{F}(x) := 1 - F(x) \sim x^{-\alpha} L(x), \quad x \to \infty, \tag{1.1}$$

where $L(x)$ a slowly varying function, so that

$$\lim_{t \to \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\alpha}, \quad x > 0, \tag{1.2}$$

where $\bar{F} = 1 - F(x)$. We distinguish three cases within F may satisfy.

(i) F has such a heavy tail that the mean is infinite and $0 < \alpha < 1$. Such heavy tails appear in the BU study of files sizes for the month of November (see the plots in Ref.[72]) and are reported by a Calgary study[7] of file sizes found on various servers.

(ii) F has a heavy tail with $1 < \alpha < 2$ so that the variance is infinite but the mean is finite. This has been a popular assumption for two reasons. The practical reason is the extensive traffic measurements of *on* periods reported in Ref.[86] where measured values of α were almost always in the interval (1,2). The theoretical reason is that mathematical analysis of models has been based on renewal theory and without a finite mean, stationary versions of renewal processes do not exist and (uncontrolled) buffer content stochastic processes would not be stable. See for example Ref.[39].

(iii) F has relatively thin tails so that the variance is finite. This includes classical models for telecommunication.

Section 2 defines the infinite source Poisson model and defines the basic descriptor processes: number of active sources at time t, cumulative inputted traffic to the sever in $[0, t]$, traffic rate, buffer content at t, time for buffer overflow of level $\gamma > 0$. The traffic rate process is the cumulative inputted traffic in small time intervals. This is obtained from the cumulative traffic by differencing. This section also considers known Gaussian and jump process approximations to the basic description. For Brownian motion approximations we could not find a proof in the literature, and hence have provided one.

In Section 3 the statistical methods we have used are presented.

Subsequent sections analyze four data sets in order to see what features of the model are consistent with the data. The data sets are:

- The Boston University data recording http sessions in two labs between November 1994 and February 1995. This data is available at http://ita.ee.lbl.gov/html/traces.html. We analyzed a 8h 20min part of the trace, with mean traffic rate 30 kbit/s.
- The UC Berkeley data with an 18 day trace collected in Nov. 1996 which contains the home IP HTTP traffic processed by UC Berkeley during this period. It is available at http://ita.ee.lbl.gov/html/contrib/UCB.home IP HTTP.html. Here we analyzed a three hour peak portion of the data, with mean traffic rate of 341 kbit/s.
- A low resolution and a high resolution data set from the two universities in Munich which were kindly supplied to us by Helmuth Gogl. The low resolution traces were collected around the clock on Wednesday, November 12, 1997 and Wednesday, December 17, 1997, and consist of the total number of cells which passed an ATM link in every two second interval, in the transmit and receive directions with mean traffic rates 5.6 and 8.8 Mbit/s, respectively. The high resolution trace records 1690729 inter arrival times of ATM cells from a measurement that captured all traffic in the sending direction of the link over a period of approximately 137 seconds with a resolution of 1 microsecond. The measurement was done on Tuesday December 23, 1997 starting at 14:48:15 and ending at 14:50:33 with a mean traffic rate of 5 Mbit/s.
- The number of bits transmitted and time stamps of starts and completions of HTTP file transfers to and from a corporate WWW Ericsson server, collected on Thursday and Friday, October 15 and 16, 1998. the trace was very non stationary, and we restricted the analysis to a 33 minute part of the trace, with a mean traffic rate 273 kbit/s.

For each data set we attempt to assess compatibility of the data with the model. As expected, there is not always a perfect fit to say the least. Depending on the richness of the available data, we wished to sort out the following issues for each data set. Limitations of time and manpower meant we were not always completely successful.

(1) How do you identify time points which are statistically verified to form a Poisson process? We examined initializations of sessions and beginnings of data bursts. When humans log in, it is plausible and widely believed to be an Internet invariant that the initialization times form a Poisson process but totally implausible that machine generated downloads triggered by an initial http request would follow the Poisson assumption. However, even for events caused by humans, aggregation beyond some scale will have to take place in order for Poisson behavior to be in force.

(2) We verify heavy tails for such quantities as file sizes, and file transfer times. Lengths of time depending on human activity such as the length of a login session at a student lab in the BU study is less likely to exhibit heavy tails. We expect to observe that fine scalings appropriate for measuring machine generated activity (milliseconds) are inappropriate for measuring human activity where seconds or even minutes are appropriate and this may explain why file transmission time

look heavy tailed but login sessions do not. Further investigations into the nature of distributions where we attempt to fit more than the tail are worthwhile and useful but it should be noted that it is likely that many macroscopic characteristics of a network will only be sensitive to tail behavior.

(3) We seek to study the distribution of transmission rates which can be defined as file size divided by the transfer time of the file. A more difficult question concerns dependence between transmission rates and file size and to what extent the dependence on network load is a measurable.

(4) We seek to understand the long range dependence behavior and the local Hölder behavior of combined traffic and relate this to the heavy tails found in item (2). An interesting statistical question is whether estimates of the tail parameter α, or equivalently of $\gamma = 1/\alpha$, based on such techniques as maximum likelihood estimation in generalized Pareto models and QQ plots[41,71,54,11] are consistent with estimates of the Hurst coefficient H and the Hölder exponent, since theoretically, the model guarantees that these two can be expressed as simple functions of α for the basic limit approximation. We estimate H and the Hölder exponent using wavelet and quadratic variation based methods
Note that for fractional Brownian motion (fBm), H and the Hölder exponent are identical so if in practice the estimates of these two quantities differ significantly, we have reason to doubt that fbm is the appropriate model.

(5) Many of the descriptors in the infinite source Poisson model have Gaussian or in some cases jump process approximations. We seek to examine the data to see if there are characteristics compatible with these approximations and, where possible, decide which type of approximation is more accurate.

(6) Much of the data exhibits evident non stationarities. (For example, loads on the Internet are heavy in afternoons, light in early morning.) How do you analyze non stationary data? One simple technique, because of the abundance of data, is to restrict attention to a subset of the data where behavior is likely to be more stationary. Although this is our primary technique, a more complete analysis could be performed by trying to view a natural period of time (one day?) as a period and to then deseasonalize the data.

2. BACKGROUND AND BASIC MODELS

For later convenience, we first collect some basic concepts, beginning with discussions of self similarity, Hurst and Hölder exponents and long range dependence. We define Lévy stable motion and fractional Brownian motion. We then define the infinite source Poisson model and give basic properties and descriptor quantities, and asymptotic approximations.

The main parameters of the infinite source Poisson model are the connection rate, λ, which is the intensity of starts of file transfers, and the tail parameter α of the transmission times. The three cases $\alpha < 1$, $1 < \alpha < 2$, and $2 < \alpha$ lead to different asymptotic approximations which below are treated in separate sections. The approximations are for large time intervals, $[0, T]$ and for the two last cases also assume that λ is large. A further

issue is that in the middle case the nature of the approximation depends on the relation between λ and T. These results are summarized in Table 2.1 below.

2.1. Self-Similarity, Hurst and Hölder Exponents, Long-Range Dependence

A stochastic process $\{X(t); 0 \leq t < \infty\}$ is self similar if there exists a constant H, the Hurst parameter, such that the finite dimensional distributions of the time changed and rescaled process $\theta^{\,H}X(\theta t)$ are the same as for the original process, i.e. in formulas, if

$$\theta^{\,H}X(\theta\cdot) \overset{fidi}{=} X(\cdot), \quad \text{for } 0 < \theta. \tag{2.1}$$

In particular, since a centered Gaussian process is uniquely determined by its covariance function, it is self similar if and only if its covariance function C satisfies

$$\theta^{\,2H}C(\theta s, \theta t) = C(s,t) \quad \text{for } 0 < \theta. \tag{2.2}$$

A second order process, Gaussian or not, which satisfies Eq. (2.2) is called second order self similar.

A fractional Brownian motion (fBm) B_H is a centered continuous Gaussian process with covariance function

$$C(s,t) = \frac{\sigma^2}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}), \quad 0 < H < 1.$$

Since C satisfies Eq. (2.2), B_H is self similar with Hurst parameter H. By setting $H = 1/2$ an ordinary Brownian motion, with independent increments, is obtained. It can be seen from the form of C that fBm has stationary (but not independent) increments also for $H \neq 1/2$. It follows that fractional Gaussian noise (fGn), i.e. the difference sequence $Y_k = X((k+1)\Delta) - X(k\Delta)$ is stationary (here Δ is the fixed length of differencing). More details can be found for instance in Ref.[77].

As discussed in the introduction, network traces looked at on widely varying time scales above a certain resolution are in the literature often claimed to have similar statistical properties, and all the limiting process which are discussed below are

Table 2.1. Limiting regimes and model numbers.

Range of α	Limiting regime	1st order approximation	Distributional approximation of A	Self-similarity parameter for approximation of A	Model number
$0 < \alpha < 1$	λ fixed, $T \to \infty$	$Tm(T)$	Gaussian	$(3 \quad \alpha)/2$	1
$1 < \alpha < 2$	λ fixed, $T \to \infty$	$\lambda\mu T$	Stable	$1/\alpha$	2
	$\lambda = \lambda(T)$, $T \to \infty$ slow growth	$\lambda(T)\mu T$	Stable	$1/\alpha$	
	$\lambda = \lambda(T)$, $T \to \infty$ fast growth	$\lambda(T)\mu T$	FBM	$(3 \quad \alpha)/2$	3
$\alpha > 2$	λ fixed, $T \to \infty$	$\lambda\mu T$	BM	1/2	4

self similar. This is completely as should be expected. If a process is obtained as a distributional limit by dilating time linearly and scaling space, then it has to be asymptotically self similar. See Refs.[56,25].

The semivariogram V of a second order process is defined by

$$V(t,\tau) = \frac{1}{2}E(X(t+\tau) - X(t))^2.$$

It is easy to see that C can be computed from V and vice versa. For a self similar process, $\theta^{2H}V(\theta t, \theta\tau) = V(t,\tau)$. If the process has stationary increments, V does not depend on its first variable, $V(t,\tau) = V(\tau)$, and if it also is self similar, then $V(\tau) = V(1)\tau^{2H} := c\tau^{2H}$. If the process is also a centered Gaussian process, this means that it is a fBm. In general, a process is said to have the local Hölder (mean square) index H_0 at t if the semivariogram satisfies

$$V(t,\tau) = c\tau^{2H_o} + o(\tau^{2H_o}), \quad \text{as } \tau \to 0, \tag{2.3}$$

for each t. For a fBm, $H = H_o$, as is easily seen. The Hölder index of a Gaussian process gives precise information on sample paths (see *e.g.* Adler,[5] Ibragimov and Rozanov[43]), on the rate of convergence of non parametric estimates of the covariance function (Istas and Laredo[46]) and on the asymptotic behavior of wavelet coefficient of X (Istas,[44]). There are other definitions of Hölder indices more suitable to the study of path properties of non Gaussian processes such as multifractals (cf.[75]). Data network applications are discussed in Refs.[28,31,76].

A process X_α is an α stable Lévy motion if it has stationary independent increments which follow a non normal stable distribution with index α, $0 < \alpha < 2$. Clearly $\theta^{1/\alpha}X_\alpha(\theta\cdot)$ also has stationary increments, for any $\theta > 0$. Further, by the characterizing property of the stable distributions, $\theta^{1/\alpha}X_\alpha(\theta t)$ has the same distribution as $X_\alpha(t)$ for any $t > 0$. It follows that X_α is self similar with Hurst parameter $H = 1/\alpha$. An iid sequence of α stable random variables is called α stable noise. Thus, in particular the sequence $X_\alpha((k+1)\Delta) - X_\alpha(k\Delta)$; $k = 0, 1, \ldots$ is α stable noise for $\Delta > 0$. Extensive information about stable processes may be found in Ref.[77].

Since stable variables have infinite variances, the variogram and the local Hölder exponent are not defined for stable Lévy motions.

A (centered) second order stationary process $X(t)$ displays long range dependence (LRD) if its covariance function decreases at a polynomial rate at large lags,

$$C(\tau) \sim K\tau^{\beta-1}, \quad \tau \to \infty, \tag{2.4}$$

where K is a constant and $0 < \beta < 1$. Another way to think about this is that the spectral density $\Gamma(\nu)$ has a polynomial divergence at the origin,

$$\Gamma(\nu) \sim K'|\nu|^{-\beta}, \quad \nu \to 0,$$

for some other constant K'. The paradigm of such processes in fGn with $1/2 < H < 1$. The LRD parameter then is related to H by $\beta - 1 = 2H - 2$, i.e. $H = (\beta+1)/2$. By analogy with fractional Gaussian noise, the quantity $(\beta+1)/2$ is sometimes referred to as the "Hurst parameter" of the process as soon as the behavior (2.4) is observed, even if the process is not self similar.

2.2. The Infinite Source Poisson Model

We now review the elements of a data transmission model used in Refs.[40,48,49,72,63]. Let $\{\Gamma_k,\ k \geq 1\}$ be the points of a rate λ homogeneous Poisson process on $\mathbb{R}_+ = [0, \infty)$ so that $\{\Gamma_{k+1} - \Gamma_k,\ k \geq 1\}$ is a sequence of iid exponentially distributed random variables with parameter λ. (In the stationary case the Poisson process instead should be defined on $\mathbb{R} = (-\infty, \infty)$, which leads to some straightforward change of notation.) We imagine that a communication system has *sources* or nodes, and at time Γ_k a *connection* is made and a source begins a transmission at unit rate to or from the server. The duration of this transmission is a random variable L_k with distribution F, usually of the form $\bar{F}(x) = x^{\alpha}\ell(x)$. When F has a finite first moment, it is convenient to set

$$\mu = E(L_1) = \int_0^\infty xF(dx).$$

The input rate could be made to deterministically vary over the transmission time of length L_k as in Ref.[53] but we do not discuss this and other variants. See Refs.[61,62]. We sometimes refer to the Poisson rate λ as the *connection rate*.

The counting function

$$M = \sum_{k=1}^{\infty} \epsilon_{(\Gamma_k, L_k)}, \tag{2.5}$$

on $\mathbb{R}_+ \times (0, \infty)$ corresponding to the points $\{(\Gamma_k, L_k),\ k \geq 1\}$, is a two dimensional Poison process on $\mathbb{R}_+ \times (0, \infty]$ with mean measure $\lambda\mathbb{L} \times \mathrm{F}$, where $\mathbb{L}$ stands for Lebesgue measure. (cf.[67].)

This model is stable under aggregation. If the traffic from two independent infinite source Poisson models are superposed, then the result also follows an infinite source Poisson model, with a connection intensity which is equal to the sum of the two intensities, and a transmission length distribution which is a mixture of the two transmission length distributions.

2.2.1. Descriptors

Define $N(t)$, the *number of active sources* at time t by

$$N(t) = \sum_{k=1}^{\infty} 1_{[\Gamma_k \leq t < \Gamma_k + L_k]} = M(\{(\gamma, l) \in \mathbb{R}_+ \times (0, \infty] : \gamma \leq t < \gamma + l\}).$$

The second expression makes it clear that for each t, $N(t)$ is a Poisson random variable with parameter

$$\lambda\mathbb{L} \times F(\{(\gamma, l) \in \mathbb{R}_+ \times (0, \infty] : \gamma \leq t < \gamma + l\}) = \lambda \int_0^t \bar{F}(t - \gamma) d\gamma$$

$$= \lambda \int_0^t \bar{F}(s) ds =: m(t). \tag{2.6}$$

During a transmission, the transmitting source is sending data to the server at unit rate. The total *cumulative traffic* in $[0, t]$ is

$$A(t) := \int_0^t N(s)ds = \sum_{\Gamma_k \le t} (L_k \wedge (t - \Gamma_k)), \tag{2.7}$$

which expresses A as a shot noise process.[52] Since an active node transmits at unit rate, the overall transmission rate at time t is $N(t)$ and a surrogate for this, which is more easily measurable, is the *traffic rate process* defined as

$$\{A((k+1)\Delta) - A(k\Delta), k \ge 0\},$$

for some fixed $\Delta > 0$. Assume the server works at constant *output rate* r. The *buffer content* at time t, $X(t)$, satisfies the storage equation

$$dX(t) = N(t)dt - r1_{[x(t)>0]}dt, \tag{2.8}$$

or[38,9,67]

$$X(t) = \bigvee_{s=0}^{t} [A(t) - A(s) - r(t-s)] = \bigvee_{s=0}^{t} \int_s^t (N(s) - r)ds, \tag{2.9}$$

where we have assumed the initial condition $X(0) = 0$.

We now describe known behavior of the basic descriptors for the three cases discussed in the introduction.

2.2.2. *The Case* $0 < \alpha < 1$

The model has been studied in Ref.[72] in this α regime where the mean transmission time is infinite, assuming the connection rate λ is constant and the time interval $[0, T]$ expands to infinity. The basic descriptor processes are all explosive and we have in probability as $T \to \infty$ that

$$\frac{N(T)}{m(T)} \xrightarrow{P} 1, \quad \frac{A(t)}{Tm(T)} \xrightarrow{P} \frac{1}{2-\alpha}, \quad \frac{X(T)}{Tm(T)} \xrightarrow{P} \frac{1}{2-\alpha}.$$

Recall $m(T)$ was defined in Eq. (2.6). Though measurement studies[7,72] report file size estimates with tails having pareto parameter in this range $\alpha < 1$, this case can only describe explosive system behavior prior to the time when the influence of system controls is felt.

Second order behavior gives Gaussian process approximations for basic descriptors in terms of $\{G(t), t \ge 0\}$, a continuous path Gaussian process with covariance function

$$C(s,t) = (s \bigvee t)^{1\ \alpha} - |t-s|^{1\ \alpha}, \quad 0 \le s \le t. \tag{2.10}$$

As $T \to \infty$ we have for $t > 0$

$$\frac{N(Tt) - m(Tt)}{\sqrt{m(T)}} \Rightarrow G(t), \quad \frac{A(Tt) - \int_0^{Tt} m(s)ds}{T\sqrt{m(T)}} \Rightarrow \int_0^t G(s)ds,$$

and $X(\cdot)$ has a weak limit expressed in terms of $\int_0^t G(s)ds$.

2.2.3. *The Case* $1 < \alpha < 2$

For this case, different approximations have been studied under different limiting regimes. It is possible to seek approximations when

(a) λ fixed and $T \to \infty$.[54,74]
(b) $\lambda = \lambda(T) \to \infty$ so that λ is allowed to increase with T.[63]
(c) $\lambda \to \infty$ and then $T \to \infty$ in that order.[79]
(d) $T \to \infty$ and then $\lambda \to \infty$ in that order.[79]

When (a) holds, one can check that conditions in Ref.[52] for convergence of $A(\cdot)$ to a limiting self similar Gaussian process fail. Furthermore, we known $N(T)$, the number of active sources at T, is Poisson distributed with parameter $m(T)$ but since for the case $1 < \alpha < 2$, $m(T) \to m(\infty) = \lambda\mu < \infty$, it follows that as $T \to \infty$,

$$N(T) \Rightarrow \mathrm{PO}(\lambda\mu),$$

a Poisson random variable with parameter $\lambda\mu$ and as a family of processes $N(T\cdot)$ becomes asymptotically uncorrelated. So one cannot get the asymptotic behavior of $A(\cdot)$ from $N(\cdot)$.

