Non Gaussian and Long Memory Statistical Modeling of Internet Traffic

Antoine Scherrer, Nicolas Larrieu, Pierre Borgnat, Philippe Owezarski,
Patrice Abry

HAL Id: ensl-00175418
https://ens-lyon.hal.science/ensl-00175418

Submitted on 28 Sep 2007

Non Gaussian and Long Memory Statistical Modeling of Internet Traffic.

A. Scherrer[1], N. Larrieu[2], P. Borgnat[3], P. Owezarski[2], P. Abry[3]

[1] LIP (UMR CNRS), ENS Lyon, France
[2] LAAS-CNRS, Toulouse, France
[3] Physics Lab. (UMR CNRS), ENS Lyon, France

Abstract. Due to the variety of services and applications available on today's Internet, many properties of the traffic stray from the classical characteristics (Gaussianity and short memory) of standard models. The goal of the present contribution is to propose a statistical model able to account both for the non Gaussian and long memory properties of aggregated count processes. First, we introduce the model and a procedure to estimate the corresponding parameters. Second, using a large set of data taken from public reference repositories (Bellcore, LBL, Auckland, UNC, CAIDA) and collected by ourselves, we show that this stochastic process is relevant for Internet traffic modeling for a wide range of aggregation levels. In conclusion we indicate how this modeling could be used in IDS design. **Key Words:** Statistical traffic modeling, Non Gaussianity, Long Memory, Aggregation Level.

1 Motivation: Internet traffic times series modeling

The Internet traffic grows in complexity as the Internet becomes the universal communication network and conveys all kinds of information: from binary data to real time video or interactive information. Evolving toward a multi-service network, its heterogeneity increases in terms of technologies, provisioning, and applications. The various networking functions (such as congestion control, traffic engineering, buffer management) required for providing differentiated and guaranteed services involve a variety of time scales which, added to the natural variability of the Internet traffic, generate traffic properties that deviates from those accounted for by simple models. Therefore, a versatile model for Internet traffic able to capture its characteristics regardless of time, place, and aggregation level, is a step forward for monitoring and improving the Internet: for traffic management, for charging, for injecting realistic traffic in simulators, for Intrusion Detection Systems,...
• **Traffic Modeling : A Brief Overview**. Packets arrival processes are natural fine-grain models of computer network traffic. They have long been recognised as not being Poisson (or renewal) processes, insofar as the inter-arrival delays are not independent [1]. Therefore, non-stationary Point processes [2] or stationary Markov Modulated Point processes [3, 4] have been proposed as models. However, the use of this fine granularity of modeling implies to take into account the

large number of packets involved in any computer network traffic, hence huge data sets to manipulate. So a coarser description of the traffic is more convenient: the packet or byte aggregated count processes. They are defined as the number of packets (resp., bytes) that lives within the k-th window of size $\Delta > 0$, i.e., whose time stamps lie between $k\Delta \leq t_l < (k+1)\Delta$; it will be noted $X_\Delta(k)$. In the present work, we concentrate on this aggregated packet level. More detailed reviews on traffic models can be found in, e.g., [5, 6]. Our objective is the modeling of $X_\Delta(k)$ with a stationary processes. Marginal distributions and auto-covariance functions are then the two major characteristics that affect network performance.

• **Marginal Distributions.** Due to the point process nature of the underlying traffic, Poisson or exponential distributions could be expected at small aggregation levels Δ; but this fails at larger aggregation levels. As $X_\Delta(k)$ is by definition a positive random variable, other works proposed to describe its marginal with common positive laws such as log-normal, Weibull or gamma distributions [5]. For highly aggregated data, Gaussian distributions are used in many cases as relevant approximations. However none of them can satisfactorily model traffic marginals both at small and large Δs. As it will be argued in the current paper, empirical studies suggest that a Gamma distribution $\Gamma_{\alpha,\beta}$ capture best the marginals of the X_Δ for a wide range of scales, providing a unified description over a wide range scales of aggregation Δ.

