Budapest University of Technology and Economics
Department of Telecommunications and Media Informatics

Characterisation of Self-Similar Traffic in Data Networks

András Gefferth

Scientific supervisors
Dr. Sándor Molnár and Dr. Darryl N. Veitch

Budapest, 2005

1 Introduction

In order to design telecommunication networks, network protocols and applications it is important to know the properties of the traffic intensity of the network. This intensity can be characterised using stochastic descriptors. Stochastic models for telephone networks have been developed in the beginning of the 20th century and by now well established methods exist for the dimensioning and planning of such networks. It was found that the traffic intensity of computer networks differs significantly from that of telephone networks. The traffic of the former has a very strong correlation resulting in long-range dependent and self-similar characteristics.

Long-range dependent stochastic processes have been discovered in various natural or man-made systems. Their first discovery relates to hydrology. In the 1950s Hurst analysed the historical data of the water level of the river Nile which was available for several hundred years. Hurst observed the presence of a very strong correlation. Since then long-range dependence has been encountered in various other fields such as agriculture, physics, soil science and also in telecommunications networks [1].

Several different definitions can be found in the literature for long-range dependence. Some of the most relevant will be given in Section 4.3.2 on page 11 and more are listed and discussed in the dissertation. Here we present a few important properties that are the most characteristic of these processes and as such are shared by almost all of the definitions.

- Long-range dependence is usually defined for discrete time, stationary stochastic processes.

- The autocovariance function $\gamma(k)$ has a slow power-law decay and so its infinite sum $\sum_{i=0}^{\infty} \gamma(k)$ is infinite.

- The variance of the sample mean calculated as $\bar{X} := \frac{\sum_{i=1}^{m} X(i)}{m}$ decreases slowly, i.e. slower than a constant times m^{-1}, where m is the sample size, and X is the long-range dependent process.

- The autocorrelation function, which describes the qualitative behaviour of the correlation structure, converges pointwise to a constant function when viewing the same process on successively larger time scales. This behaviour is referred to as asymptotic self-similarity, and is illustrated in Figure 1. Here we see that the process is visually similar on all the depicted time scales.

In telephone networks the increase in utilisation or the multiplexing of several traffic flows results in a smoother traffic intensity. Because of this smoothness it is possible to achieve a high utilisation of the network, since the peak rate and average rate are close to each other. But in data networks, an increase in utilisation does not lead to a smoother traffic. Therefore in order to achieve high utilisation long packet buffers are needed, which introduce long delay. Although there are several other differences between voice and data networks this example shows that a different approach is needed for designing and dimensioning of data networks.
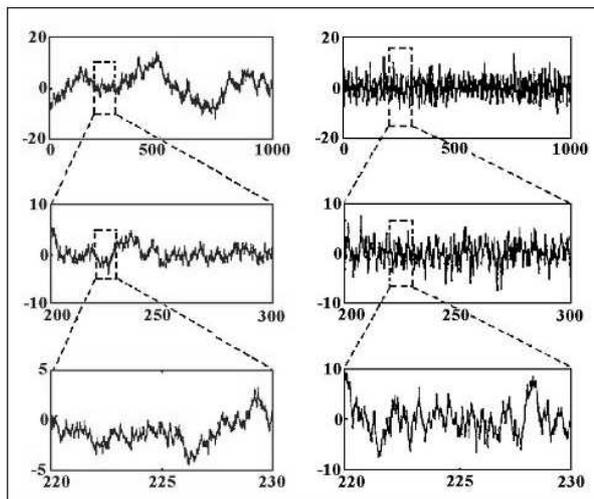
Figure 1: Depicting the same process on different time scales. If the time scale is chosen to be small (bottom row) both processes show high variability. As the time scale increases the on on the right-hand side shows a smooth curve. In contrast the left-hand side process has the burst-within-burst structure and so the bursts do not disappear even if the time granularity is coarser (middle and top rows).

Besides the effect of long-range dependence its cause was also investigated. The heavy-tailed file size distribution, or the effect of the TCP[1] mechanisms were among the possible causes [5, 10]. Also different estimators have been developed to test the presence and estimate the strength of long-range dependence [1].

The contribution of this work relates to the underlying theory of long-range dependent and self-similar stochastic processes.

