

LRD of Fractional Brownian Motion and Application in Data Network

Bole Yang

Advisor: Prof. Flaschka

The University of Arizona

Program in Applied Mathematics

May 2008

Abstract

Long-range dependence (LRD) can be observed in the large scale behavior of almost all the data network traffics, but the explanation of this phenomenon is not good if we still use models without considering the current network flows as multi-layer ones. Thus in this paper the relevant idea including fractional Brownian motion (FBM) and self-similarity will be introduced first. Then an idealized on-off source model will be established to explain the LRD phenomenon at the application layer of the common data network traffic, especially the global Internet.

1 Introduction

Mathematically, long-range dependent (LRD) processes ([14]) are characterized by their auto-covariance functions. The decay of the auto-covariance function follows a power law and decays more slowly than exponential ones, and the area under the curve is infinite. Long-range dependence is closely related to the self-similarity property and it is easily observed in nature, such as economics, hydrology, etc. We are interested in the application of long-range dependence to the area of communication networks.

There has been a big advancement in the practical studies of high quality and high volume data sets of traffic measurements ([11]) for the past several decades, especially of different links within the most common used Internet. They are implemented to describe statistical characteristics of the dynamics of 'packet flows' (or called 'bit rate processes' [15]). Those studies demonstrated long-range dependence nature of the measured packet traffic. Thus when viewed within a large scale, the traffic appears to be self-similar. In this paper, we will put an effort in exploiting the effects of human interactions within the network, and providing mathematical explanation to the observed long-range dependence phenomenon that is related to the implicit structure of the network ([13]).

In the next section we will introduce the stochastic process and the 'regular' Brownian motion (BM). Section 3 will be devoted to the mathematical concept of self-similarity, the fractional Brownian motion (FBM) and its typical properties, and long-range dependence as well. Then starting from a brief sketch in Section 4.1 of the basic design principles behind the modern data networks, especially the notion of predominant transfer protocol-TCP/IP protocol architecture, we construct a model of data network traffic at the application level in Section 4.2 to formulate some new and relevant research problems about LRD and head towards a further understanding of the users' influence upon the highly complex nature of packet flows in large scale data networks (e.g. Internet).

2 Brownian Motion

This erratic motion is named after its physics discovery, Robert Brown. See the computer simulation Figure 1 and Figure 2. Besides, Norbert Wiener proved the existence ¹ of this type of motion in a mathematical way. Brownian motion is one of the simplest continuous-time stochastic processes. The universality of Brownian motion is closely related to the universality of the normal distribution ².

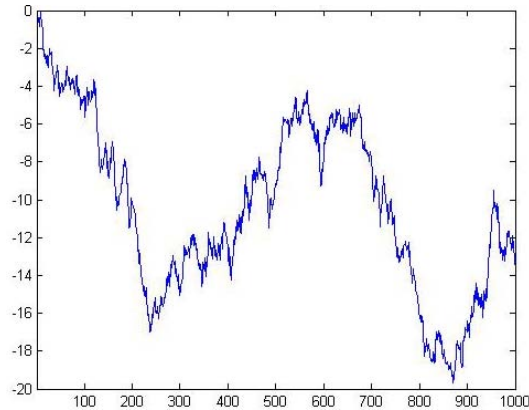


Figure 1: A Matlab illustration of one-dimensional Brownian motion.

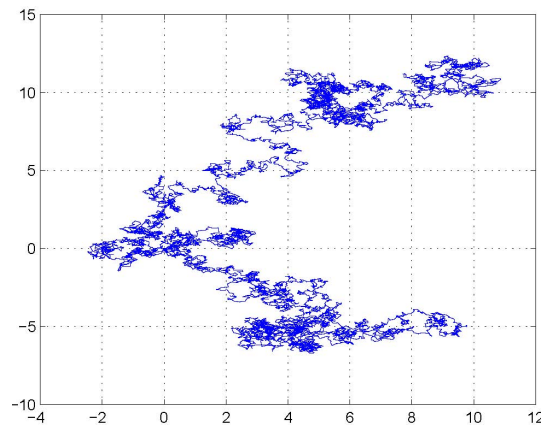


Figure 2: A Matlab illustration of two-dimensional Brownian motion.

¹The existence means that the finite-dimensional distributions define, consistently, a probability measure on the space of continuous functions.

²Let X_1, X_2, X_3, \dots be a set of n independent and identically distributed random variables having finite values of mean μ and variance $\sigma^2 > 0$. The central limit theorem states that as the sample size n increases, the distribution of the sample average approaches the normal distribution irrespective of the shape of the original distribution, i.e.,

$$\frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to $N(0,1)$ in distribution.

2.1 Stochastic Process

A stochastic process is the counterpart to a deterministic process in probability theory. Instead of dealing with merely one possible 'reality' of how the process might evolve under time, in a stochastic process there is some indeterminacy in its future evolution described by probability distributions. This means that even if the initial condition is known, there are still many possibilities the process might go to, but some paths are more probable and others less. Precisely,

Definition 1 *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a stochastic process (random process) with state space \mathcal{F} , is a collection of \mathcal{F} -valued random variables indexed by a set T ('time'). That is, a stochastic process X is a collection $\{X(t), t \in T\}$, where each $X(t)$ is an \mathcal{F} -valued random variable. Also, the mean of the process is*

$$\mu(t) = E[X(t)]$$

and the auto-covariance of the process is defined as

$$r(s, t) = Cov(X(s), X(t)) = E((X(s) - E(X(s)))(X(t) - E(X(t))))$$

□

For example, a stochastic process can amount to a sequence of random variables known as time series. Another basic type of a stochastic process is the random field, whose domain is a region of space (see [2]).