• **Covariance Functions and higher order statistics.** In Internet monitoring projects, traffic under normal conditions was observed to present large fluctuations in its throughput at all scales. This is often described in terms of long memory [7], self-similarity [6, 8], multifractality [9] that impacts the second (or higher order) statistics. For computer network traffic, long memory or long range dependence (LRD) property (cf. [10]) is an important feature as it seems related to decreases of the QoS as well as of the performance of the network (see e.g., [11]), hence need to be modeled precisely.

Let us recall that long range dependence is defined from the behaviour at the origin of the power spectral density $f_{X_\Delta}(\nu)$ of the process:

$$f_{X_\Delta}(\nu) \sim C|\nu|^{-\gamma}, \; |\nu| \to 0, \quad \text{with } 0 < \gamma < 1. \tag{1}$$

Note that Poisson or Markov processes, or their declinations, are not easily suited to incorporate long memory. They may manage to approximately model the LRD existing in a finite duration observation at the price of an increase of the number of adjustable parameters involved in the data description. But parsimony in describing data is indeed a much desired feature as it may be a necessary condition for a robust, meaningful and on the fly modeling. One can also incorporate long memory and short correlations directly into point processes using cluster point process models, yielding interesting description of the packet arrival processes as described in [12]. However this model does not seems adjustable enough to encompass a large range of aggregation window size.

An alternative relies on canonical long range dependent processes such as fractional Brownian motion, fractional Gaussian noise [13] or Fractionally Integrated Auto-Regressive Moving Average (farima) (see [6] and the references

therein). Due to the many different network mechanisms and various source characteristics, short term dependencies also exist (superimposed to LRD) and play a central role (cf. for instance [14]). This leads us to propose here the use of processes that have the same covariance as that of farima models [10], as they contain both short range and long range correlations.

• **Aggregation Level.** A recurrent issue in traffic modeling lies in the choice of the relevant aggregation level Δ. This is an involved question whose answer mixes up the characteristics of the data themselves, the goal of the modeling as well as technical issues such as real time, buffer size, computational cost constraints. Facing this difficulty of choosing a priori Δ, it is of great interest to have at disposal a statistical model that may be relevant for a large range of values of Δ, a feature we seek in the model proposed below.

• **Goals and outline of the paper.** This paper then focuses on the joint modeling of the marginal distribution and the covariance structure of Internet traffic time series, in order to capture both their non Gaussian and long memory features. For this purpose, we propose to use a non Gaussian long memory process whose marginal distribution is a gamma law, $\Gamma_{\alpha,\beta}$, and whose correlation structure is the one of a farima process, the farima(ϕ, d, θ) (Section 2). Through the application of this model to several network traffic traces of reference, we show how this 5 parameters model, $\alpha, \beta, d, \phi, \theta$ efficiently captures their statistical first and second orders characteristics. We also observe that this modeling is valid for a wide range of aggregation levels Δ (Section 3).

2 Non Gaussian Long Range Dependent stochastic Modeling: the Gamma farima model

2.1 Definitions and properties of the Gamma farima model

Given an aggregation level Δ, the notation $\{X_\Delta(k), k \in \mathbb{Z}\}$ denotes the aggregated count of packets, obtained for instance from the inter-arrival time series. The process X_Δ is assumed to be 2nd order stationary. To model it, we use the following non Gaussian, long range dependent, stationary model: the Gamma (marginal) farima (covariance) process.

• **First order Statistics (Marginals): Gamma distribution.** A $\Gamma_{\alpha,\beta}$ random variable (RV) is a positive random variable characterised by its shape $\alpha > 0$ and scale $\beta > 0$ parameters. Its probability density function (pdf) is defined as:

$$\Gamma_{\alpha,\beta}(x) = \frac{1}{\beta \Gamma(\alpha)} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \tag{2}$$

where $\Gamma(u)$ is the standard Gamma function (see e.g., [15]). Its mean and variance read respectively: $\mu = \alpha\beta$ and $\sigma^2 = \alpha\beta^2$. Hence, for a fixed α, varying β amounts to vary proportionally μ and σ. For a fixed β, the pdf of a $\Gamma_{\alpha,\beta}$ random variable evolves from a highly skewed shape ($\alpha \to 0$) via a pure exponential ($\alpha = 1$) towards a Gaussian law ($\alpha \to +\infty$). Therefore, $1/\alpha$, can be read as an index of the distance between $\Gamma_{\alpha,\beta}$ and Gaussian laws. For instance, skewness

and kurtosis read respectively as $2/\sqrt{\alpha}$ and $3 + 6/\alpha$. The proposed model assumes that first order statistics of the data follow a Gamma distribution with parameters that are indexed by Δ.

