Computación y Sistemas Vol. 9 Núm. 4, pp. 340-354
© 2006, CIC-IPN, ISSN 1405-5546, Impreso en México

Shot Noise Modeling of Heavy Tailed Activity Periods

Modelo de Ruido de Disparo para Periodos de Actividad con Distribución de Colas Pesadas

David Muñoz Rodríguez[1] , Francisco Davila[1], Marlene Angulo Bernal[2,3], Deni L. Torres Román[3], and John Fonseka[4]

[1] ITESM; Center for Electronics and Telecommunications; Monterrey, N.L, 64849, MEXICO.
dmunoz@itesm.mx
[2]UABC, Campus Mexicali, México
mangulo@uabc.mx
[3]CINVESTAV, Campus Guadalajara, México
{mangulo, dtorres}@gdl.cinvestav.mx
[4]University of Texas at Dallas, Engineering and Computer Sc., Richardson, USA
kjp@utdallas.edu

Article received on January 27, 2005; accepted on January 13, 2006

Abstract

For web users, transmission activity is often described in terms of the number of file requests and file sizes that have a slowly decaying survivability function. In this article we model this line activity using a flexible shot-noise representation that enables us to analyze the infrared catastrophe effect introduced by the heavy-tailed characteristics. The proposed methodology is based on the α-stable characteristics of the traffic pattern. In addition, the effective bandwidth estimation of the proposed traffic model is presented. It provides a useful approach for the data network design and the admission control mechanism. Numerical results demonstrate the impact of the different statistical parameters on the spectral performance.
Keyword: Stochastic processes, Shot noise, Traffic models.

Resumen

La actividad en la transmisión de los usuarios Web es frecuentemente descrita en términos del número de solicitudes de archivos y de sus tamaños, donde la función de sobrevivencia del tamaño de los archivos tiene un decaimiento lento. En este artículo se modela la actividad de línea utilizando una representación flexible del ruido de disparo que permite analizar el efecto catástrofe infrarroja introducido por las características de las distribuciones de cola pesada. La metodología propuesta es basada en las características α-estables del patrón de tráfico, presentando además la estimación del ancho de banda efectivo del modelo propuesto, lo cual brinda un enfoque muy útil para el diseño de redes de datos y los mecanismos de control de admisión. Los resultados numéricos demuestran el impacto de los diversos parámetros estadísticos en el desempeño espectral.
Palabras clave: Procesos estocásticos, ruido de disparo, modelos de tráfico.

1 Introduction

Although the existence of heavy-tailed phenomena in telecommunications has been recognized since the 1960s (Berger et al., 1963; Ritcher, 1960; Susman, 1963), it is only after the massive proliferation of internet services which cannot be modeled by conventional tools, that heavy-tailed telecommunications phenomena are receiving attention. Spectral characterization is important as the low frequency content relates to the long-range dependence of the process. The traffic modeling approaches are diverse. For instance, while some authors (Ryu and Lowen, 1996) propose the use of the fractal-shot-noise driven Poisson process to model self-similar traffic with long-range dependence (LRD), others (Crovela and Bestavros, 1997) propose a WWW traffic characterization based on *a general* heavy-tail model to describe the on and off periods.

In this paper, we propose shot noise as a flexible way to model a variety of heavy-tailed phenomena. The proposed approach is based on the alpha-stable characteristics of the non-Gaussian Paretian variables. Our approach is based on the fact that WWW traffic also presents the Noah effect (representing the variability of stable random variables): and we take advantage of the well established usefulness of the spectral analysis (Lowen and Teich, 1991).

It is important to remark that Crovela characterizes WWW traffic based on a general heavy-tail distribution to describe the on and off periods, while in this paper we deal with alpha-stable totally skewed distribution to describe the on periods, with the consequent computational advantage of simple effective bandwidth estimation.

On the other hand Ryu and Lowen paper is related to Point Processes Models (PPM). In this paper, the line activity spans are heavy tailed with starting points modeled according to a flexible PPM shot noise representation.

Although these three approaches deal with infrared noise, they tackle different problems as: Ryu and Lowen model relies on second order statistics such as covariance function, count dispersion index, Power Spectrum Density (PSD). These statistics are function of the mean arrival rate, Hurst, and fractal onset time. Crovela´s work considers asymmetrical ON/OFF period distributions in order to explain the Web traffic self-similarity. In our work the proposed methodology is based on skewed (β=1) α-stable characterization of the traffic pattern as: α, γ and μ and the occurrence rate λ of activity periods, presenting two analysis techniques: Spectral characterization, and the Effective Bandwidth. While the spectral characterization shows the impact of the alpha stability parameters on the 1/f noise, the effective bandwidth analysis allows determining the number of connections that can be multiplexed without violating any service levels guaranteed to the customers (Kelly, 1996). This approach is useful to compute the amount of resources to fulfill the user application requirements a necessary element on design of data networks as: Multiprotocol Label Switching (MPLS) and Integrated Services architectures (IntServ), and admission control mechanisms on current networks.

The remainder of this paper is organized as follows. Section two introduces concepts related to shot noise modeling, statistical properties of network traffic as well as effective bandwidth definition. Section three presents the spectral analysis of the proposed traffic model. While the effective bandwidth estimation of the proposed traffic model is presented in section four. Finally, the concluding remarks are presented.

2 Conceptual Framework

A shot-noise process *s(t)* can be described, extending the Gilbert and Pollack notation (Gilbert and Pollack, 1960), as the aggregation of multiple stochastic signals involving *k* random parameters $\bar{i} = (i_1, i_2, \ldots i_k)$. That is,

$$s(\mathrm{t}) = \sum_{\hat{i}} h_{\bar{i}}(t) = \sum_{\hat{i}} h_{i_1, i_2, \ldots, i_k}(t) \cdot \tag{1}$$

For instance, this general description allows representation of a line-data signal such as $s(\mathrm{t}) = \sum_{\hat{i}} h_{\bar{i}}(t)$ where $h_i(t)=a_i\, p(t-iT-\tau_i)$, $\{a_i\}$ is the random data sequence, *p(t)* is the transmitted wave shape, while the instants *iT* are associated with the occurrence times (often assumed to be regularly spaced), and the τ_is account for random timing errors or jitter. A similar model can also be used to represent traffic volume expressed in terms of packets per time unit. In this case, $\{a_i\}$ represents the number of arrived packets during the i-th observation span, where each observation has a duration of *T*; τ_i can be neglected, and $p(t)=1$ for $t \in [0,T]$ and $p(t)=0$ otherwise.