We need two more concepts of processes ([1]): so called 'Gaussian processes' and 'Stationary processes'. Usually it is not possible to determine a stochastic process only with the mean and the auto-covariance of a process, but there is an exception:

Definition 2 *A Gaussian process is a stochastic process which generates samples over time $\{X(t), t \in T\}$ such that no matter which finite linear combination of the $X(t)$ one takes (or, generally, any linear functional of the sample function $X(t)$), that linear combination will be normally/Gaussianly distributed.*

□

The mean vector of Gaussian process $X(t_1), \dots, X(t_n)$ is the vector $\mu(t_1), \dots, \mu(t_n)$ with component $\mu(t_j)$. The auto-covariance matrix of $X(t_1), \dots, X(t_n)$ is the matrix with entries $r(t_i, t_j)$. Thus it is clear that the mean vector and the auto-covariance matrix of jointly Gaussian random variables determine the density function, i.e., a Gaussian process can be uniquely determined by its mean and auto-covariance function.

Definition 3 *A process is stationary if for every function $f(x_1, x_2, \dots, x_n)$ the expectations can be defined as below:*

$$E[f(X(t_1), \dots, X(t_n))] = E[f(X(t_1 - s), \dots, X(t_n - s))]$$

□

Apparently, a stationary process should be considered as a stochastic process whose joint probability distribution does not change when shifted in time or space.

2.2 Brownian Motion

One of the most popular ways to mathematically introduce the Brownian motion is to consider it as a limit of a symmetric random walk process (see [1]). Roughly speaking, the random walk is a mathematical formalization of a trajectory that consists of successive steps in random directions. We might simply imagine repeatedly tossing a coin and carefully recording heads and tails to find out what is the balance.

Definition 4 *Divide time into intervals of length $\Delta t > 0$. Correspondingly, divide the space into intervals of length*

$$\Delta x = \sigma\sqrt{\Delta t}$$

Assign a random variable ξ that is ± 1 with probabilities 0.5 for either sign to each interval. Then for $t = n\Delta t$

$$W_t = \xi_1\Delta x + \xi_2\Delta x + \dots + \xi_n\Delta x$$

is called the random walk. If $n \rightarrow \infty$ and $\Delta t \rightarrow 0$ with $t = n\Delta t$ fixed, then the result limit is called the Brownian Motion (BM), denoted as $B(t)$. Sometimes the Brownian motion is also called Wiener process. \square

I take the risk not to show the existence of the Brownian motion.

The Brownian motion has some useful properties. Because W_t has mean zero and variance $n(\Delta x)^2 = \sigma^2 t$, it is not difficult to find that $B(t)$ has mean zero and variance $\sigma^2 t$ and is Gaussian. Several other claims about the Brownian motion are listed below:

Theorem 1 *For a Brownian motion $B(t)$:*

1. $B(0) = 0$.
2. $B(t)$ is a Gaussian process with mean 0 and auto-covariance function $EB(t_1)B(t_2) = \min(t_1, t_2)$
3. For each $s < t$ the random variable $B(t) - B(s)$ is Gaussian with mean zero and variance $\sigma^2(t - s)$. Besides, $B(t) - B(s)$ corresponding to disjoint intervals are independent.
4. There is a dense set T in $[0, \infty)$ such that $B(t)$ is uniformly continuous on $T \cap [0, a]$ for all a and for almost every $\omega \in \Omega$. Also almost every Brownian motion path is nowhere differentiable.

\square

Proof:

I will briefly prove the second property.

Actually, for Brownian motion, a continuous Gaussian process, its probability of the event $(B(t_1), \dots, B(t_n)) \in X$ is given by (see [2]):

$$P_n(X) = [(2\pi)^n t_1(t_2 - t_1) \dots (t_n - t_{n-1})]^{-1/2} \int_X e^{-\frac{1}{2}[\frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1} + \dots + \frac{(x_n - x_{n-1})^2}{t_n - t_{n-1}}]} dx_1 \dots dx_n$$

Suppose $t < s$, then the inverse of the auto-covariance matrix satisfies:

$$(\Lambda x, x) = \frac{x_1^2}{t} + \frac{(x_2 - x_1)^2}{s - t}$$

so the matrix Λ should be

$$\Lambda = \begin{vmatrix} \frac{s}{(s-t)t} & -\frac{1}{(s-t)} \\ -\frac{1}{(s-t)} & \frac{1}{(s-t)} \end{vmatrix}$$

that is,

$$r = \Lambda^{-1} = \begin{vmatrix} t & t \\ t & s \end{vmatrix}$$

then for $t < s$,

$$r(t, t) = t; \quad r(t, s) = t; \quad r(s, t) = t; \quad r(s, s) = s$$

Thus by symmetry, for $t, s > 0$, $r(t, s) = \min(t, s)$. □

3 Fractional Brownian Motion

As what was discussed in Section 2, we typically think of the Brownian motion as a type of random walks. In fact, we can extend the idea of semi-random walks to that of 'biased' random walks, i.e., the walks in which the steps are not independent, but each step depends on all previous steps instead. Again, the limit of such a type of biased random walk introduces a generalization of Brownian motion—the so-called fractional Brownian motion. Nevertheless, instead of deriving it within the frame of random walks, I choose a more general way.

