

# Modelagem de Tráfego de Redes Utilizando Cascata Multifractal Generalizada

Flávio Henrique Teles Vieira<sup>1</sup>  
Lee Luan Ling<sup>2</sup>

**Resumo:** Neste artigo propomos um modelo multifractal de tráfego baseado em uma cascata multiplicativa que apresenta diferentes distribuições para os multiplicadores em cada estágio da cascata. Neste modelo, os multiplicadores são obtidos através da estimação de suas densidades de probabilidade com o uso dos métodos de Kernel e Aceitação/Rejeição aplicados a tráfego real de redes. Com o intuito de validar o modelo, realizamos análises estatísticas e do comportamento de fila dos processos de tráfego gerados segundo o modelo proposto. Verificamos também o desempenho do modelo em capturar as características dos traços reais de tráfego em comparação à outros modelos multifractais.

**Abstract:** In this paper we propose a multifractal traffic model that is based on a multiplicative cascade presenting specific multiplier distributions in each cascade stage. In the proposed model, the multipliers are obtained through the estimate of their probability densities found in real network traffic by using Kernel and Acceptance/Rejection methods. Statistical analysis and queueing behavior study were carried out for the model validation. Furthermore, we verify the model performance in capturing the traffic trace characteristics in comparison to other multifractal models.

## 1 Introdução

A análise de vários fluxos de tráfego mostra que muitos deles apresentam uma estrutura de correlação que decai mais lentamente do que a exponencial. Estes decaimentos lentos (longa-dependência) são observados em fluxos de tráfego de vídeo VBR [1], de redes Ethernet [2], de Internet [3] [4] e podem ser relacionados ao conceito de auto-similaridade. Sabe-se que auto-similaridade assintótica de segunda ordem implica em longa-dependência (LRD) e vice-versa [4]. Apesar de os modelos de tráfego com curta dependência serem significativos por sua simplicidade, eles não capturam a característica de longa dependência presente nos traços reais de tráfego. Há muitos estudos que revelam a alta variabilidade do tráfego Internet,

---

<sup>1</sup>Escola de Engenharia Elétrica e de Computação, Universidade Federal de Goiás, Av. Universitária, n.1488, Qd. 86, Bloco A, 74605-010, Goiânia, GO  
{flavio@eeec.ufg.br}

<sup>2</sup>Departamento de Comunicações (DECOM), Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas Caixa Postal 6101, 13.083-852 SP, Brasil  
{lee@decom.fee.unicamp.br}

ou seja, o tráfego apresenta rajadas em uma gama de escalas de tempo, em contraste da suposição de que rajadas de tráfego só existem em escalas curtas de tempo [2][4]. Foi mostrado que estas incidências de rajadas multiescalas tem um impacto significativo no desempenho das redes [2][4][5].

O termo auto-similaridade se refere normalmente a processos assintoticamente auto-similares de segunda ordem ou monofractais [6]. A auto-similaridade de um processo é medida através do parâmetro de Hurst  $H$ . No entanto, os resultados em [5] indicam que o desempenho de fila depende mais da variabilidade do tráfego em certas escalas de tempo do que do valor de  $H$ . Sabe-se que diferentes processos com longa dependência apresentando mesmo valor para o parâmetro de Hurst podem gerar comportamentos de fila sensivelmente diferentes [7]. Portanto, modelos com novas características devem ser propostos e estudados.

Além desses fatores, constata-se que para muitos processos de tráfego de rede, a energia em escala dos coeficientes wavelet ou os gráficos variância-tempo normalmente não se comportam de forma linear [8]. Muitos destes processos têm comportamento semelhante a fragmentos fractais com parâmetro de Hurst variado em diferentes escalas de tempo, ou seja, são multifractais [8]. De fato, o desempenho de fila depende fortemente das irregularidades de tráfego em escalas de tempo pequenas devido à dinâmica complexa de redes de dados [7]. Feldmann et.al discutem que esta combinação de comportamentos em escala é melhor representada por um processo multifractal [9]. Em [10], Erramilli et al. confirmam esta afirmação e indicam que este comportamento pode ter um impacto significativo no desempenho de filas. Em [11], Lévy Véhel e Riedi mostraram que uma versão multifractal do movimento Browniano fracionário (fBm) pode refletir mais apropriadamente as propriedades do tráfego de redes. Em [12], os autores introduziram o modelo MWM (Multifractal Wavelet Model) mostrando que este modelo prediz de forma mais precisa o comportamento de traços de tráfego TCP [4] e de Ethernet [2] quando comparado ao movimento Browniano fracionário (fBm). Em [13], Gao e Rubin mostraram que ambos tempo entre chegada de pacotes e tamanho de pacotes podem ser modelados por processos multifractais. Em [14], Krishna et al. propuseram uma cascata multiplicativa em que se assume a distribuição dos multiplicadores da cascata como sendo gaussiana (VVGGM-Variable Variance Gaussian Model) para modelar o tempo entre chegada de pacotes. No modelo VVGGM, as médias das distribuições dos multiplicares possuem um único valor. Outros trabalhos envolvendo modelagem multifractal e cascatas multiplicativas têm sido recentemente propostos [15][16].