In other phenomena, τ_i instants are associated with a Poisson process while $h_i(t) = (t - \tau_i)^{-\alpha}$ for $t > \tau_i$ and 0 otherwise, with α being a non-negative constant. This latter case is known in the literature as fractal shot noise (Lowen and Teich, 1991).

It is known that line activity can be viewed as a sequence of alternating active and quiescent runs (Adler et al., 1998). In a WWW environment, activity periods are known to exhibit a long-range dependence, often described in terms of Noah/Joseph cycles (Mandelbrot, 1999a). This activity pattern can be described as a "sporadic function"[1]

[1] The "sporadic function" term holds for the class of functions not necessarily stationary, where the Wiener-Khinchin theorem can be applied (Mandelbrot, 1999b).

$s(t)=\sum h(\frac{t}{\xi_i}-\tau_i)$ where ξ_i and τ_i represent the duration and occurrence of the activity span, respectively, and *h(t)* is a shifted rectangular function (i.e., *h(t)=1* for $t\in[0,T]$ and *h(t)=0* otherwise).

Heavy Tailed Distributions

It is accepted that current Internet traffic exhibits heavy-tail characteristics.
Definition: A random variable X has a heavy-tailed distribution if

$$P[X>x]\sim x^{-\alpha} \quad \text{as } x\to\infty,\ 0<\alpha<2. \tag{2}$$

That is, regardless of the behavior of the distribution for small values of the random variable, if the asymptotic shape of the distribution is hyperbolic, it is heavy-tailed. The simplest heavy-tailed distribution is the Pareto distribution. The Pareto distribution is hyperbolic over its entire range. Heavy-tailed distributions have the following properties:

1) If $\alpha\le 2$, then the distribution has infinite variance.
2) If $\alpha\le 1$, then the distribution has infinite mean.

Thus, as alpha decreases, an arbitrarily large portion of probability mass may be present in the tail of the distribution.

α-Stable Distributions

Definition: A random variable X is said to have a stable distribution if for X_1, X_2 independent copies of X and any positive numbers A and B, there is a positive number C and a real number D such that

$$AX_1+BX_2\overset{d}{=}CX+D. \tag{3}$$

Theorem 1: For any stable random variable X, there is a number $\alpha\in(0,2]$ such that the number C in (3) satisfies:

$$C^\alpha = A^\alpha + B^\alpha. \tag{4}$$

A stable random variable X with index α is called *Alfa stable*.

The generalized theorem of Central Limit proposes the alpha-stable distributions to model the aggregated contribution of many random variables, without restricting them to have a finite variance (as central limit theorem does). The α-stable characteristic function *C(ω)* is expressed as

$$\log C_x(\omega)=\varphi_{\alpha,\beta}^{\gamma,\mu}(\omega)= j\mu\omega-\gamma|\omega|^\alpha[1+j\beta\,\mathrm{sgn}(\omega)\psi(\alpha,\omega)],$$
$$where \quad \psi(\alpha,\omega)=\begin{cases}\tan(\frac{\alpha\pi}{2}) & for \quad \alpha\neq 1\\ \frac{2}{\pi}\log|\omega| & for \quad \alpha=1\end{cases}. \tag{5}$$

Indexes α, β, γ and μ are respectively known as stability index, skewness, dispersion and location parameters, and *sgn (ω) = -1, 0* and *1* for *ω < 0, ω = 0* and *ω >0* , respectively.

α-stable variables have a characteristic function of the form $C_{\alpha,\beta}^{\gamma,\mu}(\omega)=\exp\{\varphi_{\alpha,\beta}^{\gamma,\mu}(\omega)\}$ given by (5), and they are denoted as $x\overset{d}{-}S_\alpha(\gamma,\beta,\mu)$. Special cases of alpha stability are the Gaussian $S_2(\gamma,\beta,\mu)$, Cauchy $S_1(\gamma,0,\mu)$ and Levy $S_{1/2}(\gamma,1,\mu)$

distributions (Leon-Garcia, 1994). As these special distributions have closed-form expressions, they are usually studied with other methodology, and they are not discussed in this article.

It is known (Leon-Garcia, 1994) that for a non-zero constant κ and for a set of independent α-stable variables x_i with the same stability index (i.e. $x_i \overset{d}{-} S_\alpha(\gamma_i, \beta_i, \mu_i)$), $x_i + x_j$ and kx_i are also α-stable distributions, i.e.

$$x_i + x_j \overset{d}{-} S_\alpha(\gamma, \beta, \mu), \tag{6-a}$$

where $\gamma = [\gamma_1^\alpha + \gamma_2^\alpha]^{1/\alpha}$; $\beta = \dfrac{\beta_1\gamma_1^\alpha + \beta_2\gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}$; $\mu = \mu_1 + \mu_2$, and

$$\kappa x_i \overset{d}{-} S_\alpha(|\kappa|\gamma_i, \operatorname{sgn}(\kappa)\beta_i, \kappa\mu_i) \qquad if \quad \alpha \neq 1 \cdot \tag{6-b}$$

Long Range Dependence (LRD)

The self-similar property of Internet traffic manifests itself in the autocorrelation function $\rho(k)$. The sum of all autocorrelations from any given time instant is always significant, even if individual correlations are small.

Internet traffic exhibits an LRD phenomenon. This means that Internet traffic characteristics, at time *t*, will have a long term influence. And LRD is defined as the property of some processes in which the sum of the autocorrelation values approaches infinity. This is:

$$\sum_{k=0}^{\infty} \rho(k) = \infty \cdot \tag{7}$$

Where its autocorrelation function is

$$\rho(k) \sim L_1(k)k^{-\beta} \qquad k \to 0 \qquad 0 < \beta < 1 \ , \tag{8}$$

k is the lag and L_1 is a slow-varying function[2].

Or, equivalently, as the power-law divergence at the origin of its power spectrum:

$$f(\lambda) \sim L_2(\lambda)\lambda^{-D} \qquad \lambda \to 0 \qquad 0 < D < 1 \ , \tag{9}$$

λ is the frequency and L_2 is a slow-varying function.