3.1 Self Similarity

Naively, self-similarity is a typical property of fractals. A self-similar object is exactly or approximately similar to a part of itself, i.e., the whole has the same shape as one or more of the parts. Many objects in the real world are statistically self-similar, such as plants and coastlines, see Figure A ([16]) and Figure B ([8]). There are also man-made ones. A classic example is the triadic Koch curve, see Figure C ([16]). Another one is the Sierpinski triangle, shown in Figure D ([16]).

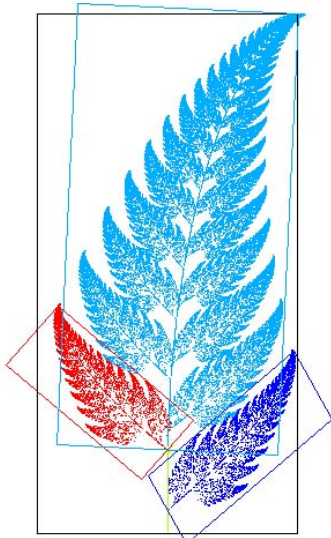


Figure A: Self-similarity in a fern

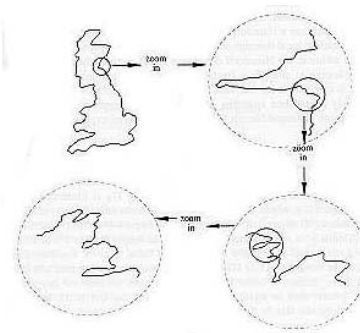


Figure B: Self-similarity in a coastline

Mathematically, let T be either \mathbb{R} , $\mathbb{R}_+ = \{t : t \geq 0\}$ or $\{t : t > 0\}$:

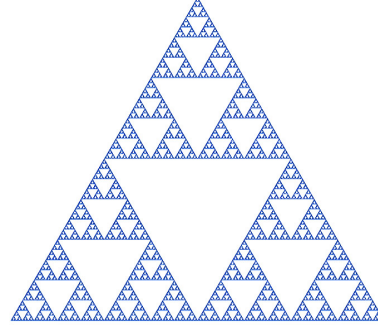
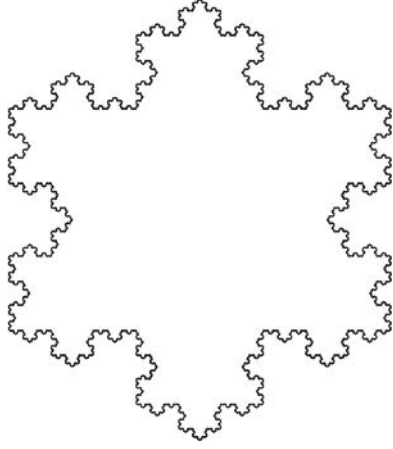


Figure C: Self-similarity in Koch curve Figure D: Self-similarity in Sierpinski triangle

Definition 5 The real-valued process $\{X(t), t \in T\}$ is called self-similar with index $H > 0$ (*H-ss*) if for all $a > 0$, the finite-dimensional distributions of $\{X(at), t \in T\}$ are identical to the finite-dimensional distributions of $a^H X(t), t \in T$:

$$(X(at_1), \dots, X(at_d)) = (a^H X(t_1), \dots, a^H X(t_d))$$

Here H is also called Hurst exponent which is supposed to be positive. □

Since the Brownian motion $\{X(t), t \in \mathbb{R}\}$ is a Gaussian process with mean 0 and auto-covariance function $EX(t_1)X(t_2) = \min(t_1, t_2)$ (by Theorem 1), then for all $a > 0$,

$$EX(at_1)X(at_2) = \min(at_1, at_2) = a \min(t_1, t_2) = E(a^{0.5} X(t_1))(a^{0.5} X(t_2))$$

Thus Brownian motion is reasonably H-ss with $H = 0.5$. Furthermore, we shall introduce the H-ss processes with stationary increments which are of great application interests because they will lead to stationary sequences with remarkable features (see [7]).

Definition 6 A real-valued process $\{X(t), t \in T\}$ has stationary increments if

$$\{X(t+h) - X(h), t \in T\} = \{X(t) - X(0), t \in T\} \quad \forall h \in T$$

□

Definition 7 The process is called H-sssi if it is self-similar with index H and has stationary increments. □

With the definition, we get a nice property below [7]:

Theorem 2 If $\{X(t), t \in \mathbb{R}\}$ is a (non-degenerate) H-sssi finite variance process. Then

$$0 < H \leq 1, \quad X(0) = 0 \text{ a.s. (almost surely)}$$

and

$$\text{Cov}(X(t_1), X(t_2)) = \frac{1}{2} \{|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\} \text{Var} X(1)$$

$\forall t \in \mathbb{R}$, we have:

$$EX(t) = 0 \text{ if } 0 < H < 1$$

$$X(t) = tX(1) \text{ if } H=1$$

□

To prove it, we need some lemmas:

Lemma 1 *If $\{X(t), t \in T\}$ is H -ss and $0 \in T$, then for each $a > 0$, $X(0) = X(a \cdot 0) = a^H X(0)$ and hence $X(0) = 0$ a.s.* □

Lemma 2 *If $\{X(t), t \in \mathbb{R}\}$ is H -sssi, then for fixed t , $X(-t) = -X(t)$.* □

Proof:

This is because $X(0) = 0$ and $X(-t) = X(-t) - X(0) = X(0) - X(t) = -X(t)$.