There is no centering and scaling to make $A(T\cdot)$ asymptotically a Gaussian process, but one gets a stable limit. Define

$$b(T) = \left(\frac{1}{1-F}\right)^{\leftarrow}(T) = T^{1/\alpha}\ell_1(T), \tag{2.11}$$

for some slowly varying function ℓ_1. A variant of (Ref.[53], Theorem 4) in particular gives, for λ fixed, $T \to \infty$, that

$$\frac{A(Tt) - \lambda\mu Tt}{b(T)} \Rightarrow X_\alpha(t), \tag{2.12}$$

in the sense of convergence of finite dimensional distributions, where $X_\alpha(\cdot)$ is an α stable Lévy motion whose marginal distribution is totally skewed to the right. See also Refs.[63,73]. So on large time scales, $A(T\cdot)$ looks like an α stable Lèvy motion. Interestingly, one gets easily from the definition of $A(T)$ that $\sigma(T) := \sqrt{Var(A(T))}$ satisfies

$$\sigma(T) \sim (\text{const})\sqrt{T^3\bar{F}(T)} \sim (\text{const})T^{(3\ \ \alpha)/2}\ell^{1/2}(T),$$

which may be compared to the asymptotic form of $b(\cdot)$ given in Eq. (2.11). Observe that for $\alpha \in (1,2)$, $\frac{1}{\alpha} < \frac{3-\alpha}{2}$.

From Eq. (2.12), we find the first order growth rate of $A(T)$ to be $\left(\text{since}\,\frac{1}{2} < \frac{1}{\alpha} < 1\right)\lambda\mu T$ so that in probability $A(T) \sim \lambda\mu T$.

Provided the constant output rate r satisfies $\lambda\mu < r$, the $X(\cdot)$ process of Eqs. (2.8) and (2.9) has negative drift and is stable. Being regenerative, $X(T)$ will have a limit distribution given by standard renewal theory.

The limit behavior in Eq. (2.12) assumes that the connection rate λ is constant and the time scale T is growing. If T is fixed and $\lambda \to \infty$, then one may expect $A(\cdot)$ to be asymptotically Gaussian. Thus, on small or moderate time scales, if the input rate is large, the cumulative inputted traffic should be approximately Gaussian. This in fact was found in Ref.[82] as well as the result that if after letting $\lambda \to \infty$ one lets $T \to \infty$, then $A(\cdot)$ is approximated by a fractional Brownian motion. On the other hand, as seen above, on large time scales, with modest connection rate, the cumulative inputted traffic is approximately a Lévy stable motion. Allowing the connection rates to vary with T produces the following result which give conditions under which either a fBm or Lévy stable motion is an appropriate approximation.[63]

Proposition 1. *Assume a family of infinite source Poisson models indexed by T, where in the T th model, $\lambda = \lambda(T)$ depends on T. Suppose the connection length distribution does not depend on T. Recall the definition of $b(\cdot)$ in Eq.* (2.11)*. For the Tth model, let $A_T(\cdot)$ be the cumulative input and $N_T(t)$ be the number of active sources at time t.*

(i) Assume the Poisson rate $\lambda = \lambda(T)$ depends on T so that one of the following equivalent slow growth conditions is satisfied:

(1) $\lim_{T\to\infty} \frac{b(\lambda T)}{T} = 0.$
(2) $\lim_{T\to\infty} \lambda T \bar{F}(T) = 0.$
(3) $\lim_{T\to\infty} Cov(N_T(0),\ N_T(T)) = 0.$

Then the process $(A_T(Tt), t \geq 0)$ describing the cumulative input in $[0, Tt], t \geq 0$, satisfies the limit relation

$$\frac{A_T(T\cdot) - T\lambda\mu(\cdot)}{b(\lambda T)} \overset{fidi}{\to} X_\alpha(\cdot),$$

where $X\alpha(\cdot)$ is a Lévy α stable motion. Here $\overset{fidi}{\to}$ dentotes convergence of the finite dimensional distributions.
(ii) Assume the Poisson rate $\lambda = \lambda(T)$ depends on T so that one of the following equivalent fast growth conditions is satisfied:

(1) $\lim_{T\to\infty} \frac{b(\lambda T)}{T} = \infty.$
(2) $\lim_{T\to\infty} \lambda T \bar{F}(T) = \infty.$
(3) $\lim_{T\to\infty} Cov(N_T(0),\ N_T(T)) = \infty.$

Then the process $(A_T(Tt), t \geq 0)$ describing the cumulative input in $[0, Tt]$, $t \geq 0$, when properly normalized as

$$\frac{A_T(T\cdot) - T\lambda\mu(\cdot)}{\sqrt{\lambda T^3 \bar{F}(T)}}$$

converges in $C[0,\infty)$ *to a fractional Brownian motion with self similarity parameter* $H = (3-\alpha)/2$.

Heavy traffic approximations for $A(\cdot)$ and $X(\cdot)$ are considered in Ref.[55].

2.2.4. *The Case* $\alpha > 2$

We again assume λ is fixed but the results undoubtedly hold also if λ is allowed to increase with T. As in the case where $1 < \alpha < 2$, $N(T)$ is asymptotically Poisson and as $T \to \infty$, $N(T\cdot)$ is asymptotically uncorrelated. Hence, we cannot expect to get the asymptotic behaviour of $A(\cdot)$ from $N(\cdot)$.

For the case $\alpha > 2$, the cumulative input $A(\cdot)$ can be approximated by a Brownian motion. We set

$$A^{(T)}(t) := \frac{A(Tt) - \lambda T t \mu}{\sqrt{\lambda T E(L_1^2)}} \tag{2.13}$$

and, in $C[0,\infty)$, the space of continuous functions on $[0,\infty)$, as $T \to \infty$,

$$A^{(T)}(\cdot) \Rightarrow B(\cdot), \tag{2.14}$$

a standard Brownian motion. Furthermore, this implies a functional central limit theorem for $X(\cdot)$: If $\lambda\mu > r$, so that the system is unstable,

$$\frac{X(Tt) - Tt(\lambda\mu - r)}{\sqrt{\lambda T E(L_1^2)}} \Rightarrow B(t) \tag{2.15}$$

in $C[0,\infty)$. If $\lambda\mu < r$, then the limit of $X(Tt)$ is 0.

These claims are proved in the Appendix.

2.2.5. Summary

We summarize some of the relevant facts about the infinite source Poisson model in Table 2.1. The column "model number" is used for referencing in Table 4.1 below.

3. ESTIMATION METHODS

This section gives a brief introduction to the statistical methods we have used. Many are standard, but will still be mentioned briefly and references will be given, in keeping with the aim to give a guide for analysis. However some are of more recent origin. Specifically, we have used rather recently developed quadratic variation and wavelet methods to estimate the Hurst and Hölder exponents. These methods will be presented in more detail.

In some situations, only traces from aggregated traffic are measured but a micro level model is needed for simulation. Hence, a further topic for the present section is to what extent it is possible to infer parameters of the infinite source Poisson model from measurements of aggregate traffic.

In practice, a frequent goal of network measurement analysis is to construct a simulator for the design of networks. A simulator needs the parameter values for the model as an input. The statistical methods we describe provide such values, which can be used directly, or as a basis for extrapolation to what parameter values may be expected in the future, for new groups of subscribers. In this section, we also make some brief comments about design of simulators.

The methods we discuss throughout assume stationarity of the measured traffic. In reality the characteristics of the traffic, especially of the process of starts of file transfers, vary substantially with the time of day, between different days in the week, and between seasons, and contain strong trends. We circumvent this problem by selecting visually stationary parts of the traces for study. In fact, this may correspond to what is practically interesting: it's the behavior during the (approximately stationary) peak periods which is of primary interest.

3.1. Testing for Independence of Heavy-Tailed Variables

The infinite source Poisson model assumes that the transfer times are independent, and similarly the stable limit for aggregated traffic implies that successive traffic rate measurements are independent.

A basic approach to testing for independence is to use the correlation function. However correlations of heavy tailed data may have a more complex behavior than in the light tailed case, see e.g. Refs.[15,24,30,70,73] and have to be interpreted with some care. A standard useful way to circumvent this problem is to make an appropriate marginal transformation before computing correlations. In the present situation this amounts to using the logarithms of the data.

A drawback with this approach is that taking logarithms obscures the impact of the very large transfer times which are of major importance. Hence it is desirable to complement with methods which do not use transformations. We employed two methods to check for independence which use the original observations. Both are based on the heavy tailed acf, i.e., on the autocorrelation function computed without subtracting means.

The first, informal, method can detect many forms of nonlinear dependence in heavy tailed observations. It simply is to split the data into parts, say two to five parts, and to compute the heavy tailed acf on each of the parts. If the observations in fact are i.i.d., or come from a linear process, then these acf's should look the same. The second method is to use simulations from the known limiting distribution of the heavy tailed acf under independence to construct confidence intervals, and to reject independence if the observed acf deviates from these intervals. For more details, see Ref.[30] and for a more formal treatment consult.[73] (An alternative, attractive, method is to use permutation tests for, say, the maximum of the heavy tailed acf over some number of lags to judge if this is larger

than what is caused by randomness alone. However, for very large data sets, such as the present ones, this method is computationally burdensome.)

3.2. Marginal Distribution and Estimation of Means of Heavy Tailed Variables

The file sizes, file transfer times, file transfer rates (= (file size)/(file transfer time)), and traffic rates (cumulative inputted traffic in small time intervals) have similar statistical properties, and are amenable to the same types of analysis.

The precise shapes of the distributions are affected by special, and rapidly changing conditions, such as, for example, automatic hookup to Netscape's homepage. Thus only nonparametric (histogram or kernel density) estimation seem reasonable for the central parts of the distribution. However, at high aggregation levels, these details are less important, and what is needed are good estimates of expected values and tails of the distributions.

For heavy tailed non negative iid data $X, X_1 \ldots, X_n$ (such as the ones analyzed below) with tails $P(X > x) \sim Lx^{-\alpha}$ with $1 < \alpha < 2$, where existing statistical techniques require that we now assume that the slowly varying function L in Eq. (1.1) is a constant L, the standard finite variance estimates of variation of the mean are not applicable. It is still possible to estimate the expected value $\mu = E(X)$ by the mean of the observation and an asymptotic confidence 100p% confidence interval for μ may be obtained as

$$\bar{X} \pm (\hat{L})^{1/\hat{\alpha}} n^{1/\hat{\alpha}-1} c_{\hat{\alpha}}^{1/\hat{\alpha}} S_{\hat{\alpha},1,1}(1-p), \tag{3.1}$$

with $c_\alpha = \Gamma(2-\alpha)|\cos(\pi\alpha/2|/(\alpha-1)$ (c.f. Ref. [77], Theorem 1.81.). Here $S_{\alpha,1,1}(1-p)$ is the $1-p$ th quantile of the completely (positively) asymmetric stable distribution with tail parameters α, in the parametrization of the cited reference. Estimates of L, α are discussed in the next section.

3.3. Tail Estimation

There is already a substantial literature on this topic, see e.g. Ref.[26] and references therein. Since one is interested in events of a size rarely encountered in the available data, a parametric model for the tail is unavoidable. The standard procedure is then to choose an appropriate high level u and to use only those observation exceeding u for the estimation of the relevant parameters. Note that estimates will be sensitive to the choice of level u.

On theoretical grounds (asymptotic theory, stability under change of level, and "the right amount of flexibility") the Poisson generalized Pareto model is preferred. In this model, the exceeding over the threshold u are i.i.d. and follow a generalized Pareto (GP) distribution of the form

$$G_\gamma(x) = 1 - \left(1 + \frac{\gamma x}{\sigma}\right)^{1/\gamma}, \quad 1 + \frac{\gamma x}{\sigma} > 0, \quad \sigma > 0, \quad \gamma \in \mathbb{R}.$$

This is a heavy tailed distribution for $\gamma > 0$ with $\alpha = 1/\gamma$; for $\gamma = 0$ this is the exponential distribution and for $\gamma < 0$ the distribution has finite upper endpoint. See for example Ref.[26]. The times at which exceeding occur follow a Poisson process which is independent of the sizes of the exceeding. This model is implied by the stronger assumption that the observations $\{X_n\}$ are iid with a regularly varying tail.

The intensity of the Poisson process is simply estimated by the number of exceedings divided by the total number of observations and the parameters of the GP distribution may be estimated by the method of maximum likelihood. S + programs for this estimation have been made available by A. McNeill, http://www.math.ethz.ch/~mcneil/. Another relevant software package, Xtremes, is a PC package for analyzing graphing extreme values and comes on a disk with the book.[66] Straighforward asymptotic confidence intervals based on the information matrix do not perform well for a small or moderate number of exceedings, but instead profile likelihood intervals do, see Ref.[78].

In addition to maximum likelihood analysis, we have used semiparametric graphical methods for variables with a Pareto like tail, such as qq plots of quantiles of a standard exponential distribution against the logarithms of the ordered exceedings and estimating α by the slope of an ordinary regression line. This method of QQ plotting does not use the detailed form of the GP distribution and is asymptotically less efficient than other similar estimators, such as the Hill estimator. However, it can be used with any standard statistical software, and has the advantage of being directly linked to a readily interpreted graphical display. It also gives a useful impression of the size of deviations from a Pareto model, since a pareto tail corresponds to a straight line. The method is discussed in detail in Ref.[54].

Finally a practical consideration, similar to the choice of bandwidth in density estimation, and a subject of much current research, is the choice of the level u, or, equivalently, the number of order statistics used for the tail estimation. In this paper we have used informal, graphical methods as discussed e.g. in Ref.[71], which look at changes in the estimators as the level u is changed.

3.4. Starts of File Transfers

That the file requests appear as a Poisson process is equivalent to the interarrival times being i.i.d. and exponentially distributed. We have investigated the Poisson process assumptions by making qq plots of the interarrival times against exponential quantiles, and by looking at the correlation functions of the interarrival times.

3.5. Stable Lévy Processes

At low aggregation levels, and for $1 < \alpha < 2$ traffic rates in the infinite source Poisson model are asymptotically stable and independent. For maximum likelihood estimation of the parameters and model checking we have followed the prescriptions in Ref.[64] and used the very convenient accompanying software which can be downloaded from http://www.cas.american.edu/~jpnolan/stable.html.

We have used the most common parametrization of stable distributions, which corresponds to the characteristic function

$$E\exp(itX) = \begin{cases} \exp\left(-\sigma^{\alpha}|t|^{a}\left(1 - i\beta\tan\left(\frac{\pi\alpha}{2}\right)\text{sign}(t)\right) + i\delta t\right) & \alpha \neq 1 \\ \exp\left(-\sigma|t|\left(1 + i\beta\frac{2}{\pi}sign(t)\ln|t|\right) + i\delta t\right) & \alpha = 1 \end{cases}.$$

3.6. The Hölder Index

For a Gaussian process $\{X(t), t \geq 0\}$ a short survey of estimators for the local Hölder index H_o is given in Lang.[57] Hall and wood[36] study the so called box counting the estimation method based directly on the capacity of fractal dimension[42] and show that the estimator has a large asymptotic bias. Hall *et al.*[37] and Feuerverger *et al.* studied estimators based on level crossings. Constantine and Hall,[16] cf. also Gladyshev[34] discussed estimators based on simple empirical quadratic variation. At scale $1/n$ and for a process with stationary increments on $[0, 1]$, this is

$$V_n = \frac{1}{2}\sum_{k=1}^{n}(X(k/n) - X((k-1)/n))^2,$$

which suitably normalized converges with probability 1,

$$\lim_{n\to\infty} n^{2H_o\ \ 1}V_n = c,$$

where c is given by Eq. (2.3), i.e. by $V(\tau) = c\tau^{2H_o} + o(\tau^{2H_o})$.

However, these quadratic variations estimators are not scale invariant and all of the estimators mentioned so far converge slowly, Istas and Lang[45] and Kent and wood[51] propose an improved method, where the "simple" difference $X(k/n) - X((k-1/n)$ is replaced by general discrete differencing of X. More precisely, they introduce the weights $a = (a_0, a_1, \ldots, a_p)$ a discrete difference operator of degree $p \geq 1$, and consider the empirical quadratic variations associated with it,

$$U_{a,n} = \sum_{j=1}^{n}{}^{p+1}\left(\sum_{i=0}^{p} a_i X((i+j-1)/n)\right)^2.$$

In this paper we use an estimator of H_o which is based on two such "general" quadratic variations. Specifically, from the sequence $\boldsymbol{a} = (1, -5, 10, -10, 5, -1)$ we define a sequence $\boldsymbol{b}$ with "double time mesh," *i.e.* by putting $b_{2i} = a_i$ and $b_{2i+1} = 0$ for $0 \leq i \leq 5$ and estimate H_o by

$$\hat{H}_o = \frac{1}{2}\ln_2\left(\frac{U_{\boldsymbol{b},n}}{U_{\boldsymbol{a},n}}\right).$$

This estimator is discussed in Istas and Lang (Ref.[45], p. 432). This sequence $\boldsymbol{a}$ corresponds to the binomial coefficients in the expansion of $(u - v)^5$, $(u, v) \in \mathbb{R}^2$. An advantage of the estimator is that it filters out polynomial trends up to order 4, and hence is rather robust to smooth non stationarities of the mean.

3.7. The Hurst Exponent

In this section we describe a recently developed wavelet estimation methodology. The method works under a variety of assumptions. If the process under investigation is stationery with finite variance, the method yields an estimate of the parameter β introduced in Eq. (2.4), with $\beta > 0$ corresponding to LRD. When applied to the increments of a selfsimilar process, the wavelet method yields an estimate of $2H - 1$, as shown below. Now, for a selfsimilar process with finite second moments and stationary increments the increments process has $\beta = 2H - 1$. Hence also for this case LRD corresponds to $\beta > 0$.

However, if the selfsimilar process has infinite second moments, $2H - 1$ can be greater than zero even if there is no LRD. For example, stable Lévy processes have Hurst parameter $H = 1/\alpha$, so that $2H - 1 > 0$ (since $0 < \alpha < 2$) even though the process has independent increments and hence is as far from LRD as possible. To summarize, for processes with finite variance, if the wavelet method gives an estimate clearly greater than zero, then there is LRD. However, for processes with heavy tails an estimate of which is greater than zero does not necessarily mean that there is LRD.

A naive estimator for β may be obtained from the empirical covariance function. For strongly correlated data this is known to be unreliable. Several alternative estimation methods have been developed: the aggregated variance, Whittle, R/S, absolute value, periodogram, variance of residuals, aggregated Whittle, local Whittle and wavelet methods. The first three methods are treated in Ref.[12]. For discussion and comparisons of all of the methods we refer to Refs.[6,3] and references therein. Here we use the wavelet method because it provides an appealing compromise between low computational cost and good statistical performance. It is also more flexible than maximum likelihood based estimators such as Whittle's estimators since it does not require an exact parametric model for the spectral density. In addition, it is based on identification of scaling in a log log diagram, which makes it possible to judge the range of scales on which the model fits. A final advantage is that it is robust to smooth non stationarities. In many cases we have also computed, as a sanity check, estimates using traditional techniques such as the R/S statistic and associated plot and the variance time plot, and in all cases agreement with the wavelet method was good.