Effective Bandwidth

Effective bandwidth is a concept in terms of throughput. It provides a measure of resource usage which adequately represents the trade-off between sources of different types, considering their statistical characteristics and Quality of Service (QoS) requirements.

The definition of the effective bandwidth $\alpha(\theta, t)$ associated with a source is (Kelly , 1996):

$$\alpha(\theta,t) = \frac{1}{\theta t}\log E[e^{\theta X[0,t]}] \qquad 0 < \theta,\ t < \infty, \tag{10}$$

[2] A slow positive and continues function (L) is slow varying if satisfies: $\lim_{x\to\infty} L(tx)/L(x) = 1 \ \ \forall t > 0$

where θ and t are the space and time scale parameters, X[0,t] is the traffic load that arrives from a source in the interval [0,t]. It is assumed that X[0,t] has stationary increments that represents the source's traffic load present on equation (10) in the interval [0,t].

Time scale parameter is the minimum scale on which the traffic needs to be observed. Typical time scale parameter goes from 10msec to 100 msec.

Space scale parameter is directly related with the QoS that a system (such a router) provides to a source. For a high QoS the reservation of system resources should tend to the source peak rate, this is a high θ-value. Otherwise, the resource reservation must be at least the mean rate, leading on a lower QoS and θ-value.

It is important to remark that equation (10) is for general application, can be used to compute the effective bandwidth of any source X[0,t]. In section four we will focus on the effective bandwidth of the Shot Noise model proposed.

3 Spectral Analysis

In the case of web page retrieval scenarios, it has been observed (Fisher and Harris, 1999; Resnick 1999) that τ_i correspond to customers' requests following a Poisson pattern while web page sizes or activity periods ξ_i can be modeled by a heavy-tailed distribution (Adler et al., 1998; Park and Willinger, 2000). Note that ξ_i values will depend not only on the file sizes but also on the network's engineering parameters.

Heavy-tailed phenomena exhibit long-range dependence, which is characterized in terms of a slowly decaying correlation function. Mandelbrot (1999a) has proposed describing these phenomena as so-called $1/f$ -noise, relating the dependence range to the low frequency content.

Autocorrelation $R_s(\tau)$ of *s(t)* can be expressed as

$$R_s(\tau) = E\{\sum_i \sum_k h_{\tau_i,\xi_i}(t) h_{\tau_k,\xi_k}(t+\tau)\} = \sum_i \sum_k \underset{(\tau_i \ \xi_i \ \tau_k \ \xi_k)}{E} \{\int h_i(t) h_k(t+\tau) dt\}, \tag{11}$$

and hence its power-spectral density function becomes $S_s(\omega) = \sum_{\bar{\tau}} \sum_{\bar{\xi}} \underset{(\bar{\tau} \ \bar{\xi})}{E} \{H_i(\omega) H_k{}^*(\omega)\}$.

Since *h(t)* has been assumed to be a shifted rectangular function

$$S_s(\omega) = \sum_i \sum_k \underset{(\tau_i \ \xi_i \ \tau_k \ \xi_k)}{E} \left\{ 2\frac{\sin\frac{\omega\xi_i}{2}}{\omega} \exp\{-j\omega(\frac{\xi_i}{2}+\tau_i)\} \quad 2\frac{\sin\frac{\omega\xi_k}{2}}{\omega} \exp\{j\omega(\frac{\xi_k}{2}+\tau_k)\} \right\} \tag{12}$$

Assuming independence among τ_is and ξ_js and using the Euler formulae (Leon-Garcia A., 1994), equation (12) can be written as:

$$S_s(\omega) = \frac{1}{\omega^2} \sum_i \sum_k C_{\tau_i-\tau_k}(\omega)[1 - C_{\xi_i}(\omega) - C^*_{\xi_k}(\omega) + C_{\xi_i-\xi_k}(\omega)] \tag{13}$$

where $C_x(\omega)$ is the characteristic function of the random variable *x*.

It has been reported in the literature that τ_is can be modeled by a Poisson processes (Fisher and Harris, 1999; Liu et al., 2000). While the activity spans ξ_is can be modeled by Paretian distributions (Mandelbrot, 1999a). This corresponds to the flexible model referred to in the literature as M/G/∞ (Parulekar and Makowsky, 1997; Kunz and Makowsky, 1998). It is known that Pareto laws belong to the so-called stable Paretian or stable non-Gaussian distributions

(Samarodnitsky and Taqqu, 1994).Therefore, since duration spans ξ_i and ξ_k are both α-stable, the difference ξ_i-ξ_k is also α-stable $\xi_i - \xi_k \overset{d}{-} S_\alpha(2^{1/\alpha}\gamma,0,0)$ and $C_{\xi_i-\xi_k}(\omega) = C_{\alpha,0}^{2^{1/\alpha}\gamma,0}(\omega) = \exp\{-2^{1/\alpha}\gamma|\omega|^\alpha\}$.

Recalling that ξ is a non-negative random variable[3], i.e. $\beta = 1$, the spectral representation becomes

$$S_s(\omega) = \frac{1}{\omega^2}\sum_i\sum_k\{1 - 2e^{-\gamma|\omega|^\alpha}\cos(\mu\omega - \gamma|\omega|^\alpha \operatorname{sgn}(\omega)\psi(\alpha,\omega)) + e^{-2^{1/\alpha}\gamma|\omega|^\alpha}\}C_{\tau_i-\tau_k}(\omega)\cdot \tag{14}$$

Considering a normalized time interval during which n number of packets arrive, (14) can be written as

$$S_s(\omega) = E_n\left\{\frac{1}{\omega^2}\sum_{i=1}^{n}\sum_{k=1}^{n}\{1 - 2e^{-\gamma|\omega|^\alpha}\cos(\mu\omega - \gamma|\omega|^\alpha \operatorname{sgn}(\omega)\psi(\alpha,\omega)) + e^{-2^{1/\alpha}\gamma|\omega|^\alpha}\}C_{\tau_i-\tau_k}(\omega)\right\},$$

$$S_s(\omega) = \frac{1}{\omega^2}\{1 - 2e^{-\gamma|\omega|^\alpha}\cos(\mu\omega - \gamma|\omega|^\alpha \operatorname{sgn}(\omega)\psi(\alpha,\omega)) + e^{-2^{1/\alpha}\gamma|\omega|^\alpha}\}E_n\left\{\sum_{i=1}^{n}\sum_{k=1}^{n}C_{\tau_i-\tau_k}(\omega)\right\}\cdot \tag{15}$$