Lemma 3 *If $X(t), t \in \mathbb{R}$ is H -sssi and*

$$P(X(1) \neq 0) > 0 \tag{1}$$

then the relation

$$E|X(1)|^\gamma < \infty$$

implies

$$0 < H < 1/\gamma \text{ if } 0 < \gamma < 1$$

and it implies

$$0 < H \leq 1 \text{ if } \gamma \geq 1$$

□

Proof:

When $0 < \gamma < 1$, then $|X(2)|^\gamma \leq |X(2) - X(1)|^\gamma + |X(1)|^\gamma$ which is strict on $\{X(1) \neq 0, X(2) - X(1) \neq 0\}$. (1) implies that $P(\{X(1) \neq 0, X(2) - X(1) \neq 0\}) > 0$, thus:

$$E|X(2)|^\gamma < E|X(2) - X(1)|^\gamma + E|X(1)|^\gamma \tag{2}$$

By the stationarity of the increments (Definition 6) and $X(0) = 0$, we have (in the sense of distribution) $X(2) - X(1) = X(1) - X(0) = X(1)$. Thus (2) means

$$E|X(2)|^\gamma < 2E|X(1)|^\gamma$$

But $E|X(2)|^\gamma < 2^{H\gamma}E|X(1)|^\gamma$ by self-similarity. So $H < 1/\gamma$

When $\gamma \geq 1$, then $E|X(1)|^p < \infty (\forall p < 1)$. Therefore $0 < H < 1/p (\forall p < 1)$, implying $0 < H \leq 1$. □

Let's go back to the original theorem.

Proof of Theorem 2:

Due to the stationarity of the increments and self-similarity,

$$\begin{aligned} EX(t_1)X(t_2) &= \frac{1}{2}\{EX^2(t_1) + EX^2(t_2) - E(X(t_1) - X(t_2))^2\} \\ &= \frac{1}{2}\{EX^2(t_1) + EX^2(t_2) - E(X(t_1 - t_2) - X(0))^2\} \\ &= \frac{1}{2}\{|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\}EX^2(1) \end{aligned} \tag{3}$$

because $X(0) = 0$ by Lemma 1, $X(1) = -X(-1)$ by Lemma 2 and $X(t) = |t|^H X(\text{sgnt})$ for fixed t (all in the sense of distribution). Since $\{X(t), t \in \mathbb{R}\}$ is a finite variance process with mean zero, we have $E|X(1)|^\gamma < \infty$ where $\gamma = 2 > 1$ here. Thus by Lemma 3, $0 < H \leq 1$.

Now, suppose $0 < H < 1$. Since

$$EX(1) = E(X(2) - X(1)) = 2^H EX(1) - EX(1) = (2^H - 1)EX(1)$$

we have $EX(1) = 0$. So, $EX(t) \equiv 0$, due to $EX(-1) = -EX(1)$ and $EX(t) = |t|^H EX(\text{sgnt})$. Thus the form

$$R_H(t_1, t_2) = \text{Cov}(X(t_1), X(t_2)) = \frac{1}{2} \{|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\} \text{Var}X(1)$$

of the auto-covariance function is true.

If $H = 1$, then (3) gives $EX(t_1)X(t_2) = t_1 t_2 EX^2(1)$ and thus

$$\begin{aligned} E(X(t) - tX(1))^2 &= EX^2(t) - 2tEX(t)X(1) - t^2EX^2(1) \\ &= (t^2 - 2t^2 + t^2)EX^2(1) \\ &= 0 \end{aligned}$$

So, $X(t) = tX(1)$ a.s. for all t . □

3.2 Fractional Brownian Motion

If we add a restriction to an H-sssi process defined in the previous subsection, it is convenient, therefore, to define the fractional Brownian motion as follows:

Definition 8 *A Gaussian H-sssi process, $0 < H \leq 1$, is called fractional Brownian motion (FBM) and will be denoted $\{B_H(t), t \in T\}$.* □

When $H = 0.5$, it is rightly the Brownian motion. Figure 3 and Figure 4 correspond to the cases of $H < 0.5$ and $H > 0.5$:

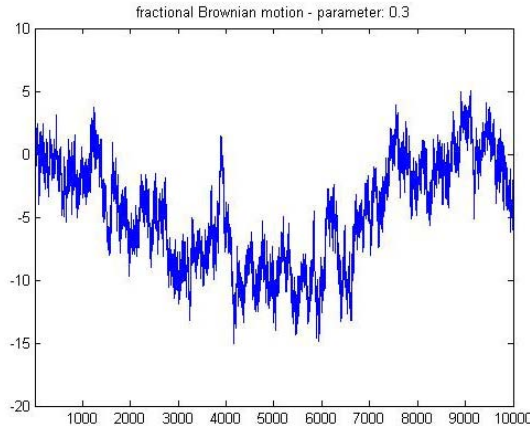


Figure 3: Fractional Brownian motion with H=0.3

There are also non-Gaussian finite variance H-sssi processes (see [5]), but we only consider the Gaussian ones in this paper.

The following corollary is an immediate consequence of Definition 8 and Theorem 2 in Subsection 3.1.