2 Research Objectives

This research originated from the task of accurately tracking the mean and detecting the change of the mean of traffic intensity in data networks. This required the study of the corresponding literature in order to get acquainted with the theoretical background of the corresponding long-range dependent traffic models. During this study I had to realise that despite network modelling has been in the focus of active research some important issues are still not completely understood. This motivated a thorough study of long-range dependent processes.

Although I was aware that from a mathematical point of view long-range dependent processes form a subset of asymptotically self-similar ones, the study of this wider class did not seem to be important for my original goals of research.

However, during my work when I analysed a family of asymptotically self-similar pro-

---

[1]Transmission Control Protocol

cesses, namely the family of the so-called fARIMA processes, I found some strange phenomenon, which I was unable to interpret using existing results. This led to the investigation of the whole class of exactly and asymptotically self-similar processes[2].

It was thus the incompleteness of the corresponding theory that made me investigate theoretical issues related to my original goal.

I started to investigate the basic properties of discrete time, exactly and asymptotically self-similar as well as long-range dependent processes, and also the relation of these. Several definitions exist in the literature for these concepts. I collected and compared these definitions. There are also many results scattered in the literature, in some cases with incomplete, incorrect or missing proofs. My goal was to collect and organise these results, provide the missing proofs, state and prove some "missing" theorems that are needed to have a complete theoretical background of the subject.

3 Methodology

All the presented results are based on analytical studies. During my work I wrote several short programs, which proved to be useful in confirming or rejecting unproved hypotheses, and also presenting new and interesting research topics in case the results of the calculations could not be interpreted by the available theorems. For the study of discrete time deterministic functions, such as the autocovariance or autocorrelation functions, an operator formalism was developed. This has allowed the separation of the study of the functional relationship between different descriptors of stochastic processes and the study of additional criteria, such as positive semi-definiteness, these functions have to satisfy.

Apart from the field of regular variation and basic knowledge of stochastic processes no other mathematical concept was required. Some number-theoretic problems have also appeared. Their solutions are also given using simple calculations.

The study of regular variation in discrete time was needed, resulting in a contribution to the clarification of their properties.

4 New Results

In this work discrete-time second-order stationary stochastic processes are investigated. Let $\{X(t), t \in \mathbf{Z}\}$ denote a discrete time second-order stationary stochastic process. The mean $\mu = \mu(t) = \mathbf{E}[X(t)]$ and variance $\mathcal{V} = \mathcal{V}(t) = \mathbf{E}[(X(t)-\mu)^2]$ of such a process are independent of t, and the *autocovariance function*, $\gamma(k) := E[(X(t)-\mu)(X(t+k)-\mu)]$, depends only on the lag k, $k \in \mathbf{Z}$, and $\gamma(k) = \gamma(-k)$. For our purposes the process is uniquely characterised by its autocovariance structure, the exact distribution of the values is not interesting. The *autocorrelation* of the process is defined as $\rho(k) := \frac{\gamma(k)}{\gamma(0)} = \frac{\gamma(k)}{\mathcal{V}}$.

In addition to the familiar descriptors of second-order structure presented above, it proved to be advantageous to work with an equivalent pair of functions, the *variance time function*

---

[2]Although the correct mathematical proof of the observed phenomenon is still subject of ongoing research, it can now be explained in view of the current results.

and its normalised form, the *correlation time function* (CTF). The variance time function is defined as

$$\omega(n) = \sum_{k=0}^{n-1} \sum_{i=-k}^{k} \gamma(i) = n\gamma(0) + 2\sum_{i=1}^{n-1} i\gamma(n-i), n = 1, 2, 3, \cdots, \tag{1}$$

while the CTF is $\phi(n) = \frac{\omega(n)}{\omega(1)} = \frac{\omega(n)}{\mathcal{V}}$.

The autocovariance function can also be expressed in terms of the covariance time function as

$$\gamma = \mathbf{D}\{\omega\}, \tag{2}$$

where $\mathbf{D}$ is the double-differencing operator, which is defined as:

$$\mathbf{D}_i\{f(i)\}(n) = \begin{cases} f(1) & : \text{ for } n = 0 \\ \frac{1}{2}(f(2) - 2f(1)) & : \text{ for } n = 1 \\ \frac{1}{2}(f(n+1) - 2f(n) + f(n-1)) & : \text{ for } n > 1. \end{cases} \tag{3}$$

The notion of positive semi-definiteness is directly related to the study of second-order stationary processes. A function $f(k)$ defined on $k = 0, 1, 2, \ldots$ is said to be *positive semi-definite* if for any $n = 1, 2, \ldots$ and for any real vector $\mathbf{a} = [a_1, a_2, \ldots, a_n]$:

$$\sum_{1 \leq i,j \leq n} a_i f(|i - j|) a_j \geq 0.$$

It can be shown that a necessary and sufficient condition for a function to be the autocovariance function of a process is that it is positive semi-definite [3].