• **Second order Statistics (covariance): Long Range Dependence.** For sake of simplicity, in the present work, we restrict the general farima(P, d, Q) to the case where the two polynomials P and Q that describe the short-range correlations properties of the time-series (that could be of any arbitrary orders) are of order 1: the farima(θ, d, ϕ). The power spectral density of a farima(θ, d, ϕ) reads, for $-1/2 < \nu < 1/2$:

$$f_X(\nu) = \sigma_\epsilon^2 |1 - e^{-i2\pi\nu}|^{-2d} \frac{|1 - \theta e^{-i2\pi\nu}|^2}{|1 - \phi e^{-i2\pi\nu}|^2}, \tag{3}$$

where σ_ϵ^2 stands for a multiplicative constant and d for a fractional integration parameter $(-1/2 < d < 1/2)$. The constant σ_ϵ^2 is chosen such that the integral of the spectrum is equal to the variance given by the marginal law, $\sigma_X^2 = \alpha\beta^2$.

The parameter d characterises the *strength* of LRD: a straightforward expansion of Eq. 3 above at frequencies close to 0 shows that the farima(θ, d, ϕ) spectrum displays LRD as soon as $d \in (0, 1/2)$, with long memory parameters (see Eq. 1) $\gamma = 2d$ and $C = \sigma_\epsilon^2$. Parameters θ and ϕ enable to design the spectrum at high frequencies and hence to model short range correlations.

• **Gamma farima model.** The $\Gamma_{\alpha,\beta}$ - farima(ϕ, d, θ) model involves 5 parameters, adjusted from the data at each aggregation level Δ. As such, it is a parsimonious model that is nevertheless versatile enough to be confronted to data. As the process defined above is not Gaussian, the specifications of its first and second order statistical properties do not fully characterise it. Room for further design to adjust other properties of the traffic remains possible in the framework of the model. However, this is beyond the scope of the present contribution.

Sample paths of this $\Gamma_{\alpha,\beta}$ - farima(ϕ, d, θ) model can be numerically generated using a circular embedded matrix based synthesis procedure. The circular embedded matrix method [16] enables to synthesise a Gaussian process with a prescribed covariance. The adaptation of the method to the non-Gaussian model relies on two ideas: first we exhibit a transformation to generates a Gamma-distributed process $X(k)$ from a Gaussian process $Y(k)$; second we deduce for this transformation the link between the covariance X and the covariance of Y so that Y is generated with the needed covariance to obtain a farima covariance for X. Specifically, random Gaussian processes $Y(k)$ with variance $\beta/2$ and covariance $\gamma_Y = \sqrt{\gamma_X/4\alpha}$ are generated. X is then obtained as the sum of the squared Y's. Choosing γ_X as the Fourier transform of (3), we obtain then a realisation of the Gamma-farima model. The method is fully described in [17].

2.2 Analysis and Estimation

• **Gamma parameter estimation.** A classical maximum likelihood method is used for the estimation of α and β [18]. Using the method mentioned above for synthesising realisations of the model, numerical simulations have been done to

Data	Date(Start Time)	T (s)	Network(Link)	IAT (ms)	Repository
PAUG	1989-08-29(11:25)	2620	LAN(100BaseT)	2.6	ita.ee.lbl.gov
LBL-TCP-3	1994-01-20(14:10)	7200	WAN(100BaseT)	4	ita.ee.lbl.gov
AUCK-IV	2001-04-02(13:00)	10800	WAN(OC3)	1.2	wand.cs.waikato.ac.nz
CAIDA	2002-08-14(10:00)	600	Backbone(OC48)	0.01	www.caida.org
UNC	2003-04-06(16:00)	3600	WAN(100BaseT)	0.8	www-dirt.cs.unc.edu
Metrosec-fc1	2005-04-14(14:15)	900	LAN(100BaseT)	0.8	www.laas.fr/METROSEC