Corollary 1 *Fix $0 < H \leq 1$ and $\sigma_0^2 = EX^2(1)$. The following statements are equivalent:*

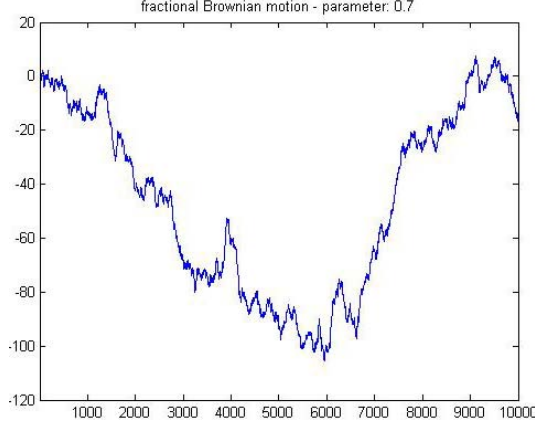


Figure 4: Fractional Brownian motion with $H=0.7$

1. $\{X(t), t \in T\}$ is Gaussian and H -sssi.
2. $\{X(t), t \in T\}$ is fractional Brownian motion with self-similarity index H .
3. $\{X(t), t \in T\}$ is Gaussian, has mean zero (if $H < 1$) and auto-covariance function

$$R_H(t_1, t_2) = Cov(X(t_1), X(t_2)) = \frac{1}{2}\{|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\}VarX(1)$$

□

It also leads to a result which will be used in Section 4:

Corollary 2 Suppose that $\{X(t), t \in \mathbb{R}\}$

1. is Gaussian with mean zero, $X(0) = 0$
2. $EX^2(t) = \sigma^2|t|^{2H}$ for some $\sigma > 0$ and $0 < H < 1$
3. has stationary increments

then $\{X(t), t \in \mathbb{R}\}$ is fractional Brownian motion

□

3.3 Long-Range Dependence

Fractal statistics ([12]) from natural processes often exhibit two properties: (1) the tendency to have wide range of variation occur in discrete jumps; (2) propensity for time series to have statistical trends and cycles while still remaining chaotic. Benoit Mandelbrot conceived two vivid name for them (see [6]): Christened Noah Effect for the first case and Joseph Effect for the second, respectively, honoring two Biblical figures: the hero of the forty-day flood, and the hero of the seven bountiful years followed by seven years of famine.

Since fractional Brownian motion is H -sssi, for any $0 < H < 1$ applying Kolmogorov's criterion ³ to FBM $B_H(t)$ with $p \geq 1/H$, we get ([10]):

$$E|B_H(t_2) - B_H(t_1)|^p = E|B_H(1)|^p|t_2 - t_1|^{pH} \quad (4)$$

³Kolmogorov(1903-1987)'s criterion: a stochastic process $\{X(t), t \in \mathbb{R}\}$ has a version with continuous paths if there are $p \geq 1$, $\eta > 1$ and a constant c such that, $\forall t_1, t_2$, we have $E|X(t_2) - X(t_1)|^p \leq c|t_2 - t_1|^\eta$.

which means the paths of FBM are continuous. Besides,

$$E\left|\frac{B_H(t_2) - B_H(t_1)}{t_2 - t_1}\right|^2 = \sigma^2|t_2 - t_1|^{2H-2} \rightarrow \infty, \text{ when } t_1 \rightarrow t_2$$

in the sense of $L^2(\Omega)$ means that FBM is not differentiable. Moreover, with the above equation, we may derive that the paths of FBM get less zigzagged as H goes from 0 to 1. Specifically, when $0 < H < 1/2$, the auto-covariance of the increments of FBM is negative, so the increments of FBM have opposite signs. Thus the particle tends to return. Such FBM is called anti-persistent. On the contrary, when $1/2 < H < 1$, the auto-covariance of the increments of FBM is positive, so the particle tends to insist on the same direction. Such FBM is called persistent. These phenomena can also be noticed in Figure 3 and Figure 4.

Obviously, the Joseph effect well agrees with those two category of fractional Brownian motion. In fact, with persistence and anti-persistence characteristics, the fractional Brownian motion has infinite dependence on the past, so it does not separate neatly into a periodic trend component and a random noise component. Thus in order to interpolate the expected value of the function at one point, we have to know the value at the times long before or after the two ending time points. This is also a big difference between the fractional Brownian motion and the regular Brownian motion.

The persistence and anti-persistence properties are together known as long-range dependence (LRD). Precisely ([10]),

Definition 9 Let $\{X_k\}$ be stationary sequences with finite variance and auto-covariance function $\gamma(k) = EX_0X_k$. Let $f(\nu) = (2\pi)^{-1}\sum_{k=-\infty}^{\infty}e^{-i\nu k}\gamma(k)$ with $\nu \in [-\pi, \pi]$ The following three definitions of LRD are common:

1. $\sum_{k=-n}^n\gamma(k) \sim n^\alpha L_1(n)$, as $n \rightarrow \infty$, and $0 < \alpha < 1$
2. $\gamma(k) \sim k^{-\beta} L_2(k)$, as $k \rightarrow \infty$, and $0 < \beta < 1$
3. $f(\nu) \sim |\nu|^{-\gamma} L_3(|\nu|)$, as $\nu \rightarrow 0$, and $0 < \gamma < 1$

Here L_1, L_2 and L_3 are so-called slowly varying functions ⁴(L_1 and L_2 at infinity and L_3 at zero). \square

According to the above discussion, self-similar processes with stationary increments can be described with long-range dependence, i.e., they form a stationary process which can display long-range dependence. Conversely, applying central limit theorem to a stationary process will yield a self-similar process with stationary increments. Example of those processes include traffic processes such as the packet inter-arrival times which will be discussed further in Section 4.