4.1 Characterisation of self-similar processes

Definitions of self-similarity are based on invariance under some kind of renormalisation. This renormalisation is defined as aggregation, followed by some scaling in amplitude.

Definition 4.1 (Self-Similarity #1 (SS1))
Let $X(t)$ be a process and define $X^{(m)}$ and $X'^{(m)}$ as

$$X^{(m)}(t) := \frac{1}{m} \sum_{j=m(t-1)+1}^{mt} X(j). \tag{4}$$

$$X'^{(m)}(t) := A_m \sum_{j=m(t-1)+1}^{mt} X(j) = mA_m X^{(m)}(t), \tag{5}$$

The process X is said to exhibit self-similarity if X and $X'^{(m)}$ have the same autocovariance functions for all $m \in \mathbf{Z}^+$, where A_m is a sequence (or a set of sequences) of predefined normalising constants.

The autocorrelation, autocovariance, etc. functions of $X^{(m)}$ will be denoted by $\gamma^{(m)}, \rho^{(m)}$, etc. respectively.

In the literature where this definition is found, A_m is given as $A_m := m^{-H}$ with $H \in [0, 1]$. Thus the first form of self-similarity applying to discrete time processes requires the equivalence of X and $m^{1-H} X^{(m)}(t)$.

It has to be mentioned that this definition in this form does not usually appear in the literature. Samorodnitsky and Taqqu [11] use a similar definition, but do not restrict equivalence to second-order properties, they require the equivalence of the complete distribution structure.

Sinai [12] and Major [9] also require the equivalence of the complete distribution structure, but do not restrict attention to the one dimensional stochastic process, $X(t)$, $t \in \mathbf{Z}$, these works also consider random fields in higher dimensions, $X(t_1, t_2, \ldots, t_d)$. Finally it has to be noted that discrete self-similarity can also be defined for non-stationary random fields [12].

The above definition will be compared to the following one:

Definition 4.2 (Self-Similarity #2 (SS2))
A process X is self-similar if X and $X^{(m)}$ have the same autocorrelation functions ($\rho = \rho^{(m)}$) for all $m \in \mathbf{Z}^+$.

This work uses this latter definition, SS2.

For SS1 processes $\gamma^{(m)}$ the autocovariance function of $X^{(m)}$ satisfies: $\gamma^{(m)} \equiv C_m \gamma$, where $C_m = (m A_m)^{-2} = m^{2H-2}$ is a constant explicitly defined by the sequence A_m. Now considering Definition 4.2: the autocorrelation function of a process differs from the autocovariance function only by a constant multiplicative factor, therefore if two different processes have the same autocorrelation function then the difference of their autocovariances is limited to a multiplicative constant. This yields that for SS2 processes $\gamma^{(m)} \equiv C_m \gamma$, where C_m is not prescribed, it can be any positive real value.

This shows that SS1 processes form a subset of SS2 processes. Whether this subset is strict or not depends on the set of the A_m functions, whether it includes all applicable processes or not.

Thesis 1   Exploring the set of self-similar processes

My goal was to describe the whole set of self-similar processes according to Definition 4.2, and describe them in terms of their autocorrelation or correlation time functions.

Besides the class of fractional noise processes, the only known self-similar processes, a new class of self-similar process was constructed. The fractional noise can be defined by its correlation time function $\phi = n^{2H}$ with $H \in [0, 1]$. It can be justified that this process is SS2 and also SS1, with $A_m = m^{-H}$. If all finite dimensional distributions of the process are Gaussian then the process is called fractional Gaussian noise. Since in this work we are only interested in second order properties we do not require the Gaussianity. These processes will be denoted by FN_H.