Table 1: **Data Description.** Description of the collection of traces. T is the duration is the trace, in second. IAT is the mean inter-arrival time, in ms.

check that this estimation procedure provides us with accurate estimates even when applied to processes with long range dependence [17]. Moreover, it behaves well, compared to the usual moment based technique, with $\hat{\beta} = \hat{\sigma^2}/\hat{\mu}$, $\hat{\alpha} = \hat{\mu}/\hat{\beta}$ where $\hat{\mu}$ and $\hat{\sigma^2}$ consist of the standard empirical mean and variance, whose quality are affected by the LRD.

• **Farima parameter estimation.** Estimation of the 3 parameters (θ, d, ϕ) boils down to a joint measurement of long memory and short-time correlation. A considerable amount of works and attention (see e.g. [19] for a thorough and up-to-date review) has been devoted to the estimation of long memory, in itself a difficult task. Joint estimations of long and short range parameters are possible, for instance from a full maximum likelihood estimation based on the analytical form of the spectrum recalled in Eq. 3, but they are computationally heavy.

We prefer to rely on a more practical two step estimation procedure: using a standard wavelet-based methodology (not recalled here, the reader is referred to [20]) d is first estimated. Then the estimation of the residual ARMA parameters θ and ϕ is performed using least squares techniques.
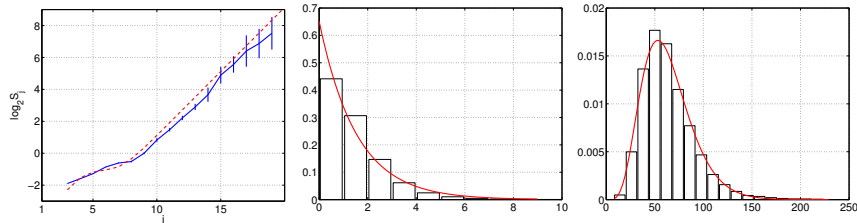

Fig. 1: $\Gamma_{\alpha,\beta}$ - farima(ϕ, d, θ) model for LBL-TCP-3: fit of covariances, fits of marginals for $\Delta = 4$ and 256 ms (from left to right). $j = 1$ corresponds to 1 ms.

3   Data analysis: Results and Discussions

3.1   The Collection of Data and Analysis

The model and analysis proposed in the present contribution are illustrated at work on a collection of standard Internet traces, described in Table 1, gathered from most of the major available Internet traces repositories (Bellcore, LBL, UNC, Auckland Univ, Univ North Carolina, CAIDA) as well as on time series collected by ourselves within the METROSEC research project. This covers a
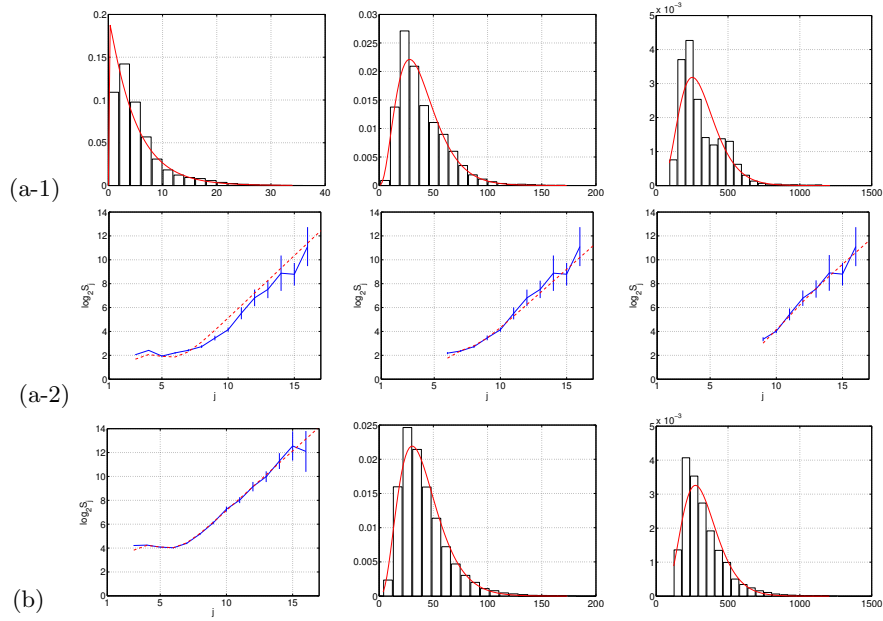
Fig. 2: $\Gamma_{\alpha,\beta}$ - farima(ϕ, d, θ) modeling of Metrosec-fc1. (a) Experimental data: fits for the marginals (top row) and covariances (bottom) for $\Delta = 4$, 32 and 256 ms (left to right). (b) Synthetic model with parameters taken from the data at $\Delta = 4$: fit of covariances, fits of marginals for $\Delta = 32$ and 256 ms (left to right). For all diagrams, $j = 1$ corresponds to 1 ms.