Defining $\chi(n) = \sum_{i=1}^{n}\sum_{k=1}^{n}C_{\tau_i-\tau_k}(\omega)$ and $\Delta_l = \tau_i - \tau_{i+l}$ and writing $\chi(n)$ as

$$\chi(n) = nC_0(\omega) + 2l\sum_{l=1}^{n-1}(n-l)C_{\Delta_l}(\omega), \tag{16}$$

$S_s(\omega)$ becomes

$$S_S(\omega) = \frac{1}{\omega^2}\{1 - 2e^{-\gamma|\omega|^\alpha}\cos(\mu\omega - \gamma|\omega|^\alpha \operatorname{sgn}(\omega)\psi(\alpha,\omega)) + e^{-2^{1/\alpha}\gamma|\omega|^\alpha}\}E_n\{\chi(n)\}\cdot \tag{17}$$

Recalling that the duration between l Poisson arrivals (Δ_l) is Gamma distributed and observing that it can take both positive and negative values, its PDF can be written as (Gradshteyn and Ryzhik, 1994).

$$f_{\Delta_l}(\Delta_l) = \frac{1}{2}\lambda\frac{e^{-\lambda|\Delta_l|}(\lambda\Delta_l)^{l-1}}{(l-1)!}, \tag{18}$$

where λ is the rate of the point Poisson process. Hence, its characteristic function $C_{\Delta_l}(\omega)$ can be written as

$$C_{\Delta_l}(\omega) = \frac{1}{2}\left[\left(\frac{\lambda}{\lambda - j\omega}\right)^l + \left(\frac{\lambda}{\lambda + j\omega}\right)^l\right], \tag{19}$$

$$C_{\Delta_l}(\omega) = \rho^l\cos(l\phi), \tag{20}$$

where $\phi = tan^{-1}(\omega/\lambda)$ and $\rho = \frac{\lambda}{\sqrt{\lambda^2 + \omega^2}}$. Substituting (19) in (16) and noticing that $C_0(\omega)$=1,

[3] For a given distribution F(x) skewness β is defined as $\lim_{x \to \infty}\frac{1 - F(x) - F(-x)}{1 - F(x) + F(-x)}$.

$$\chi(n) = n + 2\sum_{l=1}^{n-1}(n-l)\rho^{l}\cos(l\phi)\cdot \tag{21}$$

Since n is Poisson distributed, the expectation of $\chi(n)$ can be expressed as

$$E_n\{\chi(n)\} = \sum_{n=0}^{\infty}\left(n + 2\sum_{l=1}^{n-1}(n-l)\rho^{l}\cos(l\phi)\right)\frac{\lambda^n}{n!}e^{-\lambda}\cdot \tag{22}$$

This expression can be reduced to equation (23) (Gradshteyn and Ryzhik, 1994)

$$E\ \{\chi(n)\} = \lambda + \frac{2e^{-\lambda}}{1-2\rho\cos\phi+\rho^2}\sum_{n=0}^{\infty}\frac{(\rho\lambda)^n}{n!}(\rho\cos[(n-1)\phi] - \cos[n\phi]) - 2\frac{\partial}{\partial\phi}\left\{\frac{1}{1-2\rho\cos\phi+\rho}\left(\ \rho\sin[\phi] + e^{-\lambda}\sum_{n=0}^{\infty}\frac{(\rho\lambda)^n}{n!}(\rho\sin[(n-1)\phi]-\sin[n\phi])\ \right)\right\}\cdot \tag{23}$$

Using the following relationships

$$\sum_{n=0}^{\infty}\frac{(\lambda\rho)^n\cos(n\phi)}{n!} = e^{\lambda\rho\cos\phi}\cos(\lambda\rho\sin\phi), \tag{24-a}$$

and

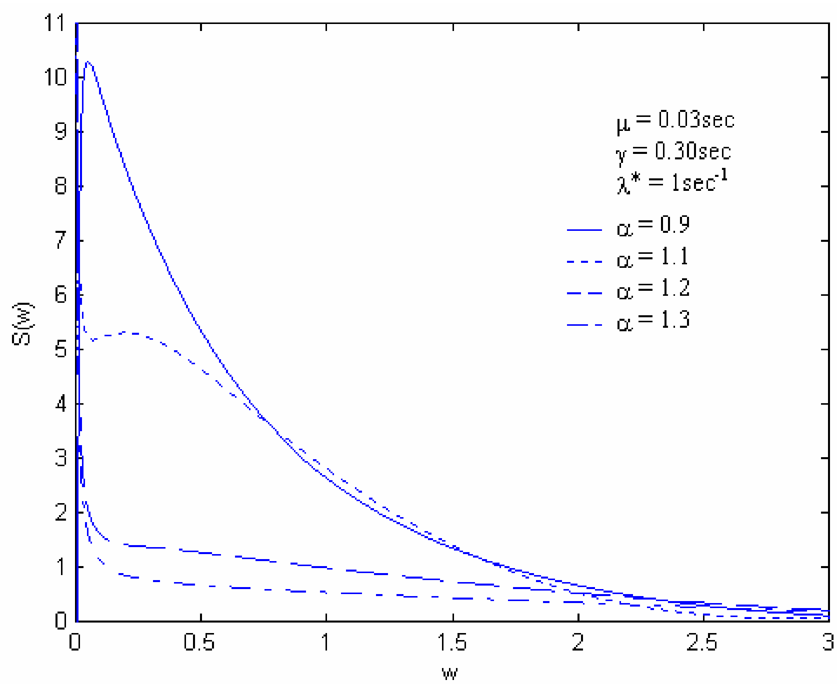
$$\sum_{n=0}^{\infty}\frac{(\lambda\rho)^n\sin(n\phi)}{n!} = e^{\lambda\rho\cos\phi}\sin(\lambda\rho\sin\phi), \tag{24-b}$$

which hold for $|\lambda\rho| \le 1$ (Mandelbrot, 1977)**,** and after some algebraic reductions $E_n\{\chi(n)\}$ can be simplified as