Thesis 1.1 (Definition of the almost periodic self-similar family $\mathbf{AP}_{q,c}$, [P4])
I defined the family of almost periodic processes as follows: Let q be a prime number, and $c \in (0, 1)$. Then the correlation time function of the $AP_{q,c}$ process, denoted by $\phi_{q,c}$, is defined
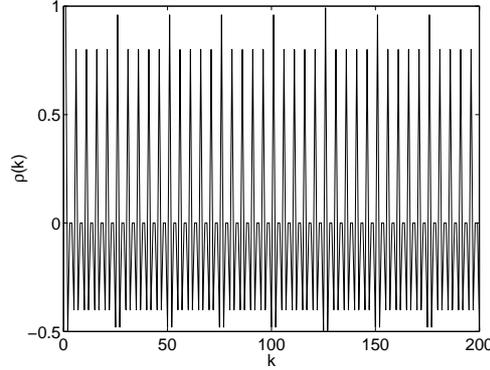
Figure 2: 200 lags of the autocorrelation function of $\text{AP}_{5,0.2}$

for p primes, as:

$$\phi_{q,c}(p) = \left\{ \begin{array}{ll} 1, & if\ p \neq q \\ c, & if\ p = q \end{array} \right.$$

At lag 1 $\phi_{q,c}(1) = 1$ and for non-primes n, $\phi_{q,c}(n)$ can be expressed as:

$$\phi_{q,c}(n) = \prod_{i=1}^{s} \phi^{r_i}(p_i),$$

where p_i are the s distinct prime factors of n, and r_i is the multiplicity of p_i.

The autocorrelation function of a member of the $\text{AP}_{q,c}$ is given in Figure 2, where its 'almost periodic' nature is readily appreciated.

Thesis 1.2 (Self-similarity of the almost periodic family, [P4])
I have shown that $\phi_{q,c}$, as defined in Thesis 1.1 is positive semi-definite, therefore the almost periodic class exists, and it is also self-similar.

Thesis 1.3 (All SS process, [P3])
I have shown that besides the two classes of self-similar processes, namely the fractional noise and the almost periodic processes no other self-similar processes exist.

4.2 Asymptotically self-similar processes

In this section the definition of asymptotical self-similarity will be given and large classes of typical examples will be shown.

Thesis 2   Limit of aggregation, definition of asymptotically self-similar processes

I investigated those processes which are not self-similar, but for which the limit $\lim_{m\to\infty} \rho^{(m)}(k) = \rho^*(k)$ exists for all $k = 0, 1, 2, \ldots$.

One can easily imagine an infinite sequence such that all elements of a sequence share a common property, but the limit of the sequence does not. I showed that this is not the

7

case with the sequence of the aggregated autocorrelation function with respect to positive semi-definiteness.

Thesis 2.1 (Positive semi-definiteness of the limiting autocorrelation , [P3])
I have shown that if ρ is a positive semi-definite autocorrelation function and $\rho^*(k) = \lim_{m \to \infty} \rho^{(m)}(k)$ exists for all $k = 0, 1, 2, \dots$ then $\rho^*(k)$ is also positive semi-definite, therefore there exists a process with autocorrelation function $\rho^*(k)$.

Thesis 2.2 (Self-similarity of the limiting autocorrelation function, [P3])
I have shown that if $\rho^*(k) := \lim_{m \to \infty} \rho^{(m)}(k)$ exists for all $k = 0, 1, 2, \dots$ then $\rho^{*(n)} \equiv \rho^*$ for all $n = 1, 2, 3 \dots$. This means that if the limit $\rho^*(k)$ exists, then it is necessary the autocorrelation function of a self-similar process.

Theses 2.1 and 2.2 allow to define asymptotically self-similar processes as:

Definition 4.3 (Asymptotically self-similar processes)
Processes for which $\rho^*(k) := \lim_{m \to \infty} \rho^{(m)}(k)$ exists are called *asymptotically self-similar*.

Some definitions of asymptotical self-similarity require the additional constraint that $\rho^*(k)$ be the autocorrelation function of the fractional noise [6]. These processes however constitute a strict subset of asymptotically self-similar processes as given in Definition 4.3.

Thesis 3   Characterisation of typical asymptotically self-similar processes

I gave large classes of typical examples for asymptotically self-similar processes. I concentrated on processes which converge to the fractional noise.

The function that expresses $\rho^{(m)}$ in terms of ρ is very complex, and therefore it is difficult to investigate its behaviour. On the other hand expressing $\phi^{(m)}$ in terms of ϕ can be done as simply as:

$$\phi^{(m)}(n) = \frac{\phi(mn)}{\phi(m)}. \tag{6}$$

Because of the simplicity of Equation (6) I gave a new definition of asymptotical self-similarity based on the correlation time function ϕ and proved its equivalence with Definition 4.3.