We here outline the main ideas and refer to Refs.[1,3,4,82,83] for an exhaustive presentation. A *wavelet* ψ is a smooth function which is well localized both in position and frequency and which satisfies the admissibility condition $0 = \int \psi = \mathcal{F}(\psi)(0)$, where $\mathcal{F}(\psi)$ is the Fourier transform of ψ. Usually, it is also required that the wavelet has some vanishing moments, $\int t^n \psi(t) = 0,\ n = 0, 1 \ldots N$. By "well localized" is meant that the function has compact support or at least is rapidly decaying. To a reference wavelet ψ, usually called the "mother wavelet", is associated a two parameter family of functions ψ_{ba} obtained by translation and dilation,

$$\psi_{ba}(t) := \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad b \in \mathbb{R}, \quad a > 0.$$

If the mother wavelet ψ is localized around some central position t_0, then it is clear form the definition that ψ_{ba} is shifted to position $at_0 + b$. Similarly, if the Fourier transform $\mathcal{F}(\psi)$ is centered around some central frequency ν_0, then $\mathcal{F}(\psi_{ba})$ is centered around $a^{-1}\nu_0$.

The admissibility condition $\hat{\psi}(0) = 0$ ensures that the function stays localized in frequency after dilation. The wavelet transform $W_\psi X$ of a process X is a function on the position frequency plane $\mathbb{R} \times \mathbb{R}^+$,

$$W_\psi X(b,a) := \int \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) X(t)dt = \int \sqrt{a}\psi(s)X(as+b)ds. \tag{3.2}$$

Roughly speaking, the filter, $W_\psi X(b,a)$ retains the part of the process which contributes to the frequency a at the position b. No information is lost, in the sense that the original signal can be retrieved from Eq. (3.2) by a reconstruction formula.

For random processes, the wavelet transform captures the stationarity and scaling properties. If the process $X(\cdot)$ is stationary, or has stationary increments then the process $W_\psi X(\cdot, a)$ is again stationary as is clear from the second formula in Eq. (3.2). If a process Y is H self similar with stationary increments and finite second moments, then it follows readily from the definition that

$$E|W_\psi Y(b,a)|^2 = a^{2H+1} E|W_\psi Y(b/a, 1)|^2$$

If we define $X(t) = Y(t+\Delta) - Y(t)$ for $\Delta > 0$, then by change of variables, and using a first order Taylor expansion and self similarity in the third step,

$$\begin{aligned} W_\psi X(b,a) &= \int \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \{Y(t+\Delta) - Y(t)\} dt \\ &= \int \sqrt{a} \left\{ \psi\left(s - \frac{b+\Delta}{a}\right) - \psi\left(s - \frac{b}{a}\right) \right\} Y(as)dt \\ &\sim \int \sqrt{a} \left\{ -\frac{\Delta}{a}\psi'(s) \right\} a^H Y(s)ds = -a^{H\ 1/2}\Delta \int \psi'(s)Y(s)ds, \quad a \to \infty, \end{aligned} \tag{3.3}$$

under suitable regularity conditions. Thus, if second moments exist

$$C(a) = E|W_\psi X(b,a)|^2 \sim K a^{2H\ \ 1}, \quad a \to \infty, \tag{3.4}$$

where $K = \Delta^2 E|\int \psi'(s)Y(s)ds|^2$. Thus, in particular this holds for fGn.

Now assume $X(\cdot)$ is any stationary process with the LRD property (2.4). The large scale behaviour of the wavelet coefficients is

$$C(a) = E|W_\psi X(b,a)|^2 \sim K_\beta a^\beta, \quad a \to +\infty, \tag{3.5}$$

where $K_\beta = K \int d\nu |\nu|^{\ \ \beta} |\mathcal{F}(\psi)(\nu)|^2$, as shown by a straightforward comoputation. This of course agrees with Eq. (3.4) for fGn, since $H = (\beta+1)/2$.

Even if the original process $X(\cdot)$ is long range dependent, the corresponding wavelet transform is short range dependent as a function of b. The idea of the wavelet estimator of the LRD parameter for LRD stationary processes is to take advantage of this decorrelation and to compute an estimator $\hat{\beta}$ using Eq. (3.5). In practice, this is done using so called multiresolution analysis, which provides a fast algorithm to compute the wavelet coefficients on a dyadic grid in the position time plane, i.e. the coefficients $d(j,k) = W_\psi X(2^j k, 2^j)$, $j,k \in \mathbb{Z}$. The information contained in these coefficients is sufficient to

reconstruct the process (see Ref.[23]). By Eq. (3.5)

$$C(2^j) = E|d(j,k)|^2 \sim K_\beta 2^{j\beta}, \quad j \to +\infty, \tag{3.6}$$

and the decrease of correlation between the wavelet coefficients $d(j,k)$ is controlled by the number of vanishing moments, N, in the following way,[81,29]

$$|Ed(j,k)d(j,k')| = O(|2^j(k-k')|^{+\beta - 1 - 2N}), \quad |2^j(k-k')| \to \infty. \tag{3.7}$$

Now, $C(2^j)$ may be estimated by the sum of the coefficients at fixed scale,

$$\hat{C}(2^j) = \frac{1}{n_j}\sum_k |d_{j,k}|^2,$$

where n_j is the number of available coefficients at scale 2^j. The parameter β in Eq. (3.5) or equivalently $H = (\beta+1)/2$ is then simply estimated from a linear regression in the log log diagram of $\widehat{C(2^j)}$.

However, taking the logarithm introduces a bias ($E\log \neq \log E$). Under the simplifying assumptions that the process is Gaussian and the wavelet coefficients are perfectly decorrelated (instead of Eq. (3.7)), $\hat{C}(2^j)$ is a sum of chi squared independent variables, and the bias can be explicitly computed and removed.

Since the variance of the wavelet coefficients increases with the scale, the quality of the estimator is improved by performing a weighted linear regression, which gives more weight to small scales. Altogether, the estimator of β is defined as, see Ref.[83],

$$\hat{\beta} = \frac{\sum y_j(jS - S_j)/\sigma_j^2}{SS_{jj} - S_j^2},$$

where

$$S = \sum 1/\sigma_j^2, \;\; S_j = \sum j/\sigma_j^2, \;\; S_{jj} = \sum j^2/\sigma_j^2,$$

with

$$g(j) = \frac{\Gamma'(n_j/2)}{\Gamma(n_j/2)\log 2} - \log_2(n_j/2) \simeq -\frac{1}{n_j \log 2},$$

$$y_j = \log(\hat{C}(2^j)) - g(j),$$

$$\sigma_j^2 = Var(y_i) = \frac{\zeta(2, n_j/2)}{\log^2 2} \simeq \frac{2}{n_j \log^2 2'}$$

Γ being the Gamma function and ζ a generalized Riemann Zeta function. Here, the sums run over some selected range of scales $[j_{min}, j_{max}]$, which is chosen a priori. The smallest scale j_{min} should be large enough for the asymptotic regime to be reached, while j_{max} is limited by the lack of coefficients at the coarsest scale. Using the relation $H = (\beta+1)/2$ the estimator of β at once gives an estimator for H.

In Ref.[83] it is shown that, if decorrelation actually were perfect, then the asymptotic variance of the estimator for β would achieve the Cramer Rao lower bound. Numerical

simulations exhibited in Ref.[83] show agreement with this approximate result. Moreover, the estimator is empirically shown to be robust with respect to some deviations from the Gaussian assumption. An approximate confidence interval for H is (see Ref.[4]):

$$\hat{H}_u - \sigma_{\hat{H}_u} z_\epsilon \leq \beta \leq \hat{H}_u + \sigma_{\hat{H}_u} z_\epsilon,$$

where z_ϵ is the $1-\epsilon$ quantile of the normal distribution and

$$\sigma^2_{\hat{H}_u} = \frac{2}{n_{j_{\min}} \log^2 2} \frac{1 - 2^{-J}}{1 - 2^{-(J+1)}(J^2+4) + 2^{-2J}},$$

$J = j_{\max} - j_{\min}$ is the number of scales which is used in the regression. This of course also gives a confidence interval for H.

We also performed a small simulation study. On 500 simulated paths of length 4096 of a fGn with $H = 0.8$ we computed the estimator of H using $j_l = 1$, $j_2 = 10$ and the Daubechies1 wavelet. The empirical means and standard deviations for the estimator were 0.797, and 0.038, where the latter should be compared with the value 0.012 obtained for the theoretical standard deviation.

An appealing feature is that the wavelet transform performs a "smooth differentiation" of the signal, with the degree of differentiation equal to the number of vanishing moments. Thus, as already mentioned, means and smooth trends are removed, and non stationary processes which have stationary increments of order N produce stationary wavelet coefficients (Ref.[14], see also Ref.[35] for the fractional increments). If such nonstationary processes exhibit a scaling of type (3.5), then the corresponding parameter can be estimated by the same procedure as before.

Suppose now X is self similar with Hurst parameter H, but not necessarily long range dependent. Using Eq. (3.3) above, it may be seen that then, also in cases with infinite variance, the wavelet method gives an estimator of H, when taking logarithms makes moments finite. In particular, for a stable Lévy motion, the method estimates $H = 1/\alpha$. This is discussed in detail in Ref.[2]. Thus, as already pointed out, an estimated value of β which is clearly different from 0.5 does not necessarily indicate LRD – it could also be caused by heavy tails. This comment also applies to, for example, the R/S statistic. See Ref.[80].

3.8. Inferring the Parameters of the Infinite Source Poisson Model from Aggregated Data

Assume that a trace of cumulative traffic has been observed during a time interval $[0, T]$ in a situation which is well described by the infinite node Poisson model with $1 < \alpha < 2$ and at a high aggregation level (i.e. with $\lambda T \bar{F}(T)$ large). Further, recall the notation from Section 2.5 and suppose the regularly varying factor in Eq. (1.1) is (at least approximately, and over the range of interest) equal to a constant, L.

By Proposition 1, the observed cumulative traffic then is approximately distributed as an expectation term $\mu(T) = \lambda \mu T$ added to a centered fBm with $H = (3-\alpha)/2$ and variance $\sigma^2(T) = \lambda T^{3-\alpha} L(4-\alpha)(2-\alpha)^{-1}(3-\alpha)^{-1}$. From the observations it is possible to find estimates of $\hat{\mu}(T), \hat{H}, \hat{\sigma}^2(T)$ of $\mu(T), H, \sigma^2(T)$ as described above. This

immediately leads to the estimate

$$\hat{\alpha} = 3 - 2\hat{H}.$$

Knowledge of the traffic rates alone does not contain enough information to make it possible to untangle λ, μ and L. We hence assume that we have more detailed information or experience from other data sets which allows a reasonable guess $\tilde{\mu}$ of the value of μ. The remaining parameters may then be estimates as

$$\hat{\lambda} = \hat{\mu}(T)/(\tilde{\mu}T)$$

and

$$\hat{L} = \hat{\sigma}^2(T)\hat{\lambda}^{\ 1}T^{\ (3\ \ \hat{\alpha})}(4-\hat{\alpha})^{\ 1}(2-\hat{\alpha})(3-\hat{\alpha}).$$

If we instead assume that the data are at a low aggregation level and the stable approximation holds, then the trace of the cumulative traffic is approximately distributed as the same mean term $\mu(T) = \lambda\mu T$ as above, plus $(\lambda T/L)^{1/\alpha}X_\alpha(\cdot)$. From the traffic rate data, using Nolan's software, we obtain estimates $\hat{\mu}(T)$, $\hat{\alpha}$, $\widehat{\lambda T/L)^{1/\alpha}}$ of the mean shape and scale. Proceeding in the same way as before this leads to the estimators

$$\hat{\lambda} = \hat{\mu}(T)/(\tilde{\mu}T)$$

and

$$\hat{L} = \widehat{(\lambda T/L)^{1/\alpha}})^{\ \hat{\alpha}}\hat{\lambda}^{\ 1}T^{\ 1}.$$

3.9. Simulation Methods

As discussed in the introduction, an important use of traffic models is to produce simulated traces, and such traces are also useful for testing estimation methods. For each of the three main models, stable noise, fGn, and the infinite source Poisson noise, simulated traces and the wavelet regression estimator for estimating the Hurst parameters are shown, in Figs. 3.1 3.3 .

Since stable noise simply consists of i.i.d. stable variables, simulation is straightforward. We used the built in simulator in the program package Splus. The parameter values were chosen as the estimated values for the UCB 10 s traffic rate trace, cf. Section 4.3.2 below. The maximum likelihood estimates which were obtained from the simulated trace were $\hat{\alpha} = 1.49 \pm .09$, $\hat{\beta} = 1.00 \pm .00$, $\hat{\sigma} = 90,000 \pm 600$, $\hat{\delta} = 446,000 \pm 1,100$, in reasonable agreement with the true parameters (given in the caption to Fig. 3.1). It may be noted that the wavelet estimate of the Hurst parameter (Fig. 3.1) is close to $1/\alpha = 0.66$ and hence is well away from 1/2 although the variables are independent, rather than long range dependent, c.f. the discussion at the end of Section 2.1.

The best available method to simulate fGn is imbedding in a circulant process. The method is described in Ref.[88], and we used software which is made available by G. Chan at http://www.maths.unsw.EDU.AU/~grace/.

The infinite source Poisson moel was simply simulated by building up traces from i.i.d. exponential starts of file transfers, and i.i.d. transfer times. For the present purpose of

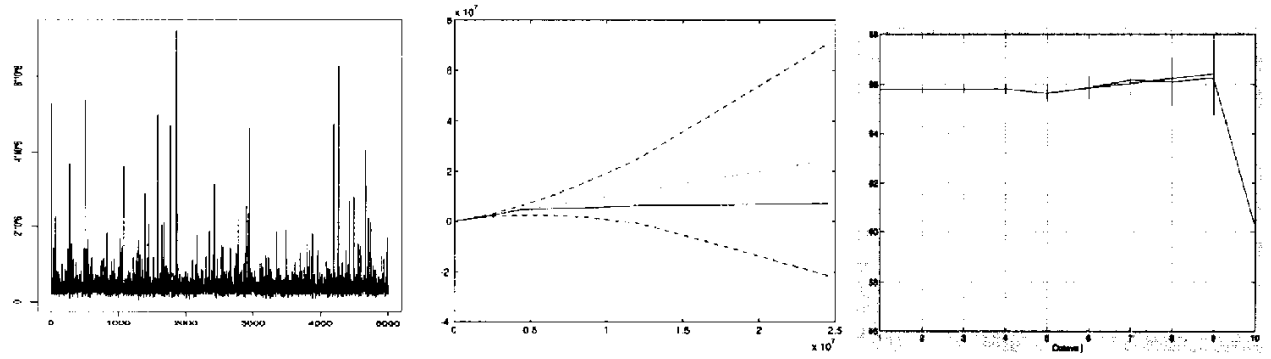


Figure 3.1. Synthetic trace from stable noise ($\alpha = 1.52$, $\beta = 1$, $\sigma = 86,000$, $\delta = 438,000$) used as the traffic rate: *left*) the traffic rate, *middle*) qq-plot (solid line) against fitted stable distribution (the dashed lines are the confidence interval), and *right*) wavelet regression estimation of the Hurst parameter, $\hat{H} = .60 \pm .06$.

model evaluation and illustration, we only used the simplest possible transfer time distributions, i.e. Pareto distributed variables with parameters roughly adjusted to match the measured traces. For use of simulation to aid in design, a more sophisticated choice would be desirable, at the very least means should be adjusted to have correct value (c.f. Section 3.8). For the purposes of Section 4, we call the resulting tace simM/G/∞.

4. DATA ANALYSIS

This section contains the statistical analysis of the traces we briefly introduced in Section 1. The plan is as follows: In the next subsection we give a quick account of our impressions of how well the models from Section 2 (see Table 2.1) fit the various data sets. The tables in subsection 4.2 summarize the results of the statistical analysis. The traces are then discussed one by one in separate subsections. This includes a more detailed description of the data, some further discussion of special issues for the separate traces, and conclusions

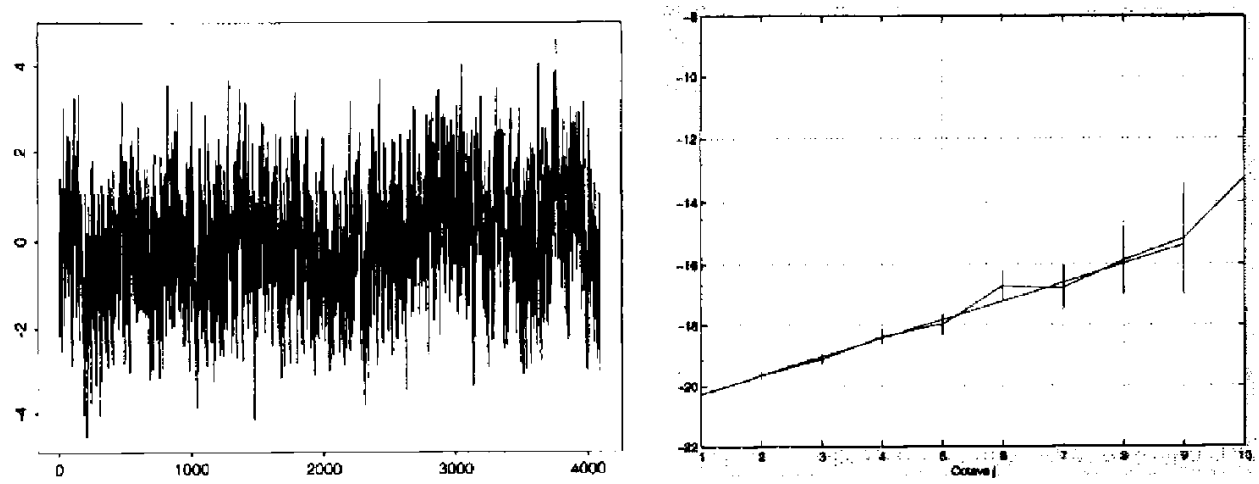


Figure 3.2. Synthetic trace from fGn, $H = .8$, used as the traffic rate: *left*) the traffic rate, and *right*) wavelet regression estimation of Hurst parameter, $\hat{H} = .81 \pm .01$.