4 Application in a Data Network Traffic Model

In [4], the authors established the experimental evidence of self-similar nature of Ethernet traffic in a statistically rigorous manner with the help of Ethernet LAN traffic data. Figure 5 is a pictorial 'proof' of self-similarity (also see [4]). In detail, the first column is Ethernet traffic (packets per time unit) on five different time scales. The second column is synthetic traffic from an appropriately chosen traditional model on the same five different time scales.

⁴A function L is slowly varying at zero (resp. infinity) if it is bounded on a finite interval and if, for all $a > 0$, $L(ax)/L(x)$ tends to 1 as x tends to zero (resp. infinity).

Actually, long-range dependence is observed in the large scale behavior of most data network traffic today, in particular related to the dynamic nature of packet flows in high-speed data networks such as the common Internet.

Since long-range dependence mostly deals with large-time scales, in this part we will explain the observed LRD phenomena with respect to the effect of user application, i.e., with the features obtained at the highest layer after decomposing the measured data network into separate levels. Thus, first we need to look into some rudimental idea of the internet protocol suite.

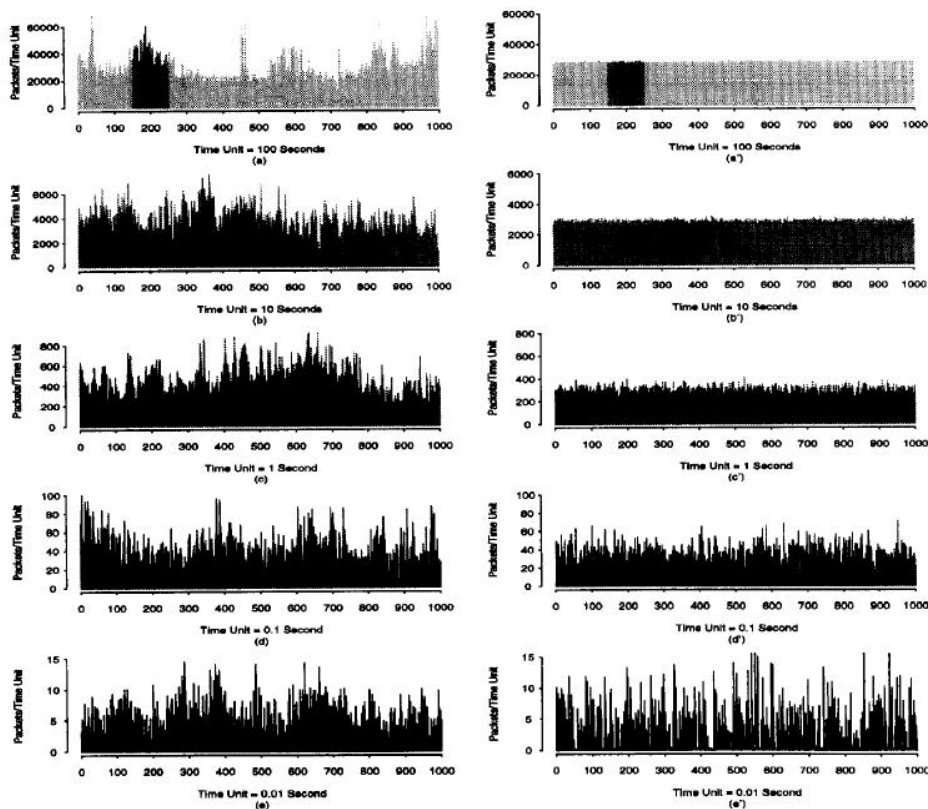


Figure 5: Pictorial evidence of self-similarity in Ethernet traffic

4.1 Internet Protocol Suite

Today when we transfer the file between the end hosts within the networks (e.g., the global Internet), the tasks are usually broken into several relatively simple subtasks which are handled at separate layers/levels ([13]). We assign each of the layers with a certain rule usually called 'protocol'. Each layer solves a set of problems. Upper layers are logically closer to the user and deal with more abstract data, relying on lower layer protocols to translate data into forms that can eventually be physically transmitted.

The most commonly used protocol stack in the global Internet is the Internet protocol suite (also called TCP/IP protocol) which can be typically viewed as four layers (from lower ones to upper ones): data-link layer, network layer, transport layer and application layer.

At the highest level, the application layer, the traffic is described in terms of session arrivals, duration and size. It is where most common network programs reside. These programs and their corresponding protocols include HTTP (The World Wide Web), FTP (File Transport), Telnet (Remote Login). In this section, our major interest lies in this level. We will see how user and application characteristics contribute to the LRD nature of data network traffics.

4.2 The On-Off Source Model

Let us consider an ideal model of network traffic at the application layer ([11]), assuming that the packets are sent at a constant rate. The users/sources cycle through periods of activity and inactivity. They communicate with each other within such networks where each individual source is modeled according to a stationary process $\{W(t), t \geq 0\}$:

$$W(t) = \begin{cases} 1 & \text{if time } t \text{ is an ON interval} \\ 0 & \text{if time } t \text{ is an OFF interval} \end{cases}$$

$W(t) = 1$ means there is a packet at time t and $W(t) = 0$ means that there is no packet then, so the source alternates between an active state and an idle one. Notice that an OFF interval always follows an ON interval. This forms 'packet trains'. We also assume that the lengths of the ON intervals are independently and identically distributed (i.i.d) and those of the OFF intervals are also i.i.d. The lengths of the ON periods and that of the OFF ones are independent as well, but those two kinds do not necessarily have identical distribution.