Thesis 3.1 (New definition of asymptotically self-similar processes, [P4, P3])
Processes for which $\phi^*(k) := \lim_{m \to \infty} \phi^{(m)}(k)$ exists are called asymptotically self-similar.

It can easily be shown that there is a one-to-one mapping between the autocorrelation function ρ and the correlation time function ϕ of a stochastic process. Here I showed that this equivalence carries over for the asymptotic properties.

Thesis 3.2 (Equivalence of ρ and ϕ, [P3])
Let ρ be the autocorrelation and ϕ be the correlation time function of a stochastic process. I have shown that $\rho^* := \lim_{m \to \infty} \rho^{(m)}$ exists if and only if $\phi^* := \lim_{m \to \infty} \phi^{(m)}$ exists and ϕ^* and ρ^* are related via the double-summing and double-differencing operators as described in Equations (1) and (2).

Based on Equation (6) the criterion for the convergence to FN_H can be written as simply as

$$\lim_{m\to\infty} \phi^{(m)}(n) = \lim_{m\to\infty} \frac{\phi(mn)}{\phi(m)} = n^{2H}. \tag{7}$$

Although in terms of the correlation time function the convergence to FN_H with any $H \in [0,1]$ can be described by a single equation ((7)) according to the value of H the processes which converge to FN_H show significantly different behaviour. Therefore different classes of examples will be constructed for $H \in \{0\}, (0, 0.5), \{0.5\}, (0.5, 1), \{1\}$.

The set autocorrelation functions which converge to a given autocorrelation function will be called the *domain of attraction* of this latter autocorrelation function. Similarly we define the domain of attraction of correlation time functions and also the domain of attraction of self-similar processes.

It has to be noted that the examples presented here do not cover the entire set of processes in a given domain of attraction.

Thesis 3.3 ($H = 0$, [P3])
I have shown that if for a process X $\lim_{k\to\infty} \rho(k)$ exists and $\rho(1) < 1$, then the differenced process $Y(i) = X(i+1) - X(i)$ is in the domain of attraction of FN_0.

Thesis 3.4 ($H \in (0, 0.5)$, [P3])
I have shown that if for a process $\sum_{i=-\infty}^{\infty} \rho(i) = 0$ and $\rho(k) \sim ck^{2H-2}$ then the process is in the domain of attraction of FN_H, if $H \in (0, 0.5)$.

Here the symbol '$\sim$' means asymptotic equivalence. The functions $f(x) \sim g(x)$ if and only if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.

Thesis 3.5 ($H = 0.5$, [P3])
I have shown that all processes for which $\sum_{i=-\infty}^{\infty} \rho(i) \in (0, \infty)$ are in the domain of attraction of $FN_{0.5}$, also called white noise. These processes are also often referred to as short-range dependent processes.

Thesis 3.6 ($H \in (0.5, 1)$, [P3])
I have shown that processes for which $\rho(k) \sim ck^{2H-2}$ are in the domain of attraction of FN_H, if $H \in (0.5, 1)$.

This class is closely related to the class of *long-range dependent* processes, which will be examined in detail in Section 4.3. For these processes $\sum_{i=-\infty}^{\infty} \gamma(i) = \infty$, showing that the influence of the past is very strong, the dependence decays very slowly.

Thesis 3.7 ($H = 1$, [P3])
Let Y be a random variable with zero mean and unit variance. Define the process $X = \{Y, aY, Y, aY, Y \cdots\}$, $a \in [-1, 1]$ where, using a fair coin independent of Y, we assign the origin of time to Y or aY to ensure stationarity. I have shown that X is in the domain of attraction of FN_1.

9

4.3 Long-range dependence

The so-called long-range dependent processes constitute an important subset of the asymptotically self-similar processes, because they appear in different scientific fields. The process that describes the intensity of network traffic is also long-range dependent.

Thesis 4   Review of long-range dependent processes

Several definitions of long-range dependence can be found in the literature [1, 4, 6], sharing many common properties and trying to capture the same phenomenon. It was found to be necessary to compare these definitions, and rigorously check and prove exactly which properties are satisfied according to the different definitions, especially since rigorous proofs for some important properties are impossible or very hard to find. Many definitions of long-range dependence rely on regular variation, whose properties and even its rigorous definition is in many cases omitted. Therefore finding the appropriate definition of discrete time regular varying functions and exploring its basic properties was relevant to the study of long-range dependent processes.