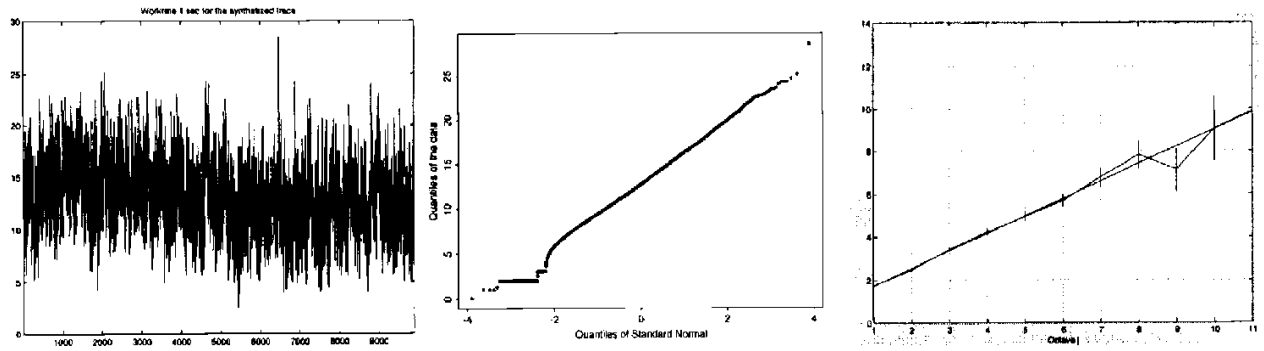


Figure 3.3. SimM/G/∞, Synthetic trace from the infinite source Poisson model, Poisson intensity 8, $\alpha = 1.2$ used as the traffic rate: *left*) traffic rate (resolution 1 second), *middle*) qqplot against normal distribution, and *right*) wavelet regression estimation of Hurst parameter, $\hat{H} = .90 \pm .01$.

trace by trace. Subsection 4.8 contains a rather extensive discussion of some of the very large numbers of issues, which arose in the analysis of these quite diverse data sets.

4.1. Does the Infinite Source Poisson Model Fit the Data

Table 4.1 lists features of the different theoretical models given in Table 2.1 and summarizes how strongly the data sets exhibit the model features. The discussion is extended in the subsections for the individual traces. The main conclusions are that the simulated trace simM/G/∞, the UCB sys 10s, the Eri syn 1s and the UCB 10s traces are well described by a suitable choice of the limiting model. However, the fact that the UCB data changes from model 2 to model 3 when the transfer rates are made constant (in the UCB sys 10s trace) indicates that the good fit of the Stable Levy motion to the UCB 10s data is not explained by the Section 2 theory. The remaining traces deviate strongly from all the models.

The second column of Table 4.1 shows an estimate of $(\lambda T \bar{F}(T)$ obtained by using λ from Table 4.2 , with T as 1 second for "BUburst 1s" and T as 10 seconds for "BUburst 10s" and "UCB 10s", etc, and $\bar{F}(T)$ estimated by #(observations $> T$)/# observations (except for simM/G/∞ where parameters were known). The next two columns contain subjective judgements of the fit of the marginal distributions to a Gaussian and a stable distribution. The family of stable distributions is quite rich, and even in the cases which are labeled "bad" in the "stable" column of the table, the visual discrepancies in the QQ plot between histograms and fitted densities were small. (In all cases the tail estimator in McNeil's EVIS program were larger than those in the mle fitted stable distributions obtained using Nolan's software.

The fifth column classifies the dependence in the traffic rate measurements. The entries are based on the estimated correlation function of the log traffic rates, using the standard 95% asymptotic confidence limits. Two scales are used. The first one is "str" if of the first 200 correlations at least 50 are $>.1$, "med" if between 20 and 50 are $>.1$, "sm" if less that 20 are $>.1$, and "tiny" means that all correlations are $<.1$. For the second scale "long" means that the area where most correlations are significantly different from 0 extend more than 200 lags, "int" that this area lasts for between 50 and 200 lags, and "sh"

Table 4.1. Traffic rates: summary of statistical analysis and model fit. Abbreviations are explained in Section 4.1.

Data set	$\lambda T\bar{F}(T)$	Gauss	Stable	Dep	Self-sim	Model	Model fit
SimM/G/∞	8	good	Bad	str/long	good	4	good
BUburst 10s	.09	med	good	str/long	bad	3	poor
BUburst 1s	.07	med	Bad	nonstat	—	3	poor
UCB 10s	16	Bad	good	indep	good	2	good
UCB syn 10 s	16	good	Bad	str/long	good	3	good
Munich lo TX	—	Bad	Bad	str/long	bad	—	—
Munich lo RX	—	Bad	Bad	nonstat	—	—	—
Munich hi .1s	—	Bad	Bad	med/int	bad	—	—
Munich hi .01s	—	Bad	Bad	sm/long	—	—	—
Ericsson	.5	Bad	Bad	nonstat	med	—	—
Eri sys 1s	.5	medium	Medium	gstr/long	good	3	medium

that the it is shorter than 50 lags. It should be noted that for different data sets, "lags" may correspond to rather different amounts of real time. The entry "indep" means the traffic rates were judged to be compatible with independence. The sixth column rates the distributional self similarity of the traces. It is based on a visual appreciation of qqplots of traces at different resolutions. The seventh column indicates the model (if any) from Table 2.1 most appropriate for the modelling of each trace. The last column gives our judgement on the quality of the fit evaluated as "poor", "intermediate" and "good" (the reasons for these valuations are presented in the detailed analysis of each trace). Finally, an entry " " means that the estimate was not available or that none of the models in Table 2.1 was appropriate.

Table 4.2. Traffic rates: summary of statistical analysis. Whenever possible, numerical results are given as point estimate ± "standard deviation." Abbreviations are explained in Section 4.1.

Data set	$\gamma(=1/\alpha)$	$\hat{H}$	H^*	$\hat{H}_o$
sim M/G/∞	.13 ± .03	.90 ± .01	.90	.88
BUburst 10s	.36 ± .13	.89 ± .02	.67 ± .01	.73
BUburst 1s	.17 ± .03	.81 ± .01	.67 ± .01	.87
UCB 10s	.05 ± .18	.58 ± .03	.62 ± .01	.65
UCB syn 10s	.60 ± .14	.95 ± .07	.62 ± .01	1.36
Munich lo TX	.09 ± .04	.89 ± .01	—	.86
Munich lo RX	.03 ± .05	.97 ± .01	—	.85
Munich h .1s	.17 ± .12	1.02 ± .03	—	.66
Munich h .01s	.10 ± .03	1.03 ± .04	—	.56
Ericsson	.47 ± .09	.88 ± .02	.86 ± .08	1.21
Eri syn 1s	.31 ± .12	1.48 ± .02	.86 ± .08	1.51

Table 4.3. Starts of transfers and file sizes: summary of statistical analysis. The columns "time," "file size" and "rate," show the shape parameter γ ($= 1/\alpha$) for the transfer times, file sizes, and transfer rates, Numerical results are given as point estimate $\pm$ "standard deviation." Abbreviations explained in Section 4.1.

Data set	Intens	Exp	Dep	Time	Dep	File size	Dep	Rate	Dep
BUburst	.14	no	Nonstat	.60 $\pm$ .02	tiny/sh	.69 $\pm$ .13	tiny/sh	1.01 $\pm$ .14	tiny/sh
UCB	6.72	yes	Tiny/long	.57 $\pm$ .02	tiny/int	.52 $\pm$ .02	tiny/sh	.79 $\pm$ .04	tiny/sh
Ericsson	1.39	no	sm/sh	.78 $\pm$ .16	tiny/sh	1.15 $\pm$ .18	sm/sh	—	sm/sh

4.2. Summary of the Results of the Data Analysis

Table 4.2 contains results for the traffic rate measurements. The left most column contains the shape parameter γ (related to the tail index α by $\alpha = 1/\gamma$) of the traffic rates estimated by maximum likelihood using a generalized Pareto model (Subsection 3.3). Standard deviations are calculated assuming independence. In all cases the top 5% of the observations were used, and the fit to a Generalized Pareto distribution, as judged by QQ plots was good (except Eri syn 1s). It might be worth recalling that a value of γ close to zero means light tails, that the variance is finite for $\gamma < 1/2$, and that the important case with finite means and infinite variance ($1 < \alpha < 2$) corresponds to $\gamma \in [1/2, 1)$.

After this come the Hurst exponent H with standard deviation, estimated by the wavelet method with the Daubechies wavelet and the number of vanishing moments and $j_{\min}, j_{\max}$ chosen to given good fit to the regression on scales (Subsection 3.6). We also calculated the Hurst exponent from the tails of the files sizes as $H^* = (3 - \alpha)/2$ with $\alpha = 1/\gamma$ taken from the file size column in Table 4.2 and with standard deviations obtained by the delta method. The following column shows the Hölder parameter H_o, estimated by the quadratic variation method (Subsection 3.7). Finally, an entry " " means that the estimate was not available.

Table 4.2 only presents the data sets which contained information about individual file transfer. In the first column, the intensity λ of starts of transfers is estimated as the total number of transfers divided by total in seconds, and in the next the fit of the interarrival times to an exponential distribution as checked by qq plots is given. In the remaining columns, the shape parameter γ is estimated and dependence is checked in the same way as for Table 4.2. The column marked "time" refers to the download time of a file and "filesize" refers to the size of the downloaded file. In the column marked rate, we computed the transmission rate as

$$\text{(file size)/(time to download the file)}$$

and then estimated the shape parameter γ. The parameter γ for the rate in the Ericsson trace could not be reasonably estimated; see the data description below.

In the following, the file sizes and traffic load will be implicitly given in bytes, and the traffic rate corresponds to the cumulative traffic load in a given time interval. For

the Munich data, the traffic is measured in number of cells, which is proportional to the number of bytes (1 cell is 53 bytes).

4.3. The BU Traces

4.3.1. Data Description

This data contains four components and are described and analyzed in Refs.[19,22]. Below we describe the BU and the derived BUburst traces, which were obtained from the posted data. Both contain times of requests, file sizes, transfer rates and transfer times, before and after "burstification" for a period of approximately 28,000 seconds (8 hours and 20 min) with the most intense traffic in the February 1995 part of the trace. The construction of the traffic rate data sets *BUburst 1s* and *Buburst 10s* is explained below. The calculation of traffic rates assumes that the transmission rate did not change during the transmission of a file, and hence the constructed traffic rates deviate from the real ones.

The traffic was generated from two rooms of users during the period October 1994 to February 1995. The statistical characteristics of data vary considerably from one month to the next. We have only used data from the room containing 32 work stations used by undergraduates, with all the cache file requests removed. This trace is the most complete among the public domain data sets we are aware of. It was recorded at a time when Mosaic was the most common browser. Unlike Netscape, the source code of Mosaic is publicly available and could be altered for measurements purposes. A followup study by a BU team is reported in Ref.[10].

The data consists of the record of all the individual sessions generated by the different users. A session is a succession of URL requests (http, ftp, gopher,...) made by one user from logging in until logout. Every request corresponds to a line in the session file which contains the following information: machine number, starting time of the request (in micro seconds since January 1, 1970), URL of the requested document, size of the document (in bytes) and transfer time (in microseconds). The cache files, that is files already stored on the user disk, are marked with a zero transmission time and file size.

The *BUburst* data resulted from the need to distinguish between file requests made by humans and machine generated requests. Typically a human initiated request, e.g. for a web page, triggers a cascade of file transfers, and hence very small intervals between file transfers usually are machine generated. If we think of users as being sources in the infinite source Poission model, we must correct for this cascading. Hence we lumped together requests which arrived less than.5 seconds apart into a single "request" which we will refer to as a "burst". The selection of the "threshold".5 seconds was based on a close look at the data. The size of the burst is then the sum of the sizes of the files lumped together, and the duration is from the beginning of the first file transmission from the BUburst data by calculating traffic rates at 1s and 10s resolutions respectively. A typical session before and after creating the bursts is presented below. The 100682 initial requests resulted in 56516 bursts, out of which the period we studied contains 4161.

BU				Buburst		
Begin download	URL	File size	Transfer time	Begin download	File size	Transfer time
43817.159177	"http://cs www.bu.edu/"	2069	2.994023	43817.159177	4591	4.903682
43820.586374	"http://cs www.bu.edu/lib/pics/bu logo.gif"	1805	0.551812			
43821.538385	"http://cs www.bu.edu/lib/pics/bu loabel. gif"	717	0.524474			
43835.427403	"http://cs www.bu.edu/courses/Home.html"	3382	0.510081	43835.427403	3382	0.948917
43836.36075	"http://cs www.bu.edu/lib/pics/bu logo.gif"	0	0.0			
43836.37632	"http://cs www.bu.edu/lib/pics/bu label.gif"	0	0.0			
43842.134286	"http://cs www.bu.edu/students/grads/tahir/ CS111/"	1065	0.404507	43842.134286	1065	0.404507
43846.456855	"http://cs www.bu.edu/students/grads/tahir/ CS111/hw6.ps"	32246	0.498012	43846.456855	32246	0.498012

4.3.2. The Number of Logged on Work Stations

Due to the diurnal cycle. (strongly visible because of the small number of users), the traffic is highly non stationary, see Fig. 4.1 which shows the number of active sessions. The periods with 0 logged on stations typically occur during nights and week ends.

4.3.3. Discussion

The interarrival times for the BU and BUbust traces over the entire month were clearly not exponentially distributed. For BUburst there was some indication of exponential like behavior for small interarrival times, but the distribution was heavy tailed as revealed by a QQ plot (see Section 3.3) yielding the estimate $\hat{\alpha} = 1.54$, cf Fig. 4.2 . In view of the low aggregation level this is not surprising.

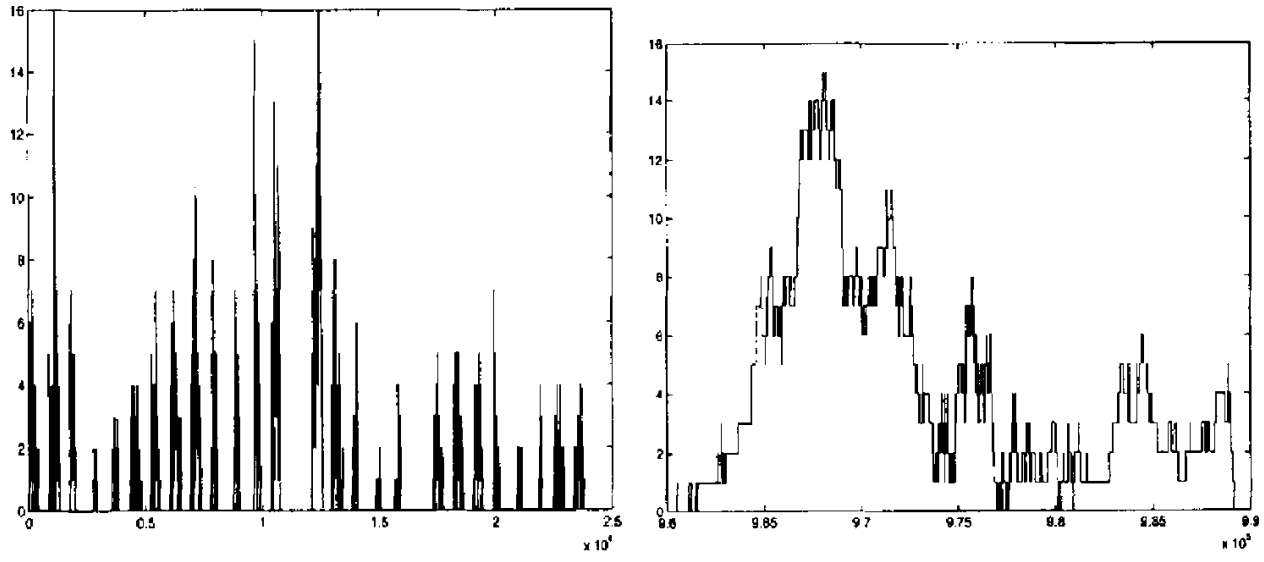


Figure 4.1. BUburst: *left*) Number of active sessions versus time (in seconds), January 1995. *right*) Detail of the left plot showing one day.

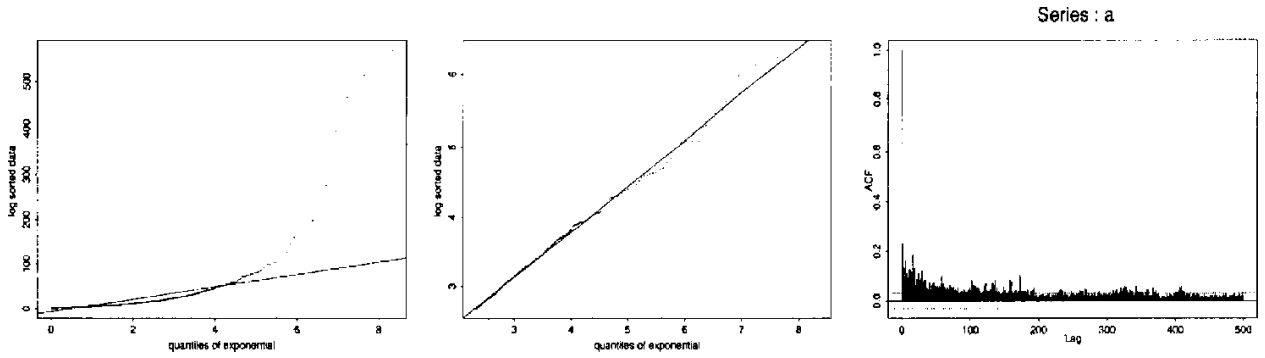


Figure 4.2. BUburst. Inter-arrival times of bursts: *left*) qq-plot against exponential distribution, *middle*) Pareto fit for the 400 largest values (last 10 %), ($\alpha = 1.54$), *right*) autocorrelation function.

For the BUburst trace the transmission times, file sizes, and transfer rates had heavy tails with $\hat{\gamma} > .5$ The tails of the transmission rates dominated the tails of the transmission times and file sizes, in contrast to our expectations.

The BUburst 1s and BUburst 10s traffic rate data sets have rather light tailed marginal distributions. However, the distributions still look much more stable than normal. Since the traffic rates are quite dependent (or non stationary), they cannot be well modeled by a stable Lévy noise.

The Hurst and Hölder estimates were rather similar, while the Hurst estimate derived from the fBm model and the $\hat{\gamma}$ for the transfer times was smaller.

We selected a short period with high traffic. Realistic models should include the variation in the number of logged on workstations.

An extensive statistical analysis of these data has been carried out by the authors of the trace.[17] In particular they present similar estimates for tails and traffic rates and explain the discrepancy with the theory through the low traffic level.

From Fig. 4.2 one sees that left tail (near 0) of the inter arrival times looks like an exponential or Weibull tail while the right looks Pareto. The autocorrelations of the interarrival times seem significant but rather small. Figures. 4.3 and 4.4 show the marginal distributions of the traffic rates and the Hurst parameter estimations. For heavy tailed data it is difficult to plot an information histogram so we show the histogram of the long the data.

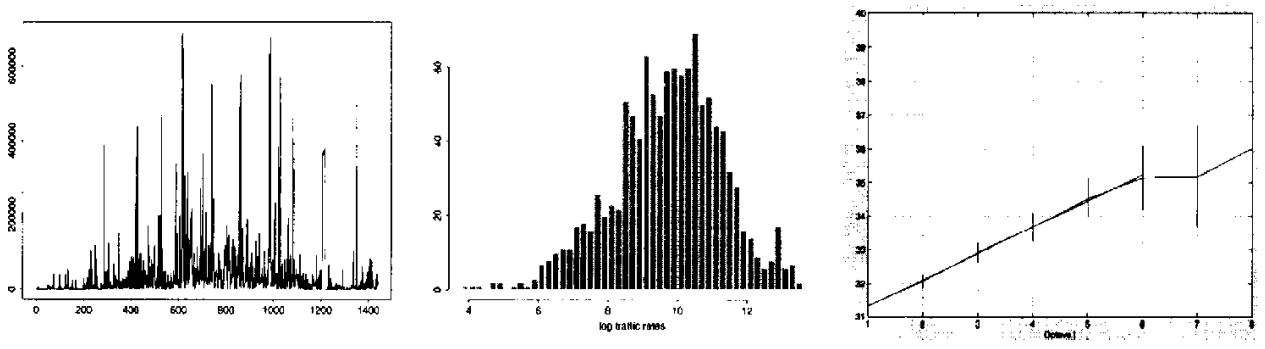


Figure 4.3. BUburst 10s. Traffic rates: *left*) trace, *middle*) histogram of log traffic rates; *right*) wavelet regression estimation of the Hurst parameter $H = .89 \pm .02$.