To specify the distributions of the ON and OFF periods, we use

$$f_1(x), \quad F_1(x) = \int_0^x f_1(u)du, \quad F_{1c} = 1 - F_1(x),$$

$$\mu_1 = \int_0^\infty x f_1(x)dx, \quad \sigma_1^2 = \int_0^\infty (x - \mu_1)^2 f_1(x)dx$$

to denote the probability density function, cumulative distribution function, tail distribution, mean length and variance of an ON interval. $f_2, F_2, F_{2c}, \mu_2, \sigma_2^2$ are respectively chosen for an OFF interval. Assume that as $x \rightarrow \infty$,

$$\text{either } F_{1c}(x) \sim l_1 x^{-\alpha_1} L_1(x) \text{ with } 1 < \alpha_1 < 2 \text{ or } \sigma_1^2 < \infty$$

and

$$\text{either } F_{2c}(x) \sim l_2 x^{-\alpha_2} L_2(x) \text{ with } 1 < \alpha_2 < 2 \text{ or } \sigma_2^2 < \infty$$

where $l_j > 0$ is a constant and $L_j > 0$ is a slowly-varying function as was talked about in Section 3.3 (we can also make it asymptotic to a constant for simplification). Notice the distribution of ON period has finite mean but infinite variance, so when $1 < \alpha_1 < 2$, the ON period can be very long with relatively high probability. At this moment, we call the distribution of the ON times as 'heavy-tailed' with exponent α_1 . Otherwise, we refer to the case $\sigma^2 < \infty$ as $\alpha_1 = 2$.

Moreover, suppose that there are M independently and identically distributed sources. Each source corresponds to its own stationary process $\{W^{(m)}(t), t \geq 0\}$. Therefore, the cumulative packet count at time t is represented by $\sum_{m=1}^M W^{(m)}(t)$. We claim:

Theorem 3 For $H = (3 - \min(\alpha_1, \alpha_2))/2$,

$$\mathcal{L} \lim_{T \rightarrow \infty} \mathcal{L} \lim_{M \rightarrow \infty} \frac{1}{T^H} \frac{1}{\sqrt{M}} \int_0^{Tt} \sum_{m=1}^M (W^{(m)}(u) - EW^{(m)}(u)) du = \sigma B_H(t)$$

Here \mathcal{L} denotes the convergence in the sense of distribution. □

Proof of Theorem 3:

Actually, the load $\int_0^t W(u)du$ by time t has variance:

$$V(t) = \text{Var}\left(\int_0^t W(u)du\right) = 2 \int_0^t \left(\int_0^v \gamma(u)du\right)dv \quad (5)$$

where $\gamma(u) = EW(u)W(0) - (EW(0))^2$ denotes the auto-covariance function of W .

Using Laplace transform, we may derive that ([9]):

$$V(t) \sim \sigma^2 t^{2H} \text{ as } t \rightarrow \infty \quad (6)$$

Here $H = (3 - \min(\alpha_1, \alpha_2))/2$ which is rightly decided by $W^{(m)}(t)$, or more precisely, by the form of the tail distributions $F_{1c}(x)$ and $F_{2c}(x)$. Let us assume (6) holds. Then we first normalize the aggregated ⁵ $\sum_{m=1}^M W^{(m)}(t)$ over the source count:

$$\mathcal{L} \lim_{M \rightarrow \infty} M^{-1/2} \sum_{m=1}^M (W^{(m)}(t) - EW^{(m)}(t)) = G(t), \quad t \geq 0 \quad (7)$$

according to the usual Central Limit Theorem.

Furthermore, since $W^{(m)}(t)$ are stationary, $\{G(t), t \geq 0\}$ is Gaussian and stationary and has mean zero and auto-covariance function $\{\gamma(t), t \geq 0\}$. We also aggregate and normalize over the time. Then with each $T > 0$ for the process $(T^{2H})^{-1/2} \int_0^{Tt} G(u)du$, by (7) we have

$$\mathcal{L} \lim_{T \rightarrow \infty} (T^{2H})^{-1/2} \int_0^{Tt} G(u)du = \sigma B_H(t), \quad t \geq 0 \quad (8)$$

The limit in (8) is Gaussian with mean zero and has stationary increments since the integral of G has these properties. And by (7), its variance is $\sigma^2 t^{2H}$. According to Corollary 1 and 2 in the previous part, the three properties imply that $\sigma B_H(t)$ is a fractional Brownian Motion with stationary increments, mean zero and variance t^{2H} .

Last, the equation (8) leads to the explicit form:

$$\mathcal{L} \lim_{T \rightarrow \infty} \mathcal{L} \lim_{M \rightarrow \infty} \frac{1}{T^H} \frac{1}{\sqrt{M}} \int_0^{Tt} \sum_{m=1}^M (W^{(m)}(u) - EW^{(m)}(u))du = \sigma B_H(t) \quad (9)$$

□

More generally, rescaling time by a factor T , consider the aggregated cumulative packet counts over the interval $[0, Tt]$:

$$W_M^*(Tt) = \int_0^{Tt} (\sum_{m=1}^M W^{(m)}(u))du$$

we should have ([9]):