4.3.1   Discrete regularly varying functions

Regularly variation in continuous time has a well developed literature [2]. Many definitions of long-range dependence use a discrete version of regular variation. It has to be noted, however, that discrete time regularly varying functions do not share all the convenient properties of their continuous equivalent. Therefore discrete time regular variation should be treated in its own right, its properties have to be stated and proved. The definition I used for the discrete case, although not stated in the same form, is equivalent to the definition of [8].

Definition 4.4 (Continuous regular variation)
A function $\widetilde{f}$ defined on $\mathbf{R}^+$ is *regularly varying* at infinity with index α if

$$\lim_{t \to \infty} \frac{\widetilde{f}(tx)}{\widetilde{f}(t)} = x^\alpha, \quad \alpha \in \mathbf{R} \tag{8}$$

for every $x \in \mathbf{R}^+$ (it is sufficient that (8) is satisfied on a dense subset of $\mathbf{R}^+$, [7] page 275). If $\alpha = 0$ the function $\widetilde{f}$ is also said to be *slowly varying*. The set of continuous regular varying functions with index α is denoted by CRV_α, and the set of slowly varying functions is denoted by CSV.

Definition 4.5 (Discrete regular variation (DRV))
A function f defined on $\mathbf{Z}^+$ is *regularly varying* at infinity with index α if there exists a function $\widetilde{f} \in \mathrm{CRV}_\alpha$ such that $f(n) = \widetilde{f}(n)$ for all $n \in \mathbf{Z}^+$. The set of discrete regular varying functions with index α is denoted by DRV_α and the set of slowly varying functions is denoted by DSV.

Thesis 4.1 (Properties of discrete regularly varying functions, [P3])
Since regular variation in discrete time is derived from its continuous equivalent it shares many, but not all of its properties. I showed that it satisfies the following ones:

- $f \in DRV_\alpha \Rightarrow \lim_{k \to \infty} f(kn)/f(k) = n^\alpha$, $\alpha \in \mathbf{R}$, $n \in \mathbf{Z}^+$

- $f \in DRV_\alpha \Leftrightarrow f(k) = s(k)k^\alpha$, $s(k) \in DSV$

- $f \in DRV_\alpha$ and $g \sim f \Rightarrow g \in DRV_\alpha$

- $f \in DRV_\alpha \Rightarrow f(k) \sim f(k + k_0)$, $\forall k_0$ constant.

- Let $K(n) \in DRV_\alpha$, and let $L(t)$ and $U(t)$ be defined as

$$L(m) := \sum_{n=0}^{m-1} K(n), \qquad\qquad U(m) := \sum_{n=m}^{\infty} K(n).$$

(a) If $\alpha \geq -1$ then $\dfrac{mK(m)}{L(m)} \to (1 + \alpha)$,   and $L \in DRV_{\alpha+1}$.

(b) If $\alpha < -1$ then $\dfrac{mK(m)}{U(m)} \to -(1 + \alpha)$,   and $U \in DRV_{\alpha+1}$.

4.3.2   Definitions of long-range dependence

Here I present some of the different definitions of long-range dependence that can be found in the literature, and propose a new definition that extends the commonly found definitions to include more processes, but preserves the spirit of the traditional definitions. More definitions are presented in the dissertation. In 4.3.3 these different definitions, their relationship and their important properties will be investigated.

Definition 4.6 (LRD1)
LRD1 processes are those whose autocovariance functions obey $\gamma(k) \sim c_\gamma k^{2H-2}$, $H \in (0.5, 1)$, $c_\gamma \in \mathbf{R}^+$.

This definition is the most frequently encountered. For example it is used in [1, 4].

Definition 4.7 (LRD2)
LRD2 processes are those whose autocovariance functions obey

$$\gamma(k) = c_\gamma(k)k^{2H-2}, \tag{9}$$

where $H \in (0.5, 1)$, $c_\gamma \in$ DSV.

This choice generalises LRD1 in a natural way, by replacing the constant c_γ, a particular slowly varying function, with a general slowly varying function.

Definition 4.8 (LRD3)
LRD3 processes are those whose covariance sums obey $\sum_{k=1}^{\infty} \gamma(k) = \infty$.

This definition, used for example in [11], nicely captures the idea of LRD, being when the sum of the past has a strong impact.