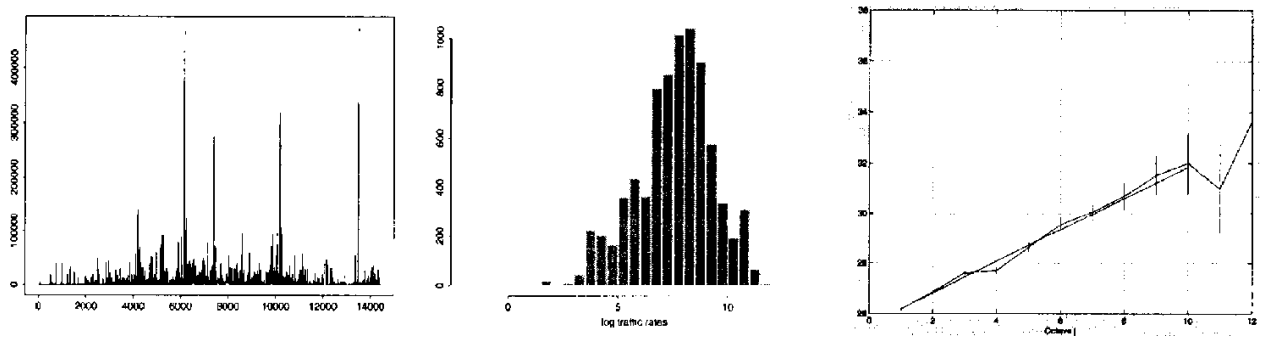


Figure 4.4. BUburst 1s. Traffic rates: *left*) trace, *middle*) histogram of log traffic rates; *right*) wavelet regression estimation of the Hurst parameter, $H = 0.81 \pm .01$.

Finally, this data set is rather old on the timescale of the Internet evolution, and in particular the paratmeter values obtained from it may be rather different from current ones.

We conclude by discussing the extent to which feature of the limiting infinite source Poisson models (Table 2.1) capture the behavior of the traffic rates. Recall that the tail index of file sizes and transfer times are in the large (1,2) and there seems to be long range dependence (or non stationarity) and also self similarity. This suggests Model 3. However, the tails of the traffic rates are heavier than normal, and overall the normal distribution does not fit the data. Hence model fit is poor.

4.4. The UCB Traces

4.4.1. Data Description

The UCB data is an 18 day trace of the home IP HTTP traffic processed by UC Berkeley during November 1996. It consists of the Internet connections through the server of the university established by individual dial up users. The traces together with a detailed description of them is available at http://ita.ee.lbl.gov/html/contrib/UCB.home IP HTTP.html.

The data content is similar to the BU traces (initiation time of a file transfer, file size, transfer times of a request IP address of client). We have not attempted to burstify the UCB data, as we did with the BU data. However, a much larger number of users are included and the traffic rate is higher. Due to the non stationarity and the diurnal cycle, we chose to restrict the analysis to several hours of peak traffic on a weekday, i.e. the period 5 8 p.m. on Thursday November 7. This part of the trace consists of about 80,000 requests. We would like to emphasize the need to carefully select the period for analysis, since some network outages occurred during the recording of the trace. E.g. one outage can be observed on November 6, from 5 pm to 8 pm.

As for the BU data, the actual traffic rates were not available and we again constructed approximate traffic rates by assuming constant (but different from transmission to transmission) transmission times.

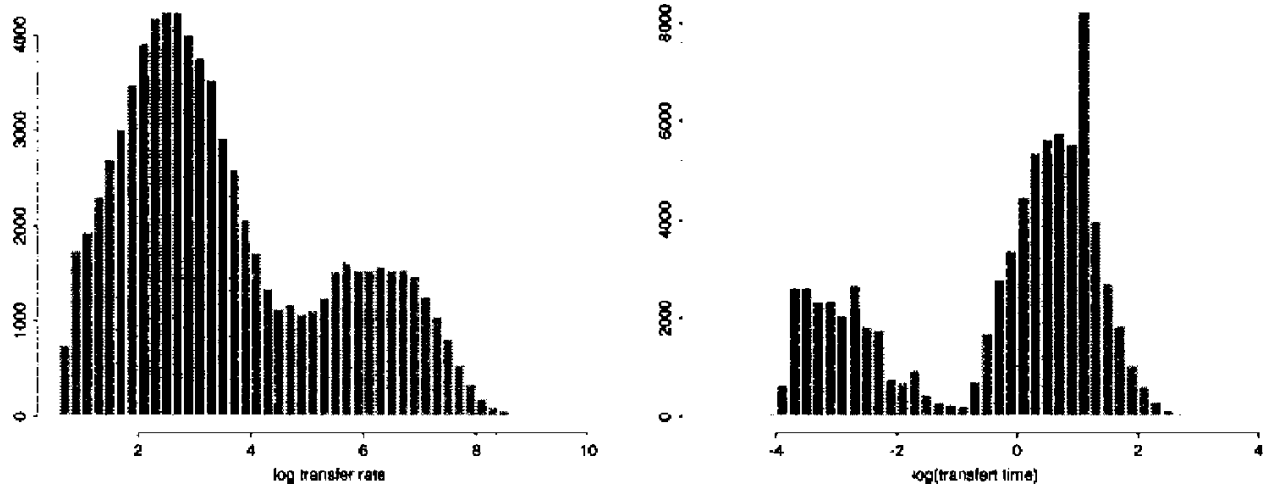


Figure 4.5. UCB Histograms of *left*) log transfer rates and *right*) log transfer time.

4.4.2. Discussion

An exponential distribution fits the inter arrival times fairly well. While rather many of the estimated correlations for the long interarrival times are significantly different from zero, they are quite small (and perhaps caused by a small nonstationarity) and independent interarrival times seems a reasonable approximation (Fig. 4.6).

The file size and transmission time distributions were close to the borderline between finite variance and infinite variance while the transmission rate was considerably more heavy tailed. The distribution of long transfer rate is clearly bi modal (Fig. 4.5, *left*), perhaps due to different modern sppeds. A natural consequence is that the distribution of long transfer times is also bi modal (Fig. 4.5, *right*). There seems to be no hope of finding simple parametric forms which accommodate these distributions.

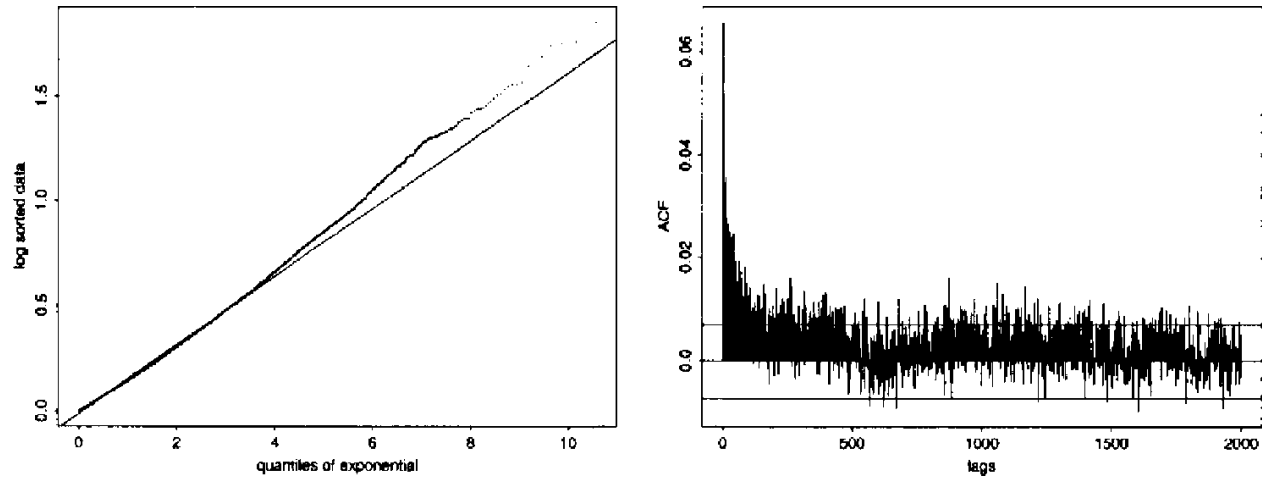


Figure 4.6. UCB inter-arrival times: *left*) qqplot against exponential distribution, *right*) autocorrelation function of interarrival times.

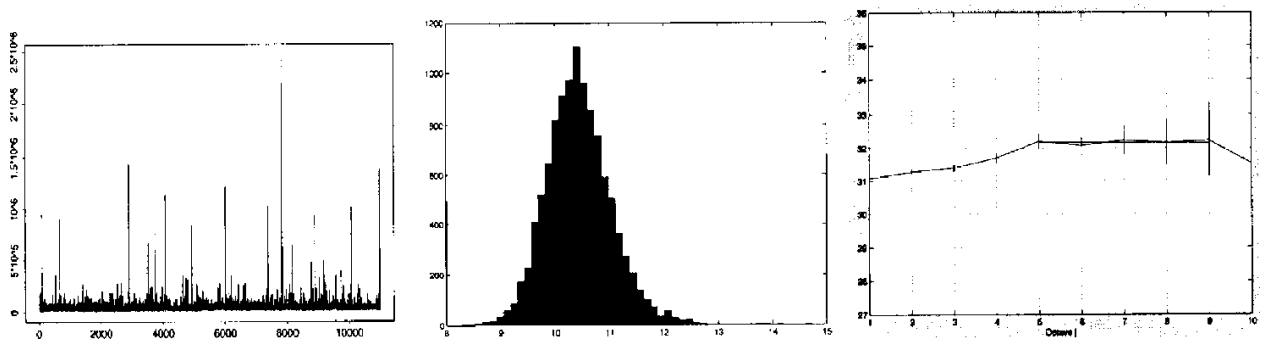


Figure 4.7. UCB, traffic rate 1 second blocks, ($\hat{\gamma}_T = 43 \pm .06$) *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent, $\hat{H} = 0.50 \pm .04$.

The UCB 10s traffic rates show a quite good fit to the stable Lévy noise model, actually the qq plot and density estimates were quite similar to those for the simulated stable noise. The ML estimates of the marginal parameters of an a stable distribution were $\alpha = 1.52 \pm .02$, $\beta = 1.00 \pm .00$, $\sigma = 86000 \pm 1200$, $\delta = 438000 \pm 2300$. Note that the value of the tail index α is very different from the value $\alpha_T = 1/.05 = 20$ estimated using the semi parametric approach described in section 3.3.

The stable Lévy motion model fitted somewhat less well for traffic rates computed using shorter time interval. This aspect of the data agrees with the analysis in Section 2.5, Proposition 1. The effect of varying block size on the marginal distribution of the traffic rates is illustrated in Fig. 4.11.

The distributional shape is similar for different sizes of the time intervals and hence indicates distributional self similarity. Figures 4.7 4.9 show the wavelet regression estimation of the Hurst parameters to be close to the independent increments Brownian motion value 0.5.

To what extent do the features of the limiting infinite source Poisson model (Table 2.1) capture the empirical behavior of the traffic rates? For the UCB 10s data, the tail index of file sizes and transfer times are in the range (1,2), the traffic rates seem independent and self similar, and the stable distribution fits the data well. Thus, Model 2 seems to describe the data well. However, $\lambda T \bar{F}(t)$ takes the rather large value 16 (Fig. 4.10).

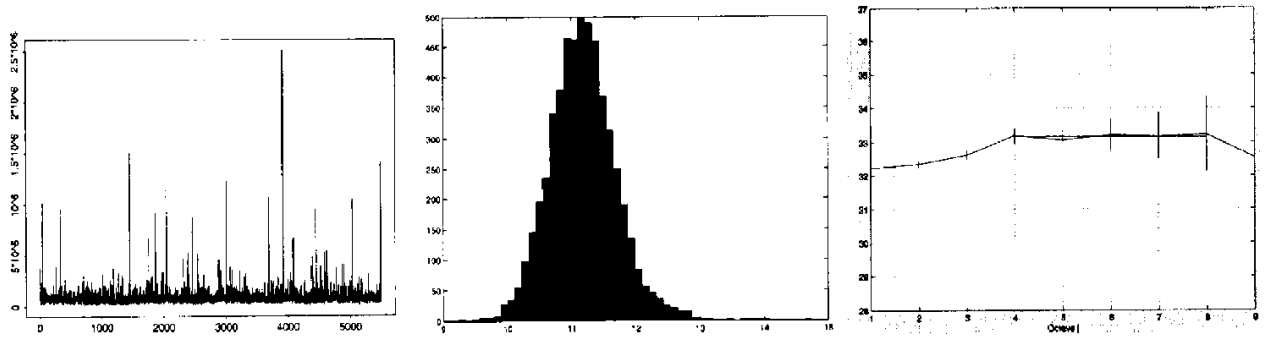


Figure 4.8. UCB, traffic rate 2 second blocks, ($\hat{\gamma}_T = .05 \pm .17$) *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent, $\hat{H} = 0.50 \pm .01$.

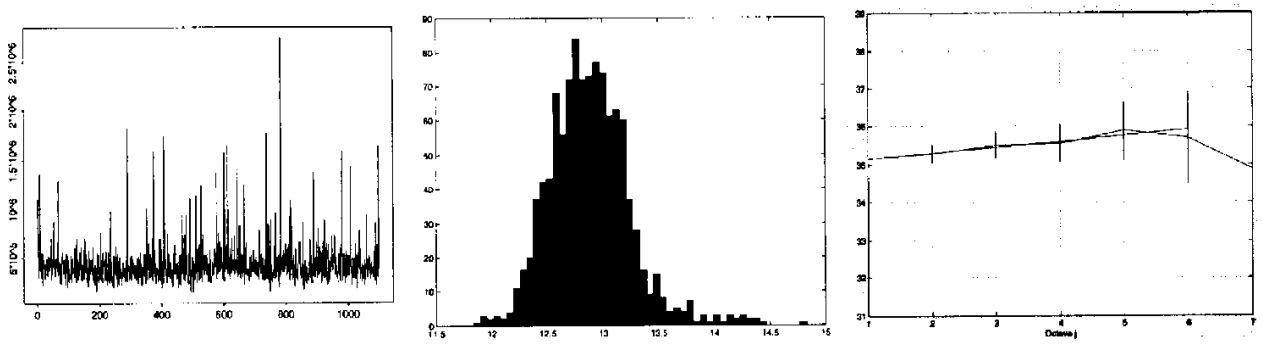


Figure 4.9. UCB, traffic rate 10 seconds blocks, ($\hat{\gamma}_T = .05 \pm .17$) *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent, $\hat{H} = 0.58 \pm .03$.

4.5. The Munich Traces

4.5.1. Data Description

The *Munich lo* data set contains measurements of cell rates for both the sending (TX) and receiving (RX) directions of an ATM link. The link, a Customer Service Switch (GDC APEX 200) with a line speed of 155 Mbit/connects the Munich University network with the German Scientific Broadband Network. The data, kindly provided to us by Helmuth Gogl, was collected round the clock on Wednesday, November 12, 1997 (TX) and Wednesday, December 17, 1997 (RX) with a temporal resolution of 2 seconds, *i.e.* the total number of cells (1 cell is 424 bits) that passed the ATM link every 2 seconds was recorded. The maximum bandwidth which was available was about 20 Mbit/s.

The traffic recorded was pure IP (mainly HTTP, FTP, and NNTP) data traffic, without any audio/video components. A shorter sample covering the period between 10 a.m. until

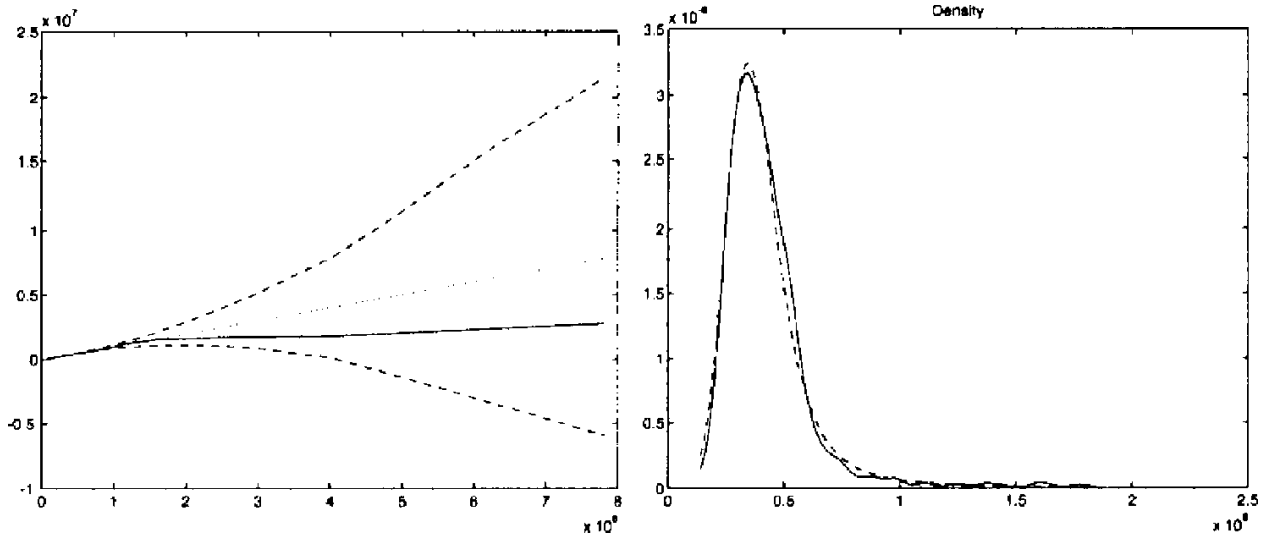


Figure 4.10. UCB 10s traffic rate: *left*) qqplot (solid line) against estimated stable distribution (dashed lines are the confidence interval), and *right*) density estimate and density of estimated stable distribution $\alpha = 1.52 \pm .02$, $\beta = 1.00 \pm .00$, $\sigma = 86000 \pm 1200$, $\delta = 438000 \pm 2300$.