Theorem 4 When $1 < \alpha_j < 2$, set $a_j = L_j(\Gamma(2 - \alpha_j))(\alpha_j - 1)$. When $\sigma_j^2 < \infty$, set $\alpha_j = 2$, $L_j = 1$ and $a_j = \sigma_j^2/2$. Let

$$b = \lim_{t \rightarrow \infty} t^{\alpha_2 - \alpha_1} \frac{L_1(t)}{L_2(t)}$$

If $0 < b < \infty$, set $\alpha_{min} = \alpha_1 = \alpha_2$,

$$\sigma_{lim}^2 = \frac{2(\mu_2^2 a_1 b + \mu_1^2 a_2)}{(\mu_1 + \mu_2)^3 \Gamma(4 - \alpha_{min})}, \text{ and } L = L_2$$

⁵Intuitively, to aggregate means lumping all connections together into a single flow.

otherwise set

$$\sigma_{lim}^2 = \frac{2\mu_{max}^2 a_{min}}{(\mu_1 + \mu_2)^3 \Gamma(4 - \alpha_{min})}, \text{ and } L = \begin{cases} L_2 & b = \infty \\ L_2 & b = 0 \end{cases}$$

Also, let $H = (3 - \alpha_{min})/2$. Then for large M and T , the aggregate cumulative packet process $\{W_M^*(Tt), t \geq 0\}$ behaves statistically like

$$TM \frac{\mu_1}{\mu_1 + \mu_2} t + T^H \sqrt{L(T)M} \sigma_{lim} B_H(t) \quad (10)$$

Or more precisely,

$$\mathcal{L}lim_{T \rightarrow \infty} \mathcal{L}lim_{M \rightarrow \infty} \frac{W_M^*(Tt) - TM \frac{\mu_1}{\mu_1 + \mu_2} t}{T^H L^{1/2}(T) M^{1/2}} = \sigma_{lim} B_H(t) \quad (11)$$

where $\mathcal{L}lim$ has the same meaning as the previous theorem. \square

In sum, with an idealized on-off source model, we use the results above to confirm that packet flows (represented by normalized and aggregated source loads) become FBM at the application layer. Therefore, it exactly explains why there exhibits LRD phenomenon at that level, related to the evidence from recent studies in other areas. Besides, it also demonstrates to some extent how users applications affect the data network traffic.

4.3 Further Discussion

There are still several interesting and beneficial points left for further thinking. First, several experiments also show that some data networks possess LRD nature even if the source population is not huge enough, so then it is not fair any longer to make the source number M go to infinity in the model. There must be some other factors leading to the modification of the model. Another problem lies in the possible exchange of sum notation in equation (9). That is, we can also take the limit for M first, or even take the two limits at the same time. It should be interesting to find out the reason if those results do not exhibit FBM properties.

5 Acknowledgements

I greatly appreciate Prof. Flaschka for his continuous help and encouragement when I was exploiting this interesting topic and working on the paper. He made the whole process full of fun.

References

- [1] William G. Faris, **Lectures on Stochastic Processes**, November 8, (2001), <http://math.arizona.edu/~faris/lecturenotes.html>
- [2] I.M.Gelfand, **Generalized Functions**, Vol.4, Academic Pr; New Ed edition (July 1977)
- [3] Judith L. Jerkins, and Jonathan L. Wang, **From Network Measurement Collection to Traffic Performance Modeling: Challenges and Lessons Learned**, J. Braz. Comp. Soc. v.5 n.3 Campinas Fev. (1999)
- [4] Will E. Leland et al., **On the Self-Similar Nature of Ethernet Traffic (Extended Version)**, IEEE/ACM Transactions on Networking, Vol. 2, pp. 1-15, 1994.
- [5] P.Major, **Multiple Wiener-Ito Integrals**, Vol.849 of Springer Lecture Notes in Mathematics, Springer-Verlag, New York, (1981).
- [6] Benoit B. Mandelbrot, **Fractals: Form, Chance, and Dimension**, W. H. Freeman and Co., (1977)
- [7] Gennady Samorodnitsky, and Murad Taqqu, **Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance**, Chapman and Hall, New York, (1994)
- [8] Jaymie Strecker, **Fractional Brownian Motion Simulation: Observing Fractal Statistics in the Wild and Raising Them in Captivity**, (2004), <http://www.cs.umd.edu/~strecker/jstreckerIS-pdflatex.pdf>
- [9] M.S.Taqqu, W.Willinger, and R.Sherman, **Proof of a fundamental result in self-similar traffic modeling**, Computer Communications Review, 27(2)(1997), 5-23
- [10] Murad S. Taqqu, **Fractional Brownian Motion and Long-Range Dependence, Theory and Applications of Long-Range Dependence**, Birkhauser Press, Boston, (2003)
- [11] Walter Willinger, Vern Paxson, Rolf H. Riedi, Murad S. Taqqu, **Long-Range Dependence and Data Network Traffic, Theory and Applications of Long-Range Dependence**, Birkhauser Press, Boston, (2003)
- [12] Fractal, Wikipedia, <http://en.wikipedia.org/wiki/Fractal>
- [13] Internet protocol suite, Wikipedia, <http://en.wikipedia.org/wiki/TCP/IP>
- [14] Long-range Dependence, Wikipedia, http://en.wikipedia.org/wiki/Long-range_dependency
- [15] Packet flows, Wikipedia, [http://en.wikipedia.org/wiki/Flow_\(computer_networking\)](http://en.wikipedia.org/wiki/Flow_(computer_networking))
- [16] Self Similarity, Wikipedia, <http://en.wikipedia.org/wiki/Self-similarity>