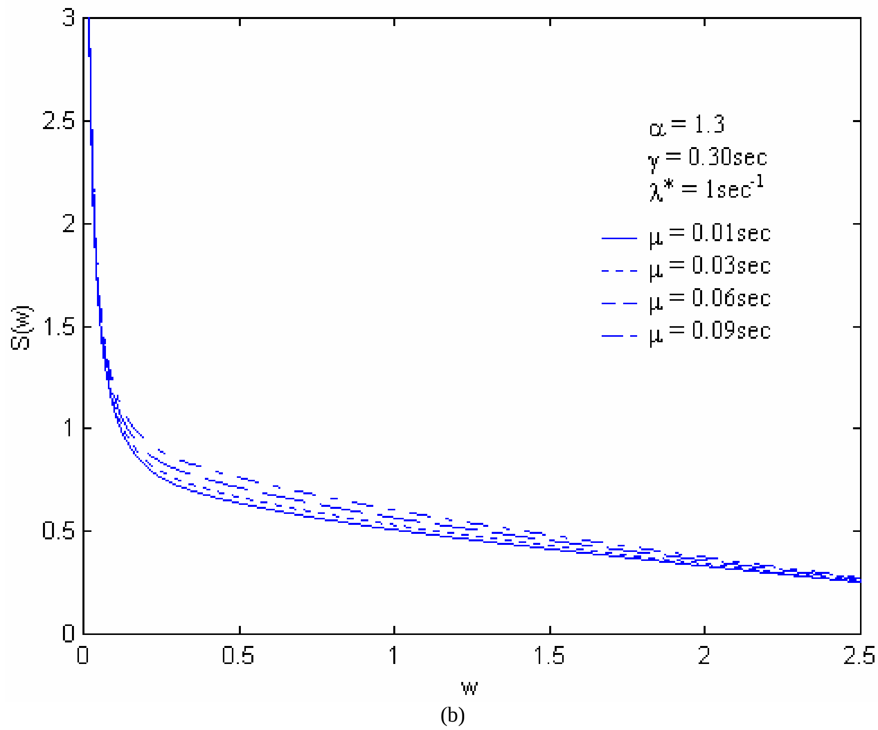
$$E_n\{\chi(n)\} = \lambda + \frac{2\lambda^2}{\omega^2} - 2\exp\{-\frac{\lambda\omega^2}{\lambda^2+\omega^2}\}\left[\frac{\lambda^2}{\omega^2}\cos\frac{\lambda^2\omega}{\lambda^2+\omega^2} + \frac{\lambda}{\omega}\sin\frac{\lambda^2\omega}{\lambda^2+\omega^2}\right]\cdot \tag{25}$$

Hence, power-spectral density can be obtained by using (25) in (17).

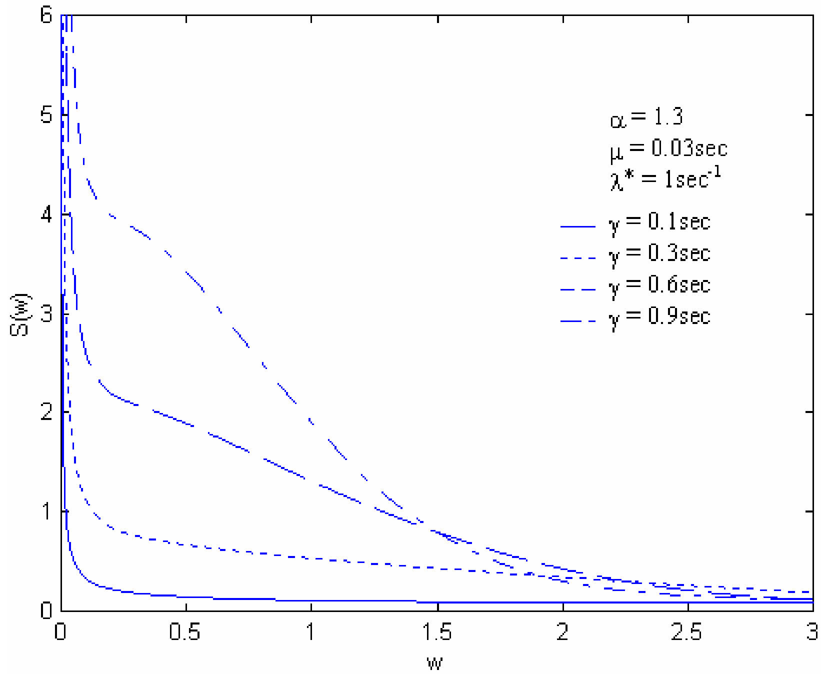
Note that in order to guarantee the convergence of equations (24) $|\lambda\rho| \le 1$ and, since $\rho = \frac{\lambda}{\sqrt{\lambda^2+\omega^2}}$, the convergence for $\omega > 0$ holds if $\lambda \le 1$. Thus, the time scale can conveniently be normalized so that event points τ_i occur at a rate of $\lambda^*=1$. This implies that ξ_i are scaled by a factor of $1/\lambda$. Thus, using property (6-b), parameters μ and γ in (15) become, respectively, $\frac{\mu}{\lambda}$ and $\frac{\gamma}{\lambda}$ while the stability index remains unchanged.


μ = 0.03sec
γ = 0.30sec
λ* = 1sec-1
α = 0.9
α = 1.1
α = 1.2
α = 1.3
S(w)
w

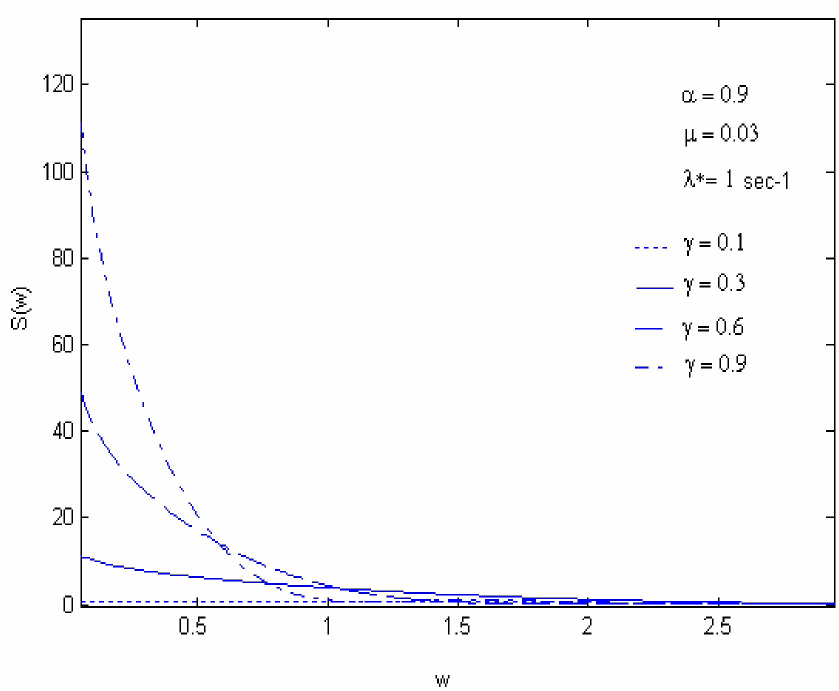
(a)