significant variety of traffic, networks (Local Area Network, Wide Area Network,... and edge networks, core networks,...) and links, collected over the last 16 years (from 1989 to 2005).

The $\Gamma_{\alpha,\beta}$ - farima(ϕ, d, θ) analysis procedures described above are applied to traffic time series X_Δ for various levels of aggregation Δ, independently. Stationarity of X is assumed, for theoretical modeling. Hence, we first perform an empirical time consistency check on the estimation results obtained from adjacent non overlapping sub-blocks, on X_Δ, for each aggregation level. Then, we analyse only data sets for which stationarity is a reasonable hypothesis. This approach is very close in spirit to the ones developed in [21, 22]. Due to obvious space constraints, we choose to report only results for the (somewhat old however massively studied) LBL-TCP-3 time series (see Fig. 1) and for the (very recent and hence representative of today's Internet, collected by ourselves) Metrosec-fc1 times series (see Fig. 2). However, similar results were obtained for all the data listed in Table 1 and are available upon request. Preliminary results (including the recent and well-known Auck-IV) were also presented in [17, 23].

3.2 Marginals and Covariance fits from the data

Fig. 1 (centre and right) and Fig. 2 (first row) consist of the model fits superimposed to empirical probability functions. It illustrates clearly the relevance of the $\Gamma_{\alpha_\Delta, \beta_\Delta}$ fits of the marginal statistics of X_Δ. Also, it shows that this adequacy

holds over a wide range of aggregation levels Δ ranging from 1ms up to 10s. Let us again put the emphasis on the fact that this is valid for various traffic conditions. The relevance of these fits has been characterised by means of χ^2 and Kolmogorov-Smirnow goodness-of-fit tests (not reported here). For some of the analysed time series or some aggregation levels, exponential, log-normal or χ^2 laws may better adjust the data. However, Gamma distributions are never significantly outperformed and they are undoubtedly the distributions that remain the most significant when varying Δ from very fine to very large.

Fig. 2 (second row) compares the logscale diagrams (LD) estimated from the analysed time series to logscale diagrams derived numerically from the combination of Eq. 3 with the estimated values of $\hat{d}_W$, $\hat{\theta}$ and $\hat{\phi}$ [4]. Increasing the aggregation level Δ is equivalent to filtering out the fine scale details while leaving the coarse scales unchanged. It is thus expected that the logscale diagram obtained for a given Δ corresponds to a truncated version of the diagrams obtained for smaller Δs, as can be seen in Fig. 2. One also sees on this figure that the coarse-scale part of the LDs is dominated by a long memory (for scales $2^j = 2^{10} \simeq 1$s).
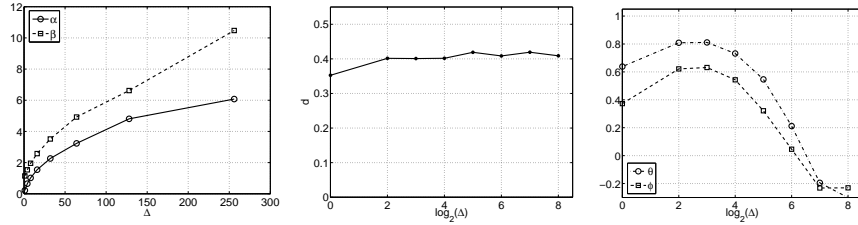