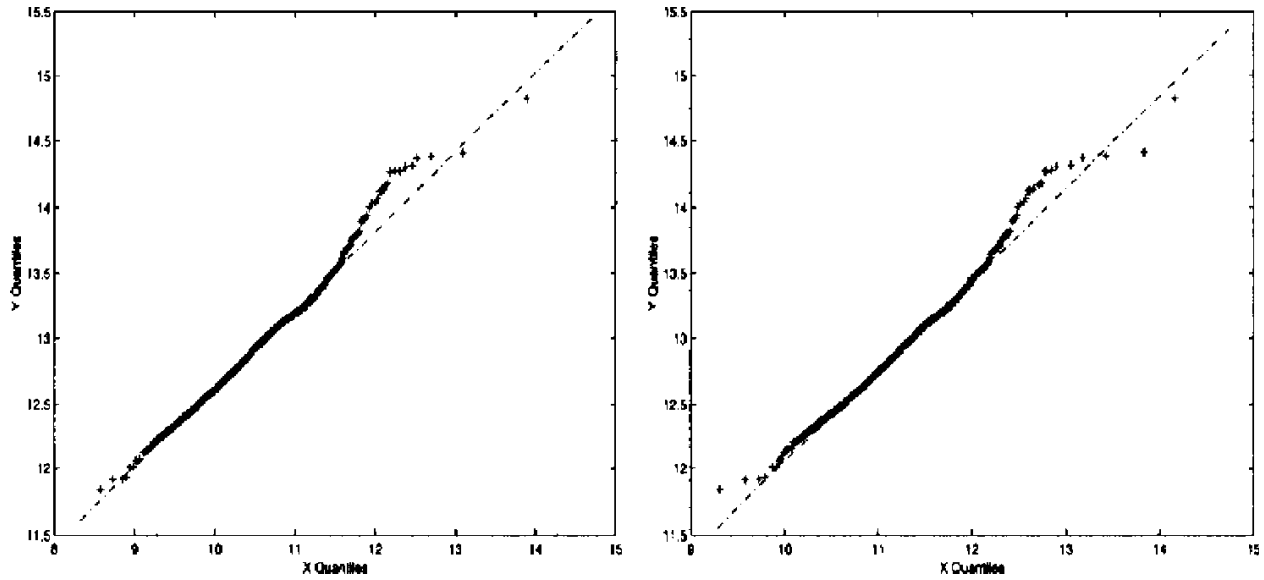


Figure 4.11. UCB: comparison of the distribution of traffic rate over time blocks of varying lengths: *left*) qqplot of log traffic rate for 1 second blocks against 10 second blocks and *right*) qqplot of log traffic rate for 2 second blocks against 10 second blocks.

1 p.m. was selected for analysis. The shorter period was chosen to obtain a roughly stationary data set. However, for the RX direction, the influence of the lunch hour is still clearly visible.

The *Munich hi* data set was recorded at the recorded at the same ATM link, but with a much higher time resolution. It contains 1690729 inter arrival times of ATM cells from all traffic in the TX direction over a period of approximatively 137 seconds with a resolution of 1 micro second. The measurement was done on Tuesday, December 23, 1997 starting at 14:48:15 and ending at 14:50:33. A mean utilization of 5.062 Mbit/s (11938 cells/s) was recorded, which corresponds to about 25% utilization (given the 20 Mbit/s maximum bandwidth). A detailed description of the data is given in Ref.[32].

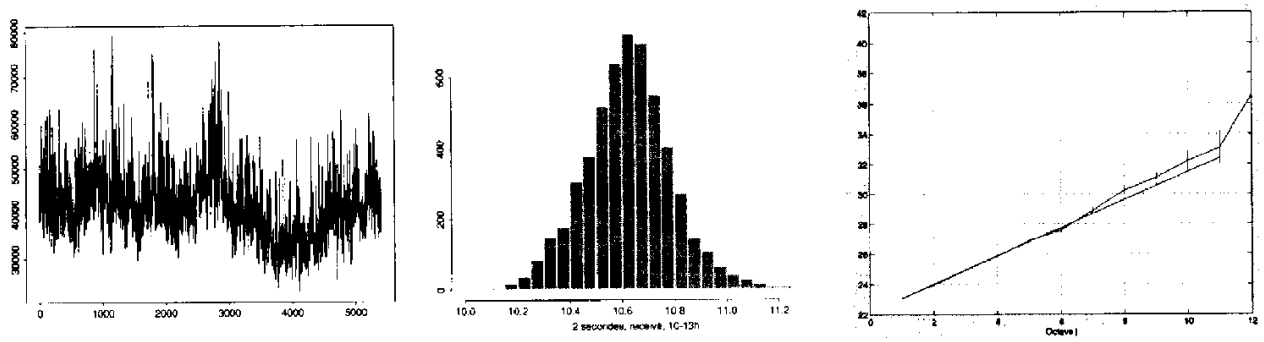


Figure 4.12. Munich lo, RX: *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent $H = .97 \pm .01$.

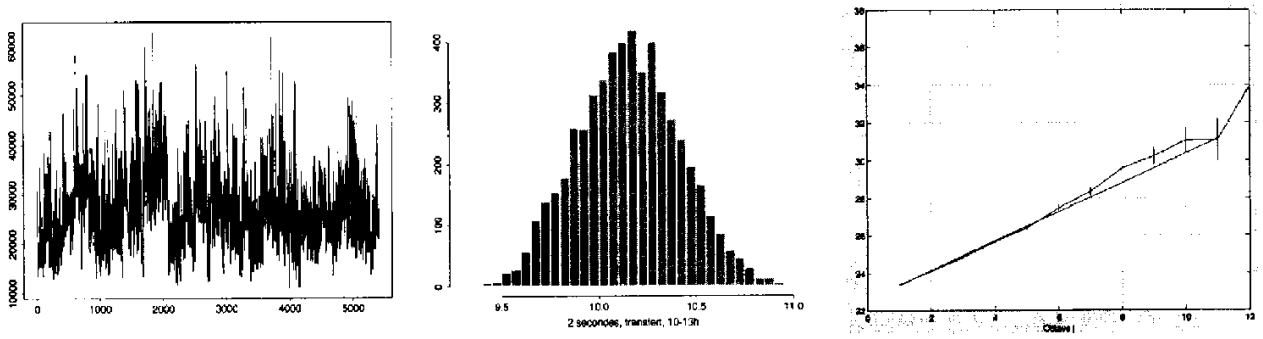


Figure 4.13. Munich lo, TX: *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent $H = .89 \pm .01$.

4.5.2. Discussion

All four Munich traces come from the same server, and similar traffic situations. The RX cells had typically passed through and been reassembled into frames in more routers than the TX traces. The temporal resolution varied by a factor 200, from the high aggregation level of 2 seconds down to a low one of .01 second. (We did not consider finer time resolutions in order to keep a certain aggregation level). The different kinds of traffic, HTTP, FTP, and NNTP, had rather different characteristics, as exhibited in figures in Ref.[32].

The traffic rate traces have similar light tails with γ_T close to zero. The distributions are neither normal nor stable, with heavier tails then for normals but not as heavy as for a stable distribution; cf. Table 4.2 and Figs. 4.12 4.15 . In fact, lognormal QQ plots showed rather good fit. The marginal distributions had different shapes for the different traces.

There is a clear and strong long range dependence, which persists over a wide range of scales as seen by autocorrelation functions and wavelet regression plots and also by the fact that for all four traces the estimates of the Hurst exponent are close to the same value, .95. In addition, the Munich lo RX trace was clearly nonstationary. However, based on theoretical considerations and numerical experiments, this is not expected to influence the Hurst and Hölder estimates significantly.

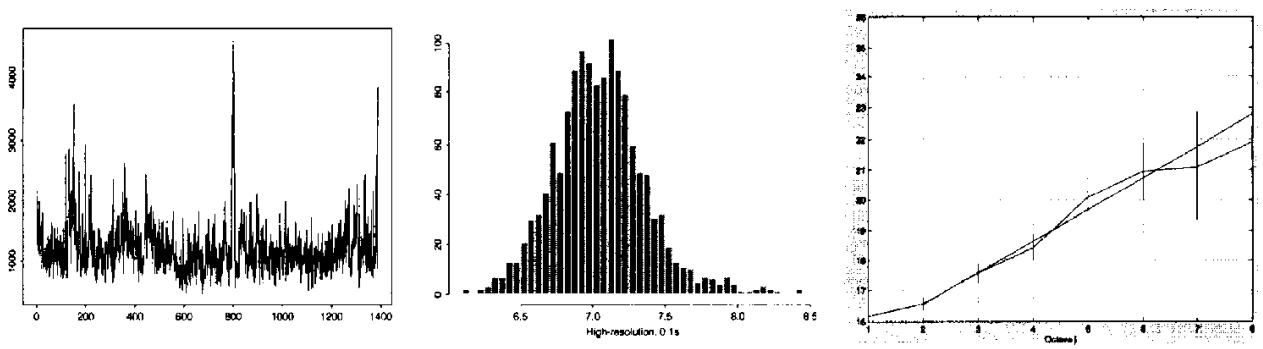


Figure 4.14. Munich hi .1s: *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent $H = 1.02 \pm .03$.

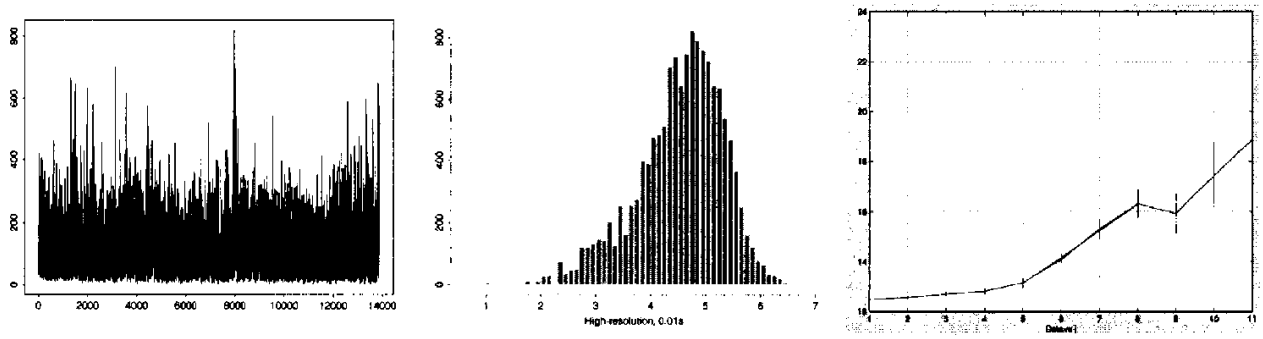


Figure 4.15. Munich hi .01s: *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent $H = .1.01 \pm .04$.

The Hölder exponents throughout are smaller than the Hurst ones, and decrease from values of around .85 corresponding to strong local dependence for the high aggregation levels down to .56 which is close to independent increments for the.01 second trace.

In the fBm model, the Hurst and Hölder exponenets are the same. Here, this is clearly not the case. One explanation could be that although the tails of the marginal distributions are lighter than Pareto, they still are rather heavy, which may inflate the Hurst exponents (recall that the i.i.d. stable variables considered in this paper have Hurst exponents $> 1/2$). However, what is not explained by the infinite source Poisson model is that the data seem more independent on a finer time resolution. A possible explanation of this discrepancy is the slow variation of λ due to a varying number of users.

An extensive analysis of these data at the cell level, and after collection into "bursts," is made in Refs.[32,33]. Both on the cell and the burst level, estimated Hurst exponents were smaller than our estimates since this analysis concerned short time distances it is in agreement with the finding that the Hölder exponents were smaller than the Hurst ones. An expected result was that the lengths of very short interarrival times between cells ($< 10\ \mu s$) were not exponential, but had a discrete distribution determined by the system clock frequency.

None of the models from Table 2.1 seem to fit the data well, and, in particular, the traffic rates are not even distributionally self similar.

4.6. The Ericsson Trace

4.6.1. Data Description

The *Ericsson* trace consist of time stamps of starts and completions of the TCP connections that correspond to HTTP file transfers to and from a corporate WWW server which holds home pages and information primarily directed to about 2000 company users. In addition, people at other Ericsson companies around the world have access to the server. The recording was started Thursday Oct 15, 1998, at 15:20 and was ended Friday Oct 16 at

15:49. The information extracted from the data gives the times of connection starts, connection durations, number of bytes transferred (from server to user as well as the opposite direction), and client identification for each connection. Approximately 2% of the connections resulted in missing data for server to client transfers. The reason was that the connection was abnormally terminated and the number of bytes transferred could not be obtained from the data saved in the trace. The data set is quite nonstationary, and hence a more stationary subset covering 2000 seconds was chosen for analysis. As for the BU and UCB traces approximate traffic rates were constructed from the data described above.

For unexplained reasons the measurement system erroneously added 30 milliseconds to about 18% of the transfer times.

4.6.2. Discussion

There were both stationarity and measurement problems with this data set, but it was also the most recent one. It had the most heavy tailed transfer times and the file sizes. However, the transfer rates were light tailed but we have not included the estimated value of γ in Table 4.2. This interarrival times were close to independent, but their distribution had heavier than exponential tails, as expected for this relatively low aggregation level.

The two Hurst estimates were similar, while the Hölder estimate was in the differentiable region.

The traffic rates were light tailed and the wavelet regression showed good agreement with the fBm model, but the marginal distributions were clearly nonnormal.

Because of the measurement and non stationarity problems, conclusions from this data set are necessarily tenuous. We still briefly discuss how features of the limiting infinite source Poisson models (Table 2.1) describe the traffic rates. The tail index of file sizes and transfer times are in the range (0,1) and (1, 2), respectively. Traffic rates exhibit evidence of non stationarity but ignoring this, neither the Gaussian nor the stable distribution gives a good fit. Hence none of the models seem to apply to the data.

4.7. The Synthetic UCB and Ericsson Traces

4.7.1. Data Description

The *UCB syn 10s* and *Eri syn 1s* traces were constructed from the same segments of the UCB and Ericsson data sets which were used in the analysis of work rates in Sections 4.6 and 4.4. The new traces were constructed as follows: File transfer started and ended as they did in the real trace, but the transfers were changed to make transfer rates equal to one. From the resulting "data sets" we then calculated 10 and 1 second traffic rates, respectively to obtain he "UCB syn 10s" and "Eri syn 1s" traces.

The infinite source Poisson model assumes that transfer rates are constant, and we hence thought it interesting to compare a situation where this indeed was the case with the real traces where the transfer rates had a heavy tailed distribution.

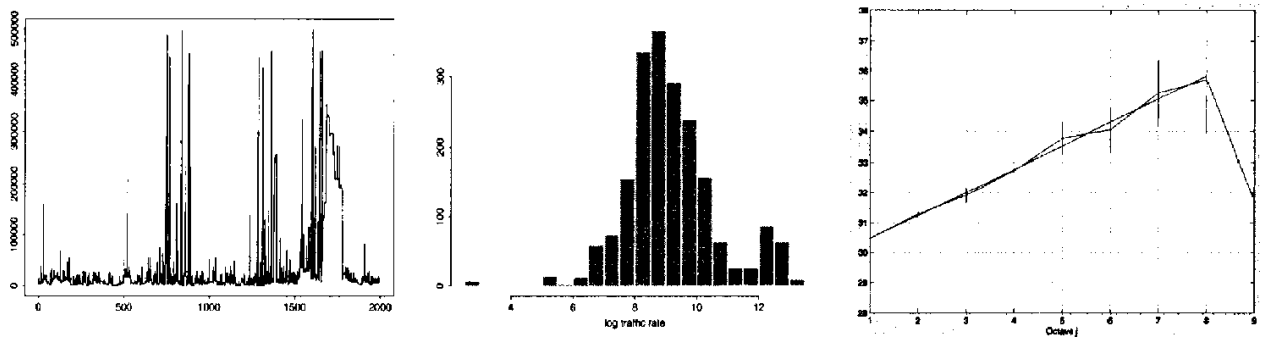


Figure 4.16. Ericsson traffic rates: *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of Hurst exponent $H = .88 \pm .02$.

4.7.2. Discussion

The UCB syn 10s traffic rate conformed with a fBm expect in that the Hölder estimate was in the differentiable region, $H_o \geq 1$. For the Eri syn 1s the marginal distribution were close to normal except for a few observations in the tails. The Hurst and Hölder estimates were both close to 1.5. We believe this is a consequence of non stationarities in the data, which clearly show up in Figs. 4.16 4.18 . Both synthetic traces were very different from the real ones.

As shown in Fig. 4.19 , large values of transfer times and transfer rates were fairly independent; the scatter plot of transfer times vs. transfer rates hugs the two axes which shows that either the two quantities are independent or if dependent, their joint distribution is multivariate regularly varying satisfying a condition of asymptotic independence (Ref.[67], page 290ff).

To conclude, Model 3 from Table 2.1 fits the UCBsyn trace well and this model also seems reasonable for the Eri syn 1s trace, although the is some discrepancy in the tail of the marginal distribution.

4.8. Summary and Conclusions

We have provided a summary in Section 2 of the description and properties of a fairly flexible model which we have called the *infinite source Poisson model*. This model

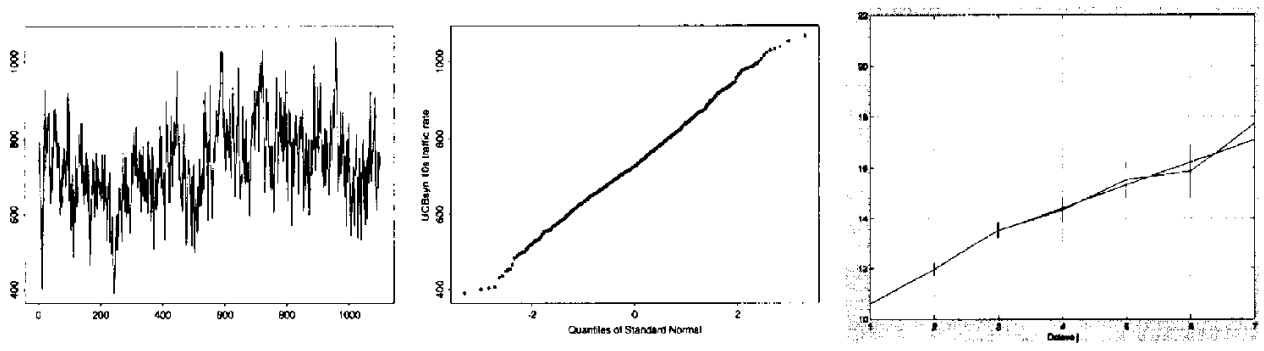


Figure 4.17. UCB syn 10 s traffic rates: *left*) trace, *middle*) qqplot vs. normal variable, *right*) wavelet regression estimation of the Hurst parameter $H = .95 \pm .07$.