α = 1.3
γ = 0.30sec
λ* = 1sec-1
μ = 0.01sec
μ = 0.03sec
μ = 0.06sec
μ = 0.09sec
S(w)
w

(b)



(c)



(d)

Fig. 1. Spectral- Power Density as Function of Call Activity Parameters

Figure 1 shows on a normalized scale, the impact of the stability index on the spectral performance. Results reported in the open literature vary widely. Results are presented by assuming a processing speed of 10Mbps and mean files size of 30MB with a dispersion of 300MB. This does not preclude the existence of larger files. This leads to μ and γ in the order of 30 and 300 msec respectively.

Figure 1(a) shows that the heavier the activity tail is (smaller α), the stronger the infrared catastrophe is[4]. Figures 1(b) and 1(c) show the impact of the mean file size (μ) and the file size dispersion (γ) on the spectral content, and it can be observed that changes due to the variation of μ are minor while the impact of the dispersion is significant. The impact of the dispersion (γ) in the low frequencies power content is more significant when the tail becomes heavier (smaller alpha), as it can be seen in (d) compared to (c).

4 Effective Bandwidth of the Proposed Model

Effective bandwidth is a concept often used in traffic analysis in order to model the channel utilization or channel availability. Shot noise modeling allows a natural representation of the workload (traffic load) in terms of the On-Off activity (considering *h(t)* of equation (1) as a shifted rectangular function).

This section provides an effective bandwidth representation as a complementary technique for the analysis of the shot noise model proposal depicted by Figure 2(a), where the parameters ξ and τ represent the duration and occurrence of the activity span respectively described in the model proposal.

The effective bandwidth of an on-off source model $\alpha(\theta,t)$ is defined by the effective bandwidth of its activity period $\alpha_1(\theta,t)$ as well as the proportion of time (p) spent in the active period. $\alpha(\theta,t)$ is given by the following equation:

$$\alpha(\theta,t) = \frac{1}{\theta t}\log[1 + p(\exp(\theta t \alpha_1(\theta,t)) - 1)], \tag{26}$$

Effective bandwidth representation

In congruence with the Shot noise model, the effective bandwidth of the *on* state is one workload unit, since p is the proportion of time spent in activity period, then $p = \frac{E\{\xi_i\}}{E\{\tau_i\}}$

Consequently, the equation (26) is simplified as follows:

$$\alpha(\theta,t) = \frac{1}{\theta\, t}\log\left[1 + \frac{E\{\xi_i\}}{E\{\tau_i\}}\exp(\theta\, t - 1)\right], \tag{27}$$

τ_i are random variables (r.v.) with a exponential distribution, i.e., its PDF is given by $\lambda e^{-\lambda x}$, and the first moment of τ_i is $1/\lambda$. ξ_i are totally (positively) skewed stable and self similar r.v. For the purpose of this work, equation (27) is valid when the stability index of ξ_i is within the interval [1,2). This section focuses on finite effective bandwidth and ξ_i has a *one-sided* Laplace transform, meaning that it could not be used to obtain low order moments.

[4] The term *infrared catastrophe* term is used in fractal theory to refer to the high spectral content at low frequencies (Mandelbrot, 1999a; Mandelbrot, 1999b).

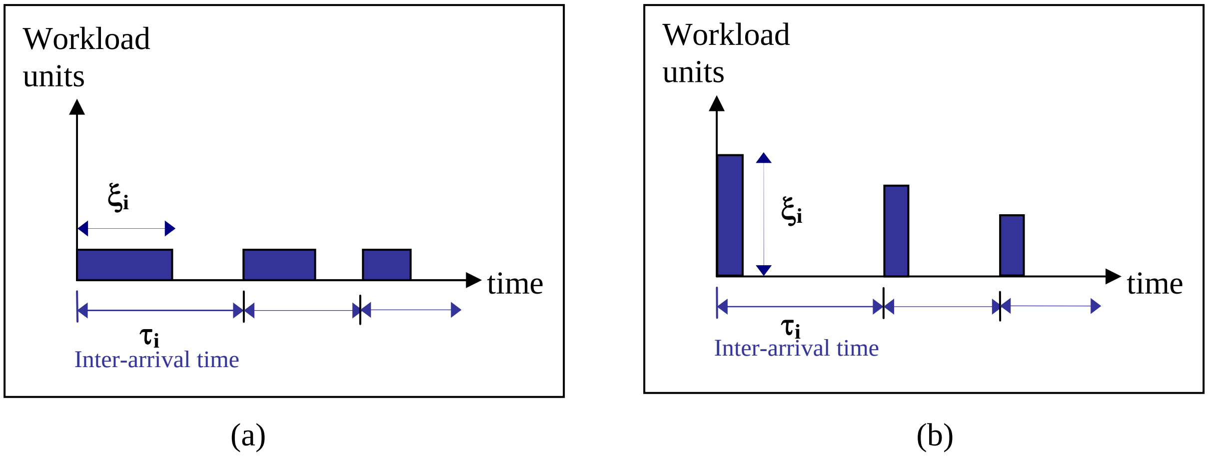


Fig. 2. Effective bandwidth of shot noise representation and equivalence

We recall that $k-order$ moments , for heavy tail distributions, are unbounded for $k > \alpha$ (Zolotarev, 1986). Moreover, the increments of a LFSM (Linear Fractional Stable Motion) present the LRD phenomenon when $H > 1/\alpha$, that is LRD when $1 < \alpha \le 2$. This work focuses on WWW traffic characterized as LRD stable process, where this process has finite mean. So, the effective bandwidth could be represented by numerical examples.