Fig. 3: LBL-TCP-3. Estimated parameters of $\Gamma_{\alpha,\beta}$ - farima(ϕ, d, θ). Left: $\hat{\alpha}_\Delta$ and $\hat{\beta}_\Delta$ (Δ in ms); middle and right: d then $\hat{\theta}$ and $\hat{\phi}$ as a function of $\log_2 \Delta$.

3.3 Adequacy of the Gamma farima model

The adequacy of the $\Gamma_{\alpha,\beta}$ class of distributions can be related to its stability under addition and multiplication properties. First, let X be a $\Gamma_{\alpha,\beta}$ RV, let $\lambda \in \mathbb{R}^{*+}$, then λX is a $\Gamma_{\alpha,\lambda\beta}$ RV. Traffic wise, this can be read as the fact that a global increase of traffic (its mean μ and standard deviation σ are both and jointly multiplied by a positive factor λ) is seen via the modification of the scale parameter β, while, conversely, it does not modify the shape parameter α. Therefore a multiplication of internet users at some time induces an increase of traffic volumes that can be very well accounted for within the $\Gamma_{\alpha,\beta}$ family and does not imply that its marginals are closer to Gaussian laws. Second, let X_i, $i = 1, 2$ be two independent $\Gamma_{\alpha_i,\beta}$ random variables, then their sum $X = X_1 + X_2$ follows a $\Gamma_{\alpha_1+\alpha_2,\beta}$ law. For traffic analysis, this is reminiscent of the aggregation procedure since one obviously has $X_{2\Delta}(k) = X_\Delta(2k) + X_\Delta(2k + 1)$. Therefore, assuming that X_Δ is a $\Gamma_{\alpha_\Delta,\beta_\Delta}$ process, independence among its

---

[4] Procedure developed with D. Veitch, cf. [24]

samples would imply that $\alpha_\Delta = \alpha_0\Delta$ and $\beta_\Delta = \beta_0$ (where $\alpha_0 = N/T$ and β_0 denote reference constants, with N the total number of packets (or bytes) seen within the observation duration T). Correlations in X_Δ will be underlined by departures of the functions $\alpha_\Delta = f(\Delta)$ and $\beta_\Delta = g(\Delta)$ from these behaviours. However, correlations do not change the fact that X_Δ is a $\Gamma_{\alpha_\Delta,\beta_\Delta}$ process for a wide range of aggregation levels ranging from very fine to very large. This covariance of the $\Gamma_{\alpha_\Delta,\beta_\Delta}$ model under a change of aggregation level Δ is therefore an argument very much in favour of its use to model Internet traffic.

Fig. 3 (left plot) shows the evolution of the estimated $\hat\alpha$ and $\hat\beta$ as a function of Δ for the LBL-TCP-3 trace (analogous features are observed for the Metrosec-fc1 time series). Practically, in our analysis $\hat\alpha$ and $\hat\beta$ are tied together via the relation $\hat\alpha\hat\beta = N\Delta/T$. Obviously, α does not follow a linear increase with Δ and $\hat\beta$ is not constant. Hence, significant departures from the *independence* behaviours are observed. The increase of $\hat\alpha$ unveils the fact that the probability density functions evolve towards Gaussian laws; we conclude that the model captures the convergence to Gaussianity. However, Fig. 3 shows that this evolution to Gaussianity goes far more slowly than it would in the case of a pure Poisson arrival process even when Δ is lower than 1s. Note that $\Delta \simeq 1$ s corresponds to the onset of long memory (as will be discussed below). The functions $\hat\alpha_\Delta$ and $\hat\beta_\Delta$ shown here accommodate then mainly for short range dependencies.