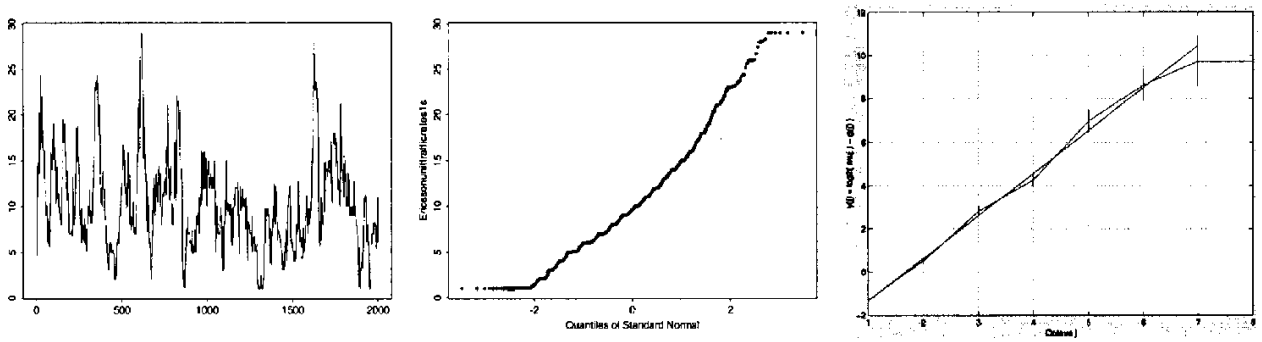


Figure 4.18. Eri syn 1 s traffic rates *left*) trace, *middle*) qqplot vs. normal variable, *right*) regression estimation of the Hurst parameter $H = 1.48 \pm .02$.

predicts that on large times scales (i.e. as $T \to \infty$), traffic will have certain properties and we catalogued these properties according to whether the regular variation index α of the connection length distribution F satisfied (i) $\alpha < 1$, (ii) $1 < \alpha < 2$ or (iii) $\alpha < 2$. We also surveyed in Section 3 statistical methods for estimating model parameters in order that we could diagnose whether our model provided a good fit to four data sets.

Our assessment of the statistical methods selected is that they were convenient to apply and aided understanding of the data. However, we have downplayed difficulties in estimating tail parameters such as α or γ. Such estimates are sensitive to either the choice of threshold or choice of number of upper order statistics as has been amply documented; see, for example, Refs.[30,68,71]. Similar choices have to be made in the estimation of Hurst parameters. More development is needed to provide reliable confidence intervals for estimates, especially of long range dependence parameters and Hölder exponents. Better understanding of the estimation of Hölder exponents would be useful as well as a clearer understanding of the relationship between treatments using the second order definition (2.3) and the pathwise treatments in, for example Ref.[75].

Global statistical properties such as heavy tails and long range dependence were amply in evidence in our data as expected and as predicted by the model. Transfer times, file sizes and transfer rates were consistently heavy tailed, usually with $1 < \alpha < 2$. (See Table 4.2 and 4.2) Traffic rates frequently displayed evidence of long range dependence as shown in Table 4.1.

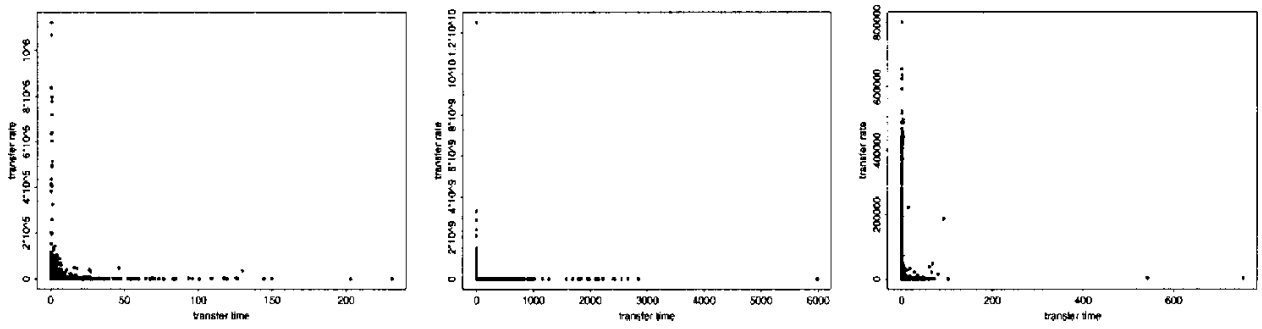


Figure 4.19. Plot of transfer times against transfer rates for *left*) the BUburst trace, *middle*) the UCB trace, and *right*) and the Ericsson trace.

However, the estimated marginal distributions of traffic rates changed markedly with aggregation level for all of the traces except for the UCB and synthetic data and thus we conclude that almost all the real traffic rates were not distributionally self similar. Either the mode does not adequately fit the data or considered time scales are too small for the asymptotic behavior discussed in Section 2.2 to hold. Our maximum aggregation level was 10 seconds. Exceeding this level, presumably, would have improved the fit of the asymptotic models but was judged to be less useful from an applied point of view.

The scaling behavior of the actual traffic rate traces as summarized by the wavelet estimator of the Hurst parameter was compatible both with the fGn model and with the stable noise model (see the column for $\hat{H}$ in Table 4.2). The fGn and stable noise model are the asymptotic limits given in Section 2.1 when $1 < \alpha < 2$. However, all measured marginal distributions were far from normal, and hence the fGn model does not fit the data. The stable distributions are much more flexible and generally fitted the data better. In one case, the UCB 10s traffic rates, the stable noise model provided a good description of the trace (Fig. 4.10). However, generally speaking, estimates of α given in Table 4.2 (remember $\alpha = 1/\gamma$) using the maximum likelihood estimation in the generalized Pareto distribution model produced lighter tails than using Nolan's maximum likelihood estimators[66] to fit stable distributions directly to the data, so there is doubt that the stable model adequately fits the tails of the data.

An overall impression is that the infinite source Poisson model struggles to adequately describe our data. The assumption of constant transfer rates in the model is at the center of the problem. This is clearly shown by the UCB syn and Eri syn traces where the real transfer rates were changed to a fixed rate (= 1). In particular, the UCB trace conformed fairly well with the limiting fBm model in the properties studied by us (Table 4.1 and Section 4.7). This was in complete contrast to the actual measured traces which had widely varying transfer rates and were quite far from being a fBM (Sections 4.4 and 4.6). Examining file transfer rates in, for example, the BU data shows that such rates should more realistically be modeled as random with a heavy tailed distribution. Furthermore, file transfer rates seem to be only asymptotically independent of file transfer times (Fig. 4.19). Non independence is expected since TCP's self clocking mechanism assures that transfer rates and times will be correlated for short connections. Implementations of TCP which include features designed to make bandwidth sharing more equitable and to improve slow start functionally may make these effects less pronounced.

Another difficulty in the model is identifying Poisson time points from the data. This is impossible with certain data sets as the Munich data and difficult with the BU data. A common paradigm is that activities initiated by humans is well modeled by Poisson processes. This was in agreement with our analyses of the UCB data (Table 4.2 and Fig. 4.6). However this simple paradigm has to be informed by the non stationarity of most collected data and by the fact that a Poisson process model is only expected (and observed) when activities initiated by many humans are aggregated. Furthermore, many types of web based activity are initiated by machine and cannot be expected to follow the Poisson process except perhaps at extremely high aggregation levels. A possible refinement of the infinite source Poisson model is a Poisson cluster model where activity triggered by humans is modeled by Poisson

cluster initiations and machine triggered actions are the cluster points associated to the Poisson cluster starts. Another alternative is the Markov modulated Poisson process or Markovian arrival process.

Here are some additional final comments.

- All the traces had very clear diurnal and weekly variations and trends. An informal technique which is widely used is to select a subset of the data for analysis which is visually stationary. An often used heuristic rule is to not consider data over time intervals greater than 4 hours. In situations where data is often copious, this waste of data may not be serious but thought needs to be given to models which incorporate the non stationarity explicitly. If data subset selection is used, the choice of subset should be determined by stationary situations of particular interest; for example the desire to model times of peak load.
- It is important to know which time scale is of interest. We are able to detect behavioral differences over time scales in the range 0.01s 10 s. Also, at very fine resolutions of the order of microseconds, protocols and clock frequencies are very influential as was seen in the Munich high resolution data. Pareto behavior of tails of transfer times is not apparent until times exceed tens of seconds and therefore correlation behavior of the cumulative input will not match those of the limiting models for lags of smaller order. Hence, one should not expect limiting approximations to be applicable at resolutions finer than tens of seconds.
- Without added refinements, the infinite source Poisson model is clearly not capable of describing behavior on very fine time scales. Fine time scale behavior is presumably affected by the passage through many protocol layers and control mechanisms such as TCP and thought to modelling the effect of such controls is urgently needed. For some results in this direction using the concept of multifractals see Refs.[75,76]. Investigating such a refinement of the model is a long term goal.
- Models should be quite different for high and low numbers of users or active nodes. For relatively low numbers of active nodes, say up to several hundred users, rather detailed models are needed. These should incorporate at least the varying number of active users, the activity levels of the users, the specific kinds of tasks of interest, and machine generated bursts of transfers of several files caused by one user request. In addition it is likely that models and approximations should be rather different at high and low utilization levels; that is when traffic rates constitute a high percentage of the maximum rate allowed by a link.

  In situations with superposition of a large number of users (thousands or more) the assumptions of the infinite source Poisson model gave a good description of user behavior during stationary periods. In fact the asymptotic fractional Brownian motion described well the simplified version of the traffic obtained by enforcing the assumption of constant transfer rates, for the UCB data, as discussed in Section 4.6 above.
- For our data, fBm was inappropriate for modeling the real traffic rates with varying transfer rates. When the transfer rates were artificially set to be constant, in particular for the UCB syn trace, there was good fit. It is clear that also at coarse resolutions

actual network traffic is strongly influenced by flow control mechanisms such as TCP, server behavior, congestion, caching strategies and other factors. To obtain useful and realistic models, these factors cannot be ignored. Since much of the network traffic passes through a large and varying number of routers and switches, simplistic modeling of the behavior of the queue in one router is unlikely to achieve acceptable levels of realism. To find realistic and useful models for highly loaded systems serving large number of users which are subject to flow control, protocol modifications and congestion is an urgent area for further modeling research. The urgency is emphasized by the rapidly changing nature of the Internet.

APPENDIX: PROOFS OF THE RESULTS IN SECTION 2.2.4

We now discuss the verification of the claims in Subsection 2.2.4 and continue to apply the results of Ref.[52] with $A_k(s) = s \wedge L_k$. We have

$$E(A(t)) = \lambda \int_0^t E(u \wedge L_1)\, du = \lambda \int_0^t \int_0^u \bar{F}(u) dv\, du \backsim \lambda \mu t,$$

$$\sigma^2(t) = E\left(\int_0^t (L_1 \wedge u)^2 du \right) = \lambda \int_{u=0}^t \left(\int_{v=0}^u \bar{F}(v) 2v\, dv \right) du$$

so that,

$$\lim_{t\to\infty} \frac{\sigma^2(t)}{t} = \lambda \lim_{u\to\infty} \int_0^u \bar{F}(v)\, 2v dv = \lambda \int_0^\infty x^2 F(dx) = \lambda E(L_1^2),$$

and we conclude

$$\sigma^2(t) \backsim t\lambda E(L_1^2). \tag{5.1}$$

We get, as $T \to \infty$, that for $s < t$,

$$E(A_1(Ts)A_1(Tt)) = E\left(L_1^2 1_{[L_1 \le s]}\right) = sE\left(L_1 1_{[s \le L_1 \le t]}\right)$$

$$= stP[L_1 > t] \to \int_0^\infty x^2 F(dx) = E(L_1^2), \tag{5.2}$$

since when $\alpha > 2$ we have $T^2 \bar{F}(Tt) \to 0$ and

$$Ts \int_{Ts}^{Tt} xF(dx) \le Ts \int_{Ts}^{Tt} \frac{x^2}{Ts} F(dx) = 1 \cdot \int_{Ts}^{Tt} x^2 F(dx) \to 0.$$

Therefore, when $\alpha > 2$, the function $E(A_1)(s)A_2(t))$ is two dimensional regularly varying with index 0. Thus (Ref.[52], Proposition 2.2) $E(A(s)A(t))$ is regularly varying with index 1 and limit function ($x < y$)

$$\hat{C}(x, y) := \int_0^x 1 du = x,$$

which is the covariance function of Brownian motion. Unlike the case $1 < \alpha < 2$, where convergence in distribution to Brownian motion did not hold, it will hold in in the present case $2 < \alpha$. This is relatively easy to verify since

$$\frac{1}{t}\int_{\epsilon\sqrt{t}}^{\infty} y \int_{u}^{t} \bar{F}(y)dudy = \frac{1}{t}\int_{\epsilon\sqrt{t}}^{\infty} y(t-y)\bar{F}(y)dy \leq \frac{t}{t}\int_{\epsilon\sqrt{t}}^{\infty} y\bar{F}(y)dy$$

$$= \int_{\epsilon\sqrt{t}}^{\infty} y\bar{F}(y)\,dy \to 0, \qquad (t \to \infty),$$

since $E(L_1^2) < \infty$. From (Ref. [52], Theorem 3.3), we get

$$A^{(T)}(t) := \frac{A(Tt) - \lambda TtE(L_1)}{\sqrt{\lambda TE(L_1^2)}} \Rightarrow B(t), \tag{5.3}$$

in the sense of convergence of finite dimensional distributions where $B(\cdot)$ is a standard Brownian motion.

A version of Eq. (5.3) in the J_1 topology (cf. Ref.[13]) also holds according to (Ref.[52], Proposition 3.4), provided we show as $T \to \infty$

$$\frac{1}{\sqrt{T}}\int_0^T (EA_1(\infty) - EA_1(u))du \to 0 \tag{5.4}$$

and

$$\frac{1}{\sqrt{T}} \bigvee_{s=0}^{T} \left| \sum_{i=1}^{M([0,s]\times\mathbb{R}_+)} (A_i(\infty) - A_i(s-\Gamma_i)) \right| \xrightarrow{P} 0. \tag{5.5}$$

Since $A_1(\infty) = L_1$, Eq. (5.4) becomes

$$\frac{1}{\sqrt{T}}\int_0^T E(L_1 - L_1 \wedge u)\,du = \frac{1}{\sqrt{T}}\int_0^T E\big((L_1 - u)1_{[L_1>u]}\big)\,du$$

which by the Schwartz inequality is bounded by

$$\leq \frac{1}{\sqrt{T}}\int_0^T \sqrt{E(L_1^2)\bar{F}(u)}du = \frac{\sqrt{E(L_1^2)}}{\sqrt{T}}\int_0^T \sqrt{\bar{F}(u)}du.$$

Since $\bar{F}$ is regularly varying, so is $\sqrt{\bar{F}}$ and its index is $-\alpha/2 < -1$. So $\sqrt{\bar{F}(u)}$ is integrable and the ratio converges to 0 as required.

We now show how to verify Eq. (5.5) for the case $\alpha > 2$. The left side of Eq. (5.5) is

$$\frac{1}{\sqrt{T}} \bigvee_{s=0}^{T} \left| \sum_{i=1}^{M([0,s]\times\mathbb{R}_+)} (L_i - (s-\Gamma_i) \wedge L_i) \right|$$

and

$$L_i - (s - \Gamma_i) \wedge L_i = \begin{cases} 0, & \text{if } s - \Gamma_i \geq L_i, \\ L_i - (s - \Gamma_i), & \text{if } s - \Gamma_i < L_i, \end{cases}$$

$$= \begin{cases} 0, & \text{if } \Gamma_i + L_i \leq s, \\ \Gamma_i + L_i - s, & \text{if } \Gamma_i + L_i > s. \end{cases}$$

Therefore Eq. (5.5) can be rewritten

$$\frac{1}{\sqrt{T}} \bigvee_{s=1}^{T} \sum_{i=1}^{M([0,s]\times\mathbb{R}_+)} (\Gamma_i + L_i - s) 1_{[\Gamma_i + L_i > s]} \xrightarrow{P} 0. \tag{5.6}$$

Since $\Gamma_n \sim n$ almost surely, it is easy to see that Eq. (5.6) is implied by

$$\frac{1}{\sqrt{n}} \bigvee_{s\in[0,\Gamma_n)} \sum_{i=1}^{M([0,s]\times\mathbb{R}_+)} (\Gamma_i + L_i - s) 1_{[\Gamma_i + L_i > s]} \xrightarrow{P} 0. \tag{5.7}$$

The left side of Eq. (5.7) is

$$\frac{1}{\sqrt{n}} \bigvee_{j=0}^{n} \bigvee_{s\in[\Gamma_j,\Gamma_{j+1}]} \sum_{i=1}^{j} (\Gamma_i + L_i - s) 1_{[\Gamma_i = L_i > s]} \leq \frac{1}{\sqrt{n}} \bigvee_{j=0}^{n} \sum_{i=1}^{j} (\Gamma_i + L_i - \Gamma_j) 1_{[\Gamma_i + L_i > \Gamma_j]}.$$

Therefore, Eq. (5.6) holds if

$$\sum_{j=0}^{n} P\left[\sum_{i=1}^{j} (\Gamma_i + L_i - \Gamma_j) 1_{[\Gamma_i + L_i > \Gamma_j]} > \epsilon\sqrt{n} \right] \to 0 \tag{5.8}$$

as $n \to \infty$ for every $\epsilon > 0$. Now the jth probability term in Eq. (5.8) is

$$P\left[\sum_{i=1}^{j} (L_i - (\Gamma_j - \Gamma_i)) 1_{[L_i > \Gamma_j \quad \Gamma_i]} > \epsilon\sqrt{n} \right] = P\left[\sum_{i=1}^{j} (L_i - \Gamma_{j \quad i}) 1_{[L_i > \Gamma_{j \quad i}]} > \epsilon\sqrt{n} \right]$$

and because $\{\Gamma_n\}$ is independent of $\{L_j\}$, this the same as

$$= P\left[\sum_{i=1}^{j} (L_i - \Gamma_i) 1_{[L_i > \Gamma_i]} > \epsilon\sqrt{n} \right] \leq P\left[\sum_{i=1}^{\infty} (L_i - \Gamma_i) 1_{[L_i > \Gamma_i]} > \epsilon\sqrt{n} \right]$$

$$= P\left[\int \int_{D^{\uparrow}} (y - s) M(dsdy) > \epsilon\sqrt{n} \right] \tag{5.9}$$

where

$$D^{\uparrow} = \{(s, y) \in \mathbb{R}_+^2 : y > s\}.$$

Note

$$EM(D^{\uparrow}) = \int_{s=0}^{\infty} \int_{y>s} \lambda ds\, F(dy) = \lambda \int_0^{\infty} \bar{F}(s) ds = \lambda\mu$$

and therefore, $M|_{D^{\uparrow}}$ is a finite point measure. It may be represented as

$$M|_{D^{\uparrow}} \stackrel{d}{=} \sum_{i=1}^{PO(\lambda\mu)} \epsilon_{(\xi_i, \eta_i)},$$

where $PO(\lambda\mu)$ is a Poisson random variable with parameter $\lambda\mu$, independent of the iid pairs $\{(\xi_i, \eta_i),\ \ i \geq 1\}$ which have common distribution

$$\left.\frac{\lambda ds F(dy)}{\lambda\mu}\right|_{\mathrm{D}^{\uparrow}} = \left.\frac{ds F(dy)}{\mu}\right|_{D^{\uparrow}}.$$

So for $x > 0$,

$$P[\eta_i - \xi_i > x] = \int \int_{\substack{s>0 \\ y>s+x}} ds \frac{F(dy)}{\mu} = \int_x^{\infty} \frac{\bar{F}(s)}{\mu} ds.$$

Set $\zeta_{\mathrm{i}} = \eta_{\mathrm{i}} - \xi_{\mathrm{i}}$. The probability in Eq. (5.9) is

$$P\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n}\right]$$

and Eq. (5.8) is bounded by

$$nP\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n}\right]. \tag{5.10}$$

If $\alpha > 2$, then $\mathrm{Var}(\zeta_i) < \infty$ and there exists $\eta > 0$ such that $E(\zeta_i^{2+\eta}) < \infty$. Pick θ such that

$$0 < \theta < \frac{\eta}{2(2+\eta)}. \tag{5.11}$$

Then

$$\begin{aligned} nP\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n}\right] &= nP\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n},\ \ PO(\lambda\mu) \leq n^{\theta}\right] \\ &\quad + nP\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n},\ \ PO(\lambda\mu) > n^{\theta}\right] \\ &\leq nP\left[\sum_{i \leq n^{\theta}} \zeta_i > \epsilon\sqrt{n}\right] + P[PO(\lambda\mu) > n^{\theta}] = I + II. \end{aligned}$$

Now by Chebychev's inequality

$$II \leq \frac{nE(e^{PO(\lambda\mu)})}{\exp\{n^{\theta}\}} = \frac{ne^{\lambda\mu(e\ \ 1)}}{\exp\{n^{\theta}\}} \to 0,$$

as $n \to \infty$. For I we have

$$I \leq \frac{nE\left(\sum_{i\leq n^{\theta}}\zeta_i\right)^{2+\eta}}{\left(\epsilon\sqrt{n}\right)^{2+\eta}} = (\text{const})\frac{E\left(\sum_{i\leq n^{\theta}}\zeta_i\right)^{2+\eta}}{n^{\eta/2}}.$$

Let $\|\zeta\|_p$ be the L_p norm of a random variable ζ and the bound in I is of the form

$$\frac{(\text{const})}{n^{\eta/2}}\left(\left\|\sum_{i\leq n^{\theta}}\zeta_i\right\|_{2+\eta}\right)^{2+\eta}$$

which by Minkowsky's inequality is dominated by

$$\frac{(\text{const})}{n^{\eta/2}}(n^{\theta}\|\zeta_1\|_{2+\eta})^{2+\eta} = (\text{const}')\frac{n^{\theta(2+\eta)}}{n^{\eta/2}} \to 0$$

as $n \to \infty$, due to Eq. (5.11).