Effective bandwidth equivalent representation

Previous subsection presents the effective bandwidth representation of the line activity depicted by figure 2(a). However, the line activity has an equivalent representation by figure 2(b), where the effective bandwidth of the *on* state, is a Totally Skewed stable distribution $S_\alpha(\gamma,1,\mu)$, instead of one workload unit as in figure 2(a). Consequently, the proportion of time spent in activity period is $p = \frac{1}{E\{\tau_i\}}$.

Since $X \sim S_\alpha(\gamma,1,\mu)$, $E\{e^{-\theta X}\}$ can be represented by the following equation (Harmantzis, 2003):

$$E\{e^{-\theta X}\} = \exp\{-[\cos(\pi\alpha/2)]^{-1}\gamma^\alpha\theta^\alpha\} \quad \theta \ge 0 \tag{28}$$

And taking t he Norros model to represent the input process $X_{H,\alpha}(t)$ of the cumulative traffic during the time interval [0,t]. The generalization of the Norros model presenting LFSM instead of Fractional Stable Motion (FSM) is described by the following equation:

$$X_{H,\alpha}(t) = mt + kL_{\alpha,H}(t), \quad t \ge 0, \tag{29}$$

where *m* is the mean input rate, $L_{\alpha,H}(t)$ is the LFSM process, with α and H as the stability and self-similarity index, respectively. In this case, the effective bandwidth of a self-similar totally skewed LFSM source is given by (Harmantzis, 2003):

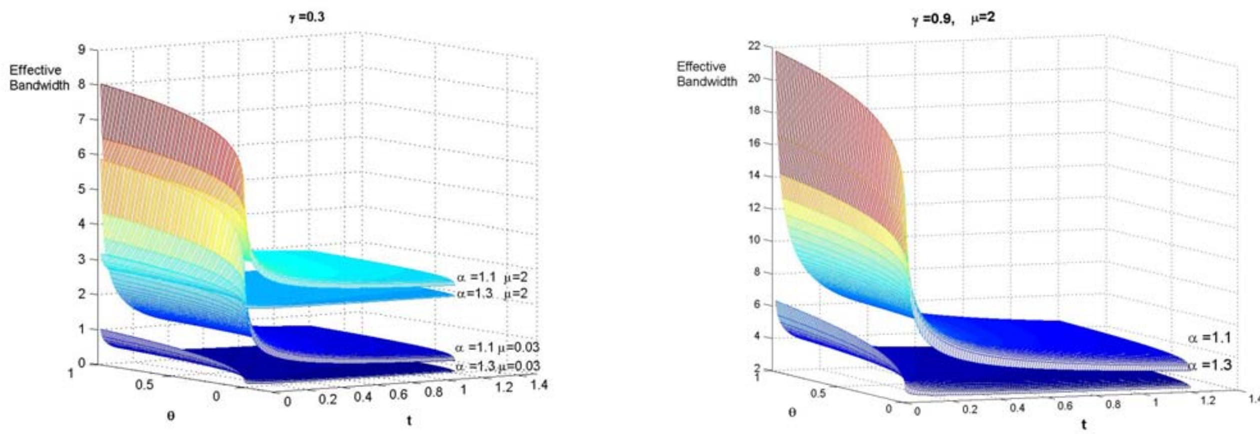
$$\alpha_1(\theta,t) = m - (\cos\pi\alpha/2)^{-1}\gamma^\alpha k^\alpha\theta^{\alpha-1}t^{\alpha H-1}. \tag{30}$$

Consequently, equation (26) for this equivalent representation is expressed as follows:

$$\alpha(\theta,t) = \frac{1}{\theta t}\log[1+\frac{1}{E\{\tau_i\}}(\exp(\theta t(m-(\cos\pi\alpha/2)^{-1}\gamma^{\alpha}k^{\alpha}\theta^{\alpha-1}t^{\alpha H-1}))-1)]. \quad (31)$$

The effective bandwidth performance of the shot noise model is illustrated, for H=0.5 and $\lambda = 1$, in figure 3. Results show the impact of the stability index on the effective bandwidth and it can be seen that the heavier the activity tail is, the more resources must be reserved in order to full-fit the same QoS. Similar effect occurs when the location parameter is increased.

Results also show that increment in the dispersion variation affect the effective bandwidth more significantly than the variation on the stability index.



(a) (b)

Fig. 3. The impact of α-stable parameters on Shot noise effective bandwidth

5 Concluding Remarks

Heavy-tailed phenomena appear in multiple WWW traffic-related problems. Because some distributions describing these phenomena lack finite moments, their study demands alternative methodologies. The characteristic function is a useful tool to deal with the addition of random variables, and in this paper the spectral performance of a line activity is presented based on the α-stable representation of the Paretian random variables. The impact of different statistical parameters such as location, dispersion, and stability index on the spectral content has been discussed. Concluding that on heavier activity of the tail is heavier, the infrared catastrophe becomes stronger. Showing that, on heavier tail activity the infrared catastrophe becomes stronger. Additionally, the impact of the mean file variation is minor while the impact of the file size dispersion is significant on the spectral density. Furthermore, the impact of the dispersion in the low frequencies power content is more significant when the tail becomes heavier. In particular, it has been shown that on/off activity exhibits a $1/f$ performance and that major changes in the spectral content are due to variation in the stability index and/or on the dispersion parameter while the sensitivity to changes of the location parameter is minor. Generalized shot-noise description is shown to be suitable for analysis of a variety of heavy-tailed phenomena.

An effective bandwidth estimation of the proposed model is presented. The model can be used as a tool to determine the amount of network resources in order to satisfy the user requirements as a design mechanism. The impact of the α-stable parameters in the effective bandwidth was studied. Moreover, the effective bandwidth analysis can be applied on current networks to guarantee the quality of existing connections by the implementation of the admission control mechanism.