As drawn on Fig. 3 (centre), the wavelet-based estimated $\hat d$s characterising the strength of the long memory remain constant for all Δ, as expected from the theoretical analysis. This is an evidence that long range dependence is a long-time feature of the traffic that has no inner time-scale. Conversely, short-time correlations have naturally short characteristic time-scales so that the filtering due to an increase of Δ smooths them out. Indeed, in Fig. 3 (right plot), one sees that $\hat\phi$ and $\hat\theta$ decrease and converge one towards another for $\Delta \geq 200$ ms, so that their influence on the spectrum cancels out. The model is then dominated by the long memory, and close to a (non Gaussian) farima$(0, d, 0)$.

Another indication of the adequacy of the model is shown on Fig. 2, through the behaviour of a synthesised Gamma-farima model. We have generated, using the method described above in section 2.1, sample paths of traffic with the same parameters as the model for Metrosec-fc1 data at scale $\Delta = 4$ms. Here the analysis of the synthetic data at coarser time-scales Δ is reported on Fig. 2, row (b), through its logscale diagram (left), and its empirical marginals and fits at $\Delta = 32$ and 256ms (centre and right). They compare well to the same fits of the experimental data (same Fig. rows (a)).

4   Conclusion and Future Researches

In the present work, we have shown that the proposed $\Gamma_{\alpha,\beta}$ - farima(ϕ, d, θ) model reproduces the marginals and both the short range and long range correlations of Internet traffic time series. This statement holds for a large variety of different traffic collected on different links and networks, including traffic collected by ourselves. The proposed model is able to fit data aggregated over a wide range of

aggregation levels Δ, accounting for the properties through the evolutions of the five parameters with Δ and enabling a smooth and continuous transition from pure exponential to Gaussian laws. This provides us with a practical solution to address the question of modeling related to the choice of the relevant aggregation level Δ. As choosing Δ a priori may be uneasy, using a process that offers an evolutive modeling with Δ is of high interest. Another central point lies in the fact that the (absolute) values of the parameters of the models obviously are likely to vary from one time series to another. In our model, the main information is contained in the evolution of the parameters with Δ, not in their absolute values. Also, the proposed model is very parsimonious as it is fully described via 5 parameters. When necessary, the short range correlations can be further designed by straightforwardly extending $\Gamma_{\alpha,\beta}$ - farima(ϕ, d, θ) model to a $\Gamma_{\alpha,\beta}$ - farima(P, d, Q) one, where P and Q are polynomials of arbitrary orders.

This contribution opens the track for two major research directions, under current explorations. First, Because a numerical synthesis exists that enables to generate numerically random realizations of the model [17], it is possible to simulate realistic background traffic that can be used to feed network simulators and hence study network performance and quality of services alterations under various simulated circumstances. Second, the monitoring with respect to time of the estimated parameters of this model may constitute a central piece in traffic anomaly detection. A change in the evolutions with Δ of one (or more) paramater(s), rather than a change of the values themselves, can be detected and used as an alarm to indicate the occurrence of an anomaly. This signature can be further used to classify anomalies into legitimate ones (flashcrowd...), illegitimate but non aggressive ones (failures...) and illegitimate and aggressive ones (Attacks...). For this latter case, it can be used to start a defense procedure. This may serve as a starting point for an Intrusion Detection System design.