Thesis 4.2 (New definition of long-range dependence, [P3])
Processes in the domain of attraction of FN_H with $H \in (0.5, 1)$ are called long-range dependent.

This definition, as it will be shown, extends definitions LRD1 and LRD2, but preserves the idea behind those definitions, namely that these processes are asymptotically equivalent and converge to a fractional noise process, with non-summable covariance sum.

4.3.3 Examining long-range dependent properties

Thesis 4.3 (LRD2 has slowly decaying variance, [P3])
I have shown that for LRD2 processes

$$\mathcal{V}^{(m)} \sim \frac{c_\gamma(m) m^{2H-2}}{H(2H-1)}. \tag{10}$$

This statement has appeared in different places including [1], but a rigorous proof was not found. This result forms the basis of a long-range dependence estimating tool, the variance time plot, which tries to detect the presence of long-range dependence indirectly through estimating the aggregated variance function, $\mathcal{V}^{(m)}$.

Thesis 4.4 (LRD2 is a subset of LRD, [P3])
I have shown that LRD2 processes form a subset of LRD processes, that is they are in the domain of attraction of FN_H.

This statement follows easily from Thesis 4.3 and Equation (7), showing the advantage of the correlation time function approach.

Thesis 4.5 (LRD2 is a *strict* subset of LRD, [P3])
The question whether the classes LRD2 and LRD are equivalent are usually not explicitly investigated, this question is generally omitted. By constructing an example in the set $LRD \setminus LRD2$ I have shown that these classes are not equivalent, LRD includes much more then LRD2.

5 Practical implications

Some general considerations about the suitability of the variance time function for the analysis of processes on multiple time scales is presented in the dissertation. There it is demonstrated how the new approach helped to reveal unknown details of the otherwise well-know fARIMA processes. Here it is presented that a clear view of the properties of the different types of long-range dependent processes (LRD1, LRD2, LRD3, LRD) contributes to the correct interpretation of the long-range dependent testing and parameter estimation methods.

As a simple example consider the variance time plot method, which is based on the slow decay of the aggregated variance of LRD processes as described in Thesis 4.3.

If the process being analysed is assumed to be LRD1 then the $c_\gamma(m)$ of (9) converges to a constant (c_γ). So Equation (10) can be written as

$$\mathcal{V}^{(m)} \sim \frac{c_\gamma m^{2H-2}}{H(2H-1)}. \tag{11}$$

Therefore the aggregated variance $\mathcal{V}^{(m)}$ follows a simple power-law for large values of m. So plotting $\mathcal{V}^{(m)}$ against m on a log-log scale the tail of the plot should align a straight line with a slope of $2H - 2$. This is depicted in Figure 3.
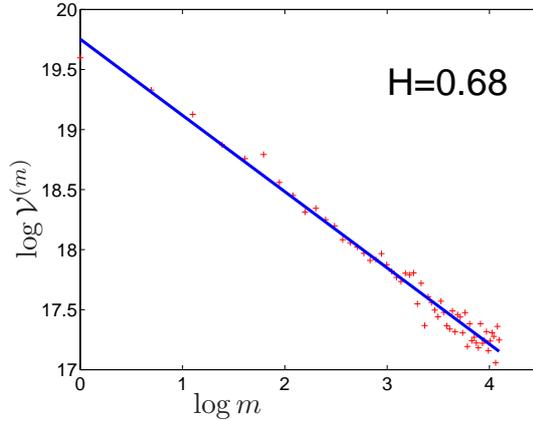
Figure 3: Parameter estimation using the Variance Time Plot method

This method is used for both detecting the presence of LRD(1) and also estimating H.

Because of the statistical nature of the process and the finite sample size the estimation will always have some inaccuracy, that can be decreased by the appropriate choice of the estimator, but can never be totally eliminated. Another important practical problem is the selection of the lower cut-off scale (m) such that Equation (11) can be applied with a reasonable level of accuracy. These issues will, however, not be investigated here. This section focuses on some theoretical aspects, and assumes that the sample size is large enough in order to get rid of the disturbing and misguiding effects.

If the process under investigation was LRD1 then the tail of the plot will align a straight line with slope $2H - 2$, so H can be estimated, by fitting a line to the tail.