Having shown Eq. (5.5) $\alpha > 2$, we get tightness and hence the functional form of Eq. (5.3).

This leads to a functional limit for the content process $X(\cdot)$ of Eq. (2.9) and verifies the claims Eq. (2.14) and (2.15) as follows. Define

$$\xi_T(t) = A(Tt) - rTt$$

so that from Eq. (5.3) we have

$$\frac{\xi_T(t) - Tt(\lambda\mu - r)}{\sqrt{\lambda TE\left(L_1^2\right)}} \Rightarrow B(t)$$

in $C[0,\infty)$. If $\lambda\mu > r$, so that the system is unstable since the input overwhelms the traffic rate, then applying (Ref.[87], Theorem 6.4 [ii]) yields

$$\frac{X(Tt) - Tt(\lambda\mu - r)}{\sqrt{\lambda TE\left(L_1^2\right)}} \Rightarrow B(t) \tag{5.12}$$

in $C[0,\infty)$. Note, the condition $\lambda\mu > r$ guarantees $A(Tt) - rTt \xrightarrow{P} \infty$ and seems necessary to get a nontrivial limit in Eq. (5.12) due to the denominator becoming infinite. If $\lambda\mu < r$, then the limit of the term $X(Tt)$ in Eq. (5.12) is the function which is identically 0 by (Ref.[84], Theorem 6.4 [iii], page 81).

ACKNOWLEDGMENT

We are very grateful to Helmut Gogl, for providing us with the Munich lo and Munich hi data sets, and to P. Abry, for letting us use his wavelet estimation program, and to both for illuminating discussions. We also want to thank M. Greiner,Å. Arvidsson, P. Abry, W. Willinger, Steve Gribble, and J. Istas for useful conversations and help with details and references.

REFERENCES

1. Abry, P. *Ondelettes et Turbulences Multirésolutions, Algorithmes de Décompo sitions, Invariance d'Echelle et Signaux de Pression*. Diderot, Editeur des Sciences et des Arts, Paris, 1997.
2. Abry, P.; Delbecke, L.; Flandrin, P. Waveletbased estimator for the self similarity parameter of α stable processes. IEEE ICASSP 99 p III 1581 1584, Phoenix, Ar, USA, March 1999.
3. Abry, P.; Flandrin, P.; Taqqu, M.S.; Veitch, D. Wavelets for the analysis, estimation and synthesis of scaling data. In *Self similar network analysis and performance evolution*; Park, K., Willinger, W., Eds.; Wiley: New York, 2000; 30 88.
4. Abry, P.; Veitch, D. Wavelet analysis of long range dependent traffic. IEEE Trans. Inform. Theory **1998**, *44*, 2 15.
5. Adler, R.J. An introduction to continuity, extrema, and related topics for general Gaussian processes. I.M.S. Lecture Notes Monogr. Series **1990**, 12.
6. Adler, R.J.; Feldman, R.; Taqqu, M.S. *A practical guide to heavy tails: statistical techniques and applications*; Brikhauser: Boston, 1998.
7. Arlitt, M.; Williamson, C.L. Web server workload characterization: the search for invariants (extended version). *Proceedings of the ACM Sigmetrics Conference, Philadelphia, PA*, 1996. Available from {mfa126,carey}@cs.usask.ca.
8. Andersen, A.T.; Nielsen, B.F. On the statistical implications of certain random permutations in Markovian Arrival Processes (MAP)s and second order self similar processes. Perform. Eval. **2000**, *41*, 67 82.
9. Asmussen, S. *Applied Probability and Queues*; Wiley: Chichester, UK, 1987.
10. Barford, P.; Bestavros, A.; Bradely, A.; Crovella, M.E. Changes in web client access patterns: characteristics and caching implications. World Wide Web, Special Issue on Characterization and Perform. Eval. **1999**, *2*, 15 28.
11. Beirlant, J.; Vynckier, P.; Teugels, J. Tail index estimation, Pareto quantile plots, and regression diagnostics. J. Am. Stat. Assoc. **1996**, *91*, 1659 1667.
12. Beran, J. *Statistics for long memory processes*; Chapman & Hall: New York, 1994.
13. Billingsley, P. *Convergence of probability measures*, 2nd Ed.; Wiley: New York, 1999.
14. Cambanis, S.; Houdré, C. On the continuous wavelet transform of second order random processes. IEEE Trans. Inform. Theory **1995**, *41*, 628 642.
15. Cohen, J.; Resnick, S.; Samorodnitsky, G. Sample correlations of infinite variance time series models: and empirical and theoretical study. J. Appl. Math. Stochast. Anal. **1998**, *1*, 255 282.

16. Constantine, A.G.; Hall, P. Characterizing surface smoothness via estimation of effective fractal dimension. J. R. Stat. Soc. B **1994**, *56*, 97 113.
17. Crovella, M.; Bestavros, A. Explaining world wide web traffic self similarity. Preprint available as TR 95 015 from {crovella,best}cs.bu.edu, 1995.
18. Crovella, M.; Bestavros, A. Self similarity in world wide web traffic: evidence and possible causes. *Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modelling of Computer Systems*, 1996.
19. Crovella, M.; Bestravros, A.; Taqqu, M. Heavy tailed distributions in the world wide web. *A practical guide to heavy tails: statistical techniques and applications*; Birkhauser: Boston, **1998**.
20. Crovella, M.; Kim, G.; Park, K. On the relationship between file sizes, transport protocols, and self similar network traffic. *Proceedings of the Fourth International Conference on Network Protocols (ICNP '96)*, 1996; 171 180.
21. Crovella, M.; Park, K.; Kim, G. On the effect of traffic self similarity on network performance. *Proceedings of the SPIE International Conference on Performance and Control of Network Systems*, 1997.
22. Cunha, C.R.; Crovella, M.; Bestavros, A. Characteristics of www client based traces. Preprint BU CS 95 010, Boston University, 1995.
23. Daubechies, I. *Ten lectures on wavelets*; SIAM: Philadelphia, PA, **1992**.
24. Davis, R.; Resnick, S.I. Limit theory for bilinear processes with heavy tailed noise. Ann. Appl. Probab. **1996**, *6*, 1191 1210.
25. Durrett, R.; Resnick, S. Weak convergence with random indices. J. Stochast. Proc. Appl. **1977**, *5*, 213 220.
26. Embrechts, P.; Klüppelberg, C.; Mikosh, T. *Modelling extremal events for insurance and finance*; Springer: Heidelberg, **1997**.
27. Erramilli, A.; Narayan, O.; Willinger, W. Experimental queueing analysis with long range dependent packet traffic. IEEE/ACM Trans. Network Comput. **1996**, *4*, 209 223.
28. Feldmann, A.; Gilbert, A.C.; Willinger, W. Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic. *Proc. of the ACM/SIGCOMM '98*, Vancouver, B.C., 1998; 25 38.
29. Flandrin, P. Wavelets analysis and synthesis of fractional Brownian motion. IEEE Trans. Inform. Theory **1992**, *38*, 910 917.
30. Feigin, P.; Resnick, S. Pitfalls of fitting autoregressive models for heavy tailed time series. Extremes **1999**, *1*, 391 422.
31. Gilbert, A.C.; Willinger, W.; Feldmann, A. Scaling analysis of conservative cascades, with applications to network traffic. IEEE Trans. Inform. Theory **1999**, *45* 3, 971 991.
32. Gogl, H. Data Acquisition and Data Analysis in ATM based Networks. Report 1998, Technical University of Munich, available at http://wwwjessen.informatik.tu muenchen.de/forschung/leistungsanalyse/ATM_Project/atm_index.html.
33. Gogl, H. Traffic statistics and application profiles in ATM networks. Report 1998, Technical University of Munich, available at http://wwwjessen.informatik.tu muenchen.de/forschung/leistungsanalyse/ATM_Project/atm_index.html.

34. Gladyshev, E.G. A new limit theorem for processes with Gaussian increments. Theory Probab. Appl. **1961**, *6*, 52 61.
35. Guérin, C.A. Wavelet analysis and covariance structure of some classes of non stationary processes. J. Fourier Anal. Appl. **2000**, *4* 6, 403 425.
36. Hall, P.; Wood, A. On the performance of box counting estimators of fractal dimension. Biometrika **1993**, *80*, 246 252.
37. Hall, P.; Wood, A.; Feuerverger, A. Estimation of fractal index and fractal dimension of a Gaussian process by counting the number of level crossings. J. Timer Ser. Anal. **1994**, *15* n°6, 587 606.
38. Michael Harrison, J. *Brownian Motion and Stochastic Flow Systems*; Wiley: New York, **1985**.
39. Heath, D.; Resnick, S.; Samorodnitsky, G. Heavy tails and long range dependence in on/off processes and associated fluid models. Math. Oper. Res. **1998**, *23*, 145 165.
40. Heath, D.; Resnick, S.; Samorodnitsky, G. How system performance is affected by the interplay of averages in a fluid queue with long range dependence induced by heavy tails. Ann. Appl. Probab. **1999**, *9*, 352 375.
41. Hill, B.M. A simple general approach to inference about the tail of a distribution. Ann. Stat. **1975**, *3*, 1163 1174.
42. Hunt, F. Error analysis and convergence of capacity dimension algorithms. SIAM J. Appl. Math. **1990**, *50*, 307 321.
43. Ibragimov, L.A.; Rozanov, Y.A. *Gaussian random processes*; Springer Verlag: Berlin, **1978**.
44. Istas, J. Wavelets coefficients of a Gaussian process and applications. Ann. Inst. Henri Poincaré **1992**, *28*, 537 556.
45. Istas, J.; Lang, G. Quadratic variations and estimation of the local Hölder index of a Gaussian process. Ann. Inst. Henri Poincaré **1997**, *33*, 407 436.
46. Istas, J.; Laredo, C. Estimation de la fonction de covariance d'un processus gaussien stationnaire par méthodes d'échelles. C. R. Acad. Sci., Tome **1993**, *316*, 495 498, Série I.
47. Jelenković P.R.; Lazar, A.A. Subexponential asymptotics of a Markov modulated G/G/1 queue. Technical report, Columbia University, 1995.
48. Jelenković P.R.; Lazar, A.A. Multiplexing on/off sources with subexponential on periods. Technical Report CTR. 457 96 23, Columbia University, 1996.
49. Jelenković, P.; Lazar, A. Asymtotic results for multiplexing subexponential on off processes. Adv. Appl. Probab. **1999**, *31*, 394 421.
50. Kaj, I. Convergence of scaled renewal processes to fractional Brownian motion. Technical report, available at http://www.math.uu.se/~ikaj/papers.html, 1999.
51. Kent, J.T.; Wood, A.T.A. Estimating the fractal dimension of a locally self similar Gaussian process by using increments. J. R. Stat. Soc. B **1997**, *59* n°3, 679 699.
52. Klüppelberg, C.; Mikosch, T. Explosive Poisson shot noise processes with applications to risk reserves. Bernoulli **1995**, *1*, 125 148.
53. Konstantopoulos, T.; S. J., Lin Macroscopic models for long range dependent network traffic. Queueing Systems Theory Appl. **1998**, *28*, 215 243.
54. Kratz, M.; Resnick, S. The qq estimator and heavy tails. Stochast. Models **1996**, *12*, 699 724.

55. Kurtz, T. Limit theorems for workload input models. In *Stochastic Networks: Theory and Applications*; Zachary, S., Kelly, F.P., Ziedins, I., Eds.; Clarendon Press: Oxford, **1996**.
56. Lamperti, J.W. Semi stable stochastic processes. Trans. Am. Math. Soc. **1962**, 62 78.
57. Lang, G. Estimation de la régularité et de la longue dépendence de processus gaussiens. Thése de l'université Toulouse III, 1994.
58. Leland, W.E.; Taqqu, M.S.; Willinger, W.; Wilson, D.V. On the self similar nature of Ethernet traffic. *ACM/SIGCOMM Computer Communications Review*, 1993. Proceedings of the ACM/SIGCOMM93. To appear.
59. Leland, W.E.; Taqqu, M.S.; Willinger, W.; Wilson, D.V. On the self similar nature of Ethernet traffic (extended version). IEEE/ACM Trans. Networking **1994**, *2*, 1 15.
60. Leland, W.E.; Taqqu, M.S.; Willinger, W.; Wilson, D.V. Statistical analysis of high time resolution Ethernet Lan traffic measurements. In. *PCmpScSt25*, 146 155, 1993
61. Maulik, K.; Resnick, S.; Rootzen, H. Asymptotic independence and a network traffic model. J. Appl. Probab. **2002**, *4*, 671 699.
62. Maulik, K.; Resnick. S. Small and large time scale analysis of a network traffic model. Technical report, available at www.orie.cornell.edu/trlist/trlist.html, 2001.
63. Miokosch, T.; Resnick, S.; Rootzén, H.; Stegeman. A. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? Preprint 1999, available as #1999:32 at http://www.md.chalmers.se/Stat/Research/Preprints.
64. Nolan, J.P. *Maximum likelihood estimation and diagnostics for stable processes. Levy processes*; Birkhaüser Boston: Boston, MA, **2001**.
65. Prabhu, N.U. *Stochastic Storage Processes: Queues, Insurance risk, Dams and Data Communication*; Springer: New York, **1998**.
66. Reiss, R.D.; Thomas, M. *Statistical analysis of extreme values*; Birghäuser: Basel, **1997**.
67. Resnick, S.I. *Extreme Values, Regular Variation and Point Processes*; Springer: New York, **1987**.
68. Resnick, S.I. Heavy tail modelling and teletraffic data. Ann. Stat. **1997**, *25*, 1805 1869.
69. Resnick, S.; Samorodnitsky, G. Performance decay in a single server exponential queuing model with long range dependence. Oper. Res. **1997**, *45*, 235 243.
70. Resnick, S.; Samorodnitsky, G.; Xue, F. How misleading can sample acf's of stable ma's be? (Very!). Ann. Appl. Probab. **1999**, *9* 3, 797 817.
71. Resnick, C. Smoothing the Hill estimator. Adv. Appl. Probab. **1997**, *29*, 271 293.
72. Resnick, S.I.; Rootzén, H. Self similar communication models and very heavy tails. Ann. Appl. Probab. **2000**, *10*, 753 778.
73. Resnick, S.; van den Berg, E. A test for nonlinearity of time series with infinite variance. Extremes **2000**, *3* 4, 145 172.
74. Resnick, S.; van den Berg, E. Weak Convergence of high speed network traffic models. J. Appl. Probab. **2000**, *37* 2, 575 597.
75. Riedi, R.H. Multifractal processes. *Long range dependence: theory and applications*; Birkhauser: Boston, **2001**.

76. Riedi, R.H.; Willinger, W. Toward an improved understanding of network traffic dynamics. *Self similar Network Traffic and Performance Evaluation*; Wiley: New York, **2000**.
77. Samorodnitsky, G.; Taqqu, M. *Stable non Gaussian random processes*; Chapman and Hall: New York, **1994**.
78. Tajvidi N. Confidence Intervals and Accuracy Estimation for Heavy tailed Generalized Pareto Distribution. Part of PhD thesis, 1996, Chalmers, available at http://www.mai.liu.se/~nataj/
79. Taqqu, M.; Willinger, W.; Sherman, R. Proof of a fundamental result in self similar traffic modelling. Comp. Commun. Rev. **1997**, *27*.
80. Mandelbrot, B.B.; Taqqu, M.S. Robust R/S analysis of long run serial correlation. *Proceedings of the 42nd Session of the International Statistical Institute Manila*, Vol.48, Book 2, pp. 69 104, Bulletin of the ISI, 1979.
81. Tewfik, A.H.; Kim, M. Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. IEEE Trans. Inform. Theory **1992**, *38*, 904 909.
82. Veitch, D.; Abry, P. Estimation conjointe en ondelettes des paramèters du phénomène de dépendence longue. *Proc. 16ième Colloque GRETSI*, Grenoble, France, 1997; 1451 1454.
83. Veitch, D.; Abry, P. A wavelet based joint estimator for the parameters of Ird. IEEE Trans. Inform. Theory. **1998**, special issue "Multiscale Statistical Signal Analysis and its Application".
84. Whitt, W. Some useful functions for functional limit theorems. Math. Oper. Res. **1980**, *5*, 67 85.
85. Willinger, W.; Taqqu, M.S.; Leland, M.; Wilson, D. Self similarity in high speed packet traffic: analysis and modelling of ethernet traffic measurements. Stat. Sci. **1995**, *10*, 67 85.
86. Willinger, W.; Taqqu, M.S.; Leland, M.; Wilson, D. Self similarity through high variability: statistical analysis of ethernet lan traffic at the source level. Comput. Commun. Rev. **1995**, *25*, 100 113, Proceedings of the ACM/SIGCOMM'95, Cambridge, MA.
87. Willinger, W.; Taqqu, M.S.; Leland, M.; Wilson, D. Wilson, D. Self similarity through high variability: statistical analysis of ethernet lan traffic at the source level (extended version). IEEE/ACM Trans. Networking **1996**, To appear.
88. Wood, A.T.A.; Chan, G. Simulation of stationary Gaussian Processes in $[0,1]^d$. J. Comput. Graphical Stat. **1994**, *3*, 409 432.