References

1. Paxon, V., Floyd, S.: Wide-area traffic: The failure of Poisson modeling. ACM/IEEE transactions on Networking **3**(3) (1995) 226–244
2. Karagiannis, T., Molle, M., Faloutsos, M., Broido, A.: A non stationary Poisson view of the internet traffic. In: INFOCOM. (2004)
3. Andersen, A., Nielsen, B.: A Markovian approach for modelling packet traffic with long range dependence. IEEE journal on Selected Areas in Communications **5**(16) (1998) 719–732
4. Paulo, S., Rui, V., António, P.: Multiscale fitting procedure using Markov Modulated Poisson Processes. Telecommunication Systems **23 (1/2)** (2003) 123–148
5. Melamed, B.: An overview of TES processes and modeling methodology. In: Performance/SIGMETRICS Tutorials. (1993) 359–393

6. Park, K., Willinger, W.: Self-similar network traffic: An overview. In Park, K., Willinger, W., eds.: Self-Similar Network Traffic and Performance Evaluation. Wiley (Interscience Division) (2000) 1–38

7. Erramilli, A., Narayan, O., Willinger, W.: Experimental queueing analysis with long-range dependent packet traffic. ACM/IEEE transactions on Networking $\mathbf{4}$(2) (1996) 209–223

8. Leland, W.E., Taqqu, M.S., Willinger, W., Wilson, D.V.: On the self-similar nature of ethernet traffic (extended version). ACM/IEEE transactions on Networking $\mathbf{2}$(1) (1994) 1–15

9. Feldmann, A., Gilbert, A., Willinger, W.: Data networks as cascades: Investigating the multifractal nature of internet WAN traffic. In: ACM/SIGCOMM conference on Applications, technologies, architectures, and protocols for computer communication. (1998)

10. Beran, J.: Statistics for Long-memory processes. Chapman & Hall, New York (1994)

11. Tsybakov, B., Georganas, N.: Self similar processes in communications networks. IEEE Trans. on Info. Theory $\mathbf{44}$(5) (1998) 1713–1725

12. Hohn, N., Veitch, D., Abry, P.: Cluster processes, a natural language for network traffic. IEEE Transactions on Signal Processing Special Issue on Signal Processing in Networking $\mathbf{8}$(51) (2003) 2229–2244

13. Norros, I.: On the use of fractional Brownian motion in the theory of connectionless networks. IEEE journal on Selected Areas in Communications $\mathbf{13}$(6) (1995)

14. Huang, C., Devetsikiotis, M., Lambadaris, I., Kaye, A.: Modeling and simulation of self-similar Variable Bit Rate compressed video: A unified approach. In: ACM SIGCOMM, Cambridge, UK (1995)

15. Evans, M., Hastings, N., Peacock, B.: Statistical Distributions. Wiley (Interscience Division) (2000)

16. Bardet, J., Lang, G., Oppenheim, G., Philippe, A., Taqqu, M.: Generators of long-range dependent processes: a survey. In Doukhan, P., Oppenheim, G., Taqqu, M., eds.: Long-Range Dependence: Theory and Applications, Boston, Birkhäuser (2003) 579–623

17. Scherrer, A., Abry, P.: Marginales non gaussiennes et longue mémoire : analyse et synthèse de trafic Internet. In: Colloque GRETSI-2005, Louvain-la-Neuve, Belgique (2005)

18. Hahn, G., Shapiro, S. In: Statistical Models in Engineering. Wiley (Interscience Division) (1994) 88

19. Doukhan, P., Oppenheim, G., Taqqu, M.: Long-Range Dependence: Theory and Applications. Birkhäuser, Boston (2003)

20. Abry, P., Veitch, D.: Wavelet analysis of long-range dependent traffic. IEEE Trans. on Info. Theory $\mathbf{44}$(1) (1998) 2–15

21. Uhlig, S., Bonaventure, O., Rapier, C.: 3D-LD: a graphical wavelet-based method for analyzing scaling processes. In: ITC Specialist Seminar, Würzburg, Germany (2003) 329–336

22. Veitch, D., Abry, P.: A statistical test for the time constancy of scaling exponents. IEEE Transactions on Signal Processing $\mathbf{49}$(10) (2001) 2325–2334

23. Borgnat, P., Larrieu, N., Abry, P., Owezarksi, P.: Détection d'attaques de "déni de services" : ruptures dans les statistiques du trafic. In: Colloque GRETSI-2005, Louvain-la-Neuve, Belgique (2005)

24. Veitch, D., Abry, P., Taqqu, M.S.: On the automatic selection of the onset of scaling. Fractals $\mathbf{11}$(4) (2003) 377–390