The algorithmic steps of the estimation and detection are the following:

1. Estimate $\mathcal{V}^{(m)}$ for $m = 1, 2, \ldots, m_{\max}$, where $m_{\max} << n$ and n is the sample size[3].

2. Draw a plot of $\log \mathcal{V}^{(m)}$ against $\log m$.

3. Fit a straight line to the tail of the slope

4. Can the line be fitted? If no then reject the hypothesis of LRD(1).

---

[3]As m is increasing we have less and less samples from the process $X^{(m)}$ so the estimation of $\mathcal{V}^{(m)}$ is less and less reliable.

5. If yes then conclude LRD1 and measure the slope and estimate H.

We have seen that this method works well for LRD1 processes, that is the assumption of LRD1 is not rejected and also the parameter can be estimated. We now investigate the behaviour of the variance time plot for non-LRD1 processes.

Assume that the process under investigation is LRD2 but not LRD1. In this case the c_γ function of Equation (10) does not converge to a constant which means that no matter how large m is chosen the plot of $\log \mathcal{V}^{(m)}$ against $\log m$ will not converge to a straight line. So because of the lack of the straight line the estimation method will not detect the presence of LRD2.

We have seen that the appearance of a straight line is not a necessary condition for being LRD(2) but on the other hand it is necessary for being LRD1. The question we will investigate now is whether it is also sufficient for being LRD1 or not.

From the presence of the straight line we can conclude that Equation (11) holds, which is a necessary, but, as can be shown, not sufficient condition for the process to be LRD1. So the straight line does not prove the presence of LRD1.

This simple example shows that it is important to know the connections of the different LRD properties in order to interpret the results of the estimators correctly. Similar considerations apply to other LRD testing and estimation methods too.

Acknowledgement

References

[1] J. Beran. *Statistics for Long-Memory Processes*. Chapman and Hall, New York, 1994.

[2] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Cambridge University Press, Cambridge England, 1987.

[3] P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods*. Springer, 1996.

[4] D. R. Cox. *in Long-range dependence: a review*, H.A. David and H.T. David editors, chapter Statistics : an Appraisal, pages 55–74. Iowa State University Press, Ames (IA), 1984.

[5] Mark E. Crovella and Azer Bestavros. Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes. In *ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Networks*, May 1996.

[6] D. L. Jagerman and B. Melamed and W. Willinger. Stochastic Modeling of Traffic Processes. In *Frontiers in Queueing: Models, Methods and Problems*. CRC Press, 1996.

[7] William Feller. *An Introduction to Probability Theory and its Applications*, volume II. John Wiley and Sons, Brisbane, second edition, 1970.

[8] J. Galambos and E. Seneta. Regularly Varying Sequences. *Proceedings of the American Mathematical Society*, 41(1):110–116, November 1973.

[9] P. Major. *Multiple Wiener-Itô Integrals*, volume 849 of *Springer Lecture Notes in Mathematics*. Springer-Verlag, New York, 1981.

[10] Kihong Park, Gi Tae Kin, and Mark E. Crovella. On the Relationship Between File Sizes, Transport Protocols, and Self-Similar Network Traffic. Technical report, Computer Science Department, Boston University, 1996. TR-96-016.

[11] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes*. Chapman and Hall, 1994.

[12] Ya. G. Sinai. Self-Similar Probability Distributions. *Theory of Probability and its Applications*, 21:64–80, 1976.

Publications

[P1] Gefferth András. WWW információs rendszer az Austria Telecomnál (WWW Information system at Austria Telecom). *Magyar Távközlés (Hungarian Telecommunications)*, February 1996.

[P2] A. Gefferth, S. Molnár, and D. Veitch. Discrete Self-Similarity. In *IFIP WG6.7 Workshop and EUNICE Summer School on Adaptable Networks and Teleservices*, Trondheim, Norway, September 2-4 2002,pp. 55-61.

[P3] A. Gefferth, D. Veitch, I. Maricza, S. Molnár, and I. Ruzsa. The Nature of Discrete Second-Order Self-Similarity. *Advances in Applied Probability*, 35(2), June 2003,pp. 395-416.

[P4] A. Gefferth, D. Veitch, I. Ruzsa, I. Maricza, and S. Molnár. A New Class of Second Order Self-Similar Processes. *Stochastic Models*, 20(3):381–389 September 2004,pp 381-389.

[P5] S. Molnár and A. Gefferth. On the Scaling and Burst Structure of Data Traffic. In *8th International Conference on Telecommunication Systems, Modelling and Analysis*, Nashville, Tennessee, USA, March 2000.