References

1. Adler R., R. Feldman, and M. Taqqu, “A Practical Guide to Heavy Tails: Statistical Techniques and Applications”, *Birkhauser,* 1998.
2. Berger, J. M. and B. Mandelbrot**,** *A New Model for Error Clustering in Telephone Circuits*, IBM Journal, 1963, pp. 224-236.
3. Crovella M. E. and A. Bestavros, “Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes”, *IEEE/ACM Trans. on Networking,* Vol 5, No. 6, 1997, pp. 835-846.
4. Fisher M.J. and C.M. Harris, *A Method for Analyzing Congestion in Pareto and Related Queues*, The Telecommunication Review, 1999, pp 15-27.
5. Gilbert E. N. and H. O. Pollack, *Amplitude Distribution of Shot Noise*, The Bell System Technical Journal, 1960, pp. 333-350.
6. Gradshteyn I.S. and I.M. Ryzhik, *Tables of Integrals Series and Products,* 5^{th} Edition, Academic Press, 1994.
7. Harmantzis, F.C., D. Hatzinakos and I. Lambadaris, “Effective Bandwidths and tail probabilities for Gaussian and Stable Self-Similar Traffic”, *IEEE International Conference on Communications* 2003, Vol. 3, 11-15 May 2003. pp. 1515 – 1520.
8. Johnson, N. L., S. Kotz and N. Balakrishnan, “Continuous Univariate Distributions”, *John Wiley*, 1994
9. Kelly F.P., “Notes on effective bandwidths”, Stochastic Networks: Theory and applications. Royal Statistical Society Lecture Notes, F.P. Kelly, S. Zachary and I.B. Ziedins Eds. London, U.K.: Oxford Univ. Press, Vol 4, 1996, pp. 141-168.
10. Krunz M. and Makowsky, **A.** *Modeling Video Traffic using M/G/∞ Input: A Compromise Between Markovian and Long Range Dependent models*, IEEE Journal on Selected Areas in Commun., Vol. 16, 1998, pp. 733-740.
11. Liu Z., N. Niclausse, and C. Jalpa-Villanueva, “Web traffic modeling and performance comparison”, *in E. Gelenbe: System Performance Evaluation Methodologies and applications*, CRC Press, 2000.
12. Leon-Garcia A., ”Probability and Random Processes for Electrical Engineering”, Addison Wesley, 1994.
13. Lowen S. B. and M. C. Teich, *Doubly stochastic Poisson Process driven by Fractal Shot Noise*, Physical Review A, Vol. 43, 1991, pp. 4192-4215.
14. Mandelbrot B. B., “The fractal Geometry of Nature”, W.H. Freeman and Company, 1977.
15. Mandelbrot B. B., “The Fractal Geometry of Nature”, W.H. Freeman, Co, NY, 1999a.
16. Mandelbrot B. B., “Multifractals and 1/f noise: Wild self-affinity in physics”, Springer Verlag, 1999b.
17. Park K. and W. Willinger**,** *Self similar Network Traffic and performance evaluation*, John Wiley, 2000.
18. Parulekar M. and A. Makowsky, *M/G/∞ Input Processes: A Versatile Class of Models for Network Traffic*, IEEE 6^{th} INFOCOM, 1997, pp.419-426.
19. Resnick S., “ Infinite Source Poisson Model with Heavy Tailed Transmission Times”, *Proceedings of the American University Conference: Heavy tails and Applications*, 1999.
20. Ritcher, J., “Application of Pareto Error Statistics to Hagelberger Codes”, *IEEE Transaction on Information Theory*, vol. 2, No. 4, 1960, pp 571-576.
21. Ryu B. **K.** and S. B. Lowen, *Point Process Approaches to the Modeling and Analysis of Self-Similar Traffic*, Proceedings of the IEEE INFOCOM, San Francisco CA., March 1996, pp. 1468-1475.
22. Samorodnitsky G. and M. S. Taqqu, *Stable non Gaussian Random Process: stochastic models with infinite variance*, Chapman and Hall, 1994.
23. Sussman S., “Analysis of the Pareto Model to Error Statistics on Telephone Circuits”, *IEEE Transaction Communications Systems*, 1963, pp. 213-221.
24. Zolotarev V.M., “One dimensional Stable Distributions”, American Mathematical Society, 1986.

David Muñoz Rodríguez *was born in Guadalajara, Mexico. He received his B.S., M.S., and Ph.D. degrees in electrical engineering from the Universidad de Guadalajara, Guadalajara, México, Cinvestav, México City, México, and University of Essex, Colchester, England, in 1972, 1976 and 1979, respectively. He has been Chairman of the Communications Department and Electrical Engineering Departament at Cinvestav, IPN. In 1992, he joined the Centro de Electrónica y Telecomunicaciones, Instituto Tecnológico y de Estudios Superiores de Monterrey, Monterrey, México, where he is the director. His research interests include transmission and personal communication systems.*

Francisco Dávila *was born in Celaya, Mexico in 1971. He received his Bachelors degree in Electronics and Communications in 1993 from Tec de Monterrey-CEM, Mexico and his M.Sc. degree in Electronics and Telecommunications in 2002 from ITESM, Mexico. From 1997 to 1999 he was in charge of Tec de Monterrey CCM Department of Telephony and from 1999 to 2003 he was in charge of Tec de Monterrey Mty Section of Telephony. He is currently working at Teledinamica Mexico, where his main activities are the design and implementation of voice and data communications solutions. His main research interests are in interference modeling for wireless voice and data networks.*

Marlenne Angulo Bernal, *was born in Mazatlán, Mexico, in 1976. She received her M.Sc. degree from the Department of Electronic and Telecommunications CICESE research center in Ensenada, Mexico. Currently, she is pursuing her Ph.D. degree at the CINVESTAV Research Center, Guadalajara, Mexico. Her research interests include Quality of Service and multimedia.*

Deni L. Torres-Roman *was born in Cuba. He received his Ph.D. degree in telecommunications from Technical University Dresden, Germany in 1986. He was a professor at the University of Oriente, Cuba. And he is the coauthor of a book about Data Transmission. He received the Telecommunications Research Prize in 1993 from the AHCIET Association and was the recipient of the 1995 Best Paper Award from the AHCIET Review, Spain. Since 1996 he has been an associate professor at the Research and Advanced Studies Center of IPN (Cinvestav- Mexico). His research interests include VLSI, hard and software designs for telecommunications area.*

Dr. John Fonseka *was born in Negombo, Sri Lanka. He received his B.Sc. (Hons.) degree in Electronics and Telecommunication Engineering from the University of Moratuwa, Sri Lanka, in 1980, his M.Eng. degree from the Memorial University of Newfoundland, Canada in 1985, and his Ph.D. degree in Electrical Engineering from Arizona State University in 1988. He is currently an Associate Professor of Electrical Engineering at the University of Texas at Dallas. His research interests include combined coded modulation, signaling through bandlimited channels, coding theory, information theory, satellite mobile communications, telecommunication networks, and coherent optical communications.*