Neste trabalho, propomos um modelo multifractal baseado em cascata multiplicativa onde os multiplicadores da cascata são gerados conforme as distribuições dos mesmos encontradas nos traços de tráfego de redes. Mais precisamente, propomos obter as distribuições dos multiplicadores usando o método de Kernel e então o método da Aceitação/Rejeição para a geração dos multiplicadores. Através de simulações com traços reais de tráfego verificamos a melhoria obtida na modelagem do tráfego de redes.

O artigo está organizado da seguinte forma: na seção 2, discursamos sobre processos multifractais baseados em cascata multiplicativa. Na seção 3, apresentamos o modelo multifractal proposto. Na seção 4, são mostrados os testes realizados para validar o modelo multifractal. Por fim, na seção 5, apresentamos as conclusões obtidas.

## 2 Processos Multifractais

O conceito de processo multifractal foi introduzido por Mandelbrot no contexto de turbulência [17]. Desde então a teoria multifractal é usada em vários campos como processamento de imagem, geofísica, etc. O tráfego de redes ao ser considerado multifractal significa que possui uma estrutura de forte dependência inerente, com incidência de rajadas em várias escalas [6]. Estas características fazem com que o desempenho de rede seja pior do que é estimado usando modelos Gaussianos e de curta-dependência. O conceito de ‘multifractalidade’ se estende facilmente a processos estocásticos:

**Definição 1** Um processo estocástico  $X(t)$  é multifractal se satisfaz a equação:

$$E(|X(t)|^q) = c(q)t^{\tau(q)+1} \quad (1)$$

onde  $t \in T$  e  $q \in Q$ ,  $T$  e  $Q$  são intervalos na reta real,  $\tau(q)$  e  $c(q)$  são funções com domínio  $Q$ . Além disso, assume-se que  $T$  e  $Q$  têm comprimentos positivos, e que  $0 \in T$ ,  $[0, 1] \subseteq Q$ .

A Definição 1 permite que descrevamos ‘multifractalidade’ em termos de momentos onde  $\tau(q)$  é a função de escala e  $c(q)$  é a fator de momento de um processo multifractal. Caso  $\tau(q)$  seja linear em  $q$  o processo é dito ser monofractal, caso contrário, é multifractal. Para processos auto-similares com parâmetro de Hurst  $H$  pode-se mostrar que  $\tau(q) = qH - 1$  e  $c(q) = E(|X(1)|^q)$  [6].

### 2.1 Cascata Binomial

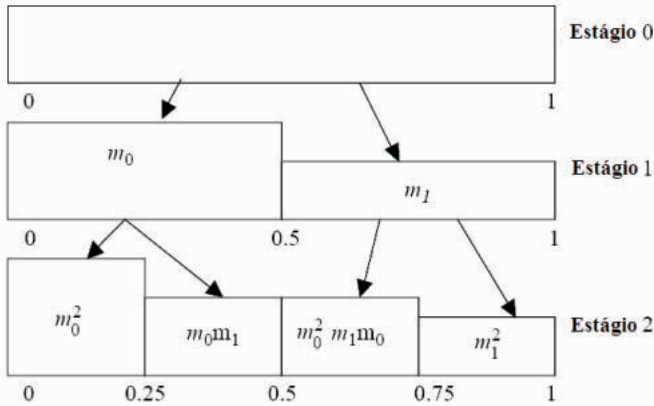
A cascata binomial é um método de se obter um processo multifractal que consiste de um procedimento iterativo no intervalo compacto  $[0,1]$ . Sejam  $m_0$  e  $m_1$  (multiplicadores da cascata) dois números positivos cuja soma é 1. No estágio  $k = 0$  da cascata, obtemos a medida inicial  $\mu_0$  do processo com probabilidade uniforme em  $[0,1]$ . No estágio  $k = 1$ , a medida  $\mu_1$  distribui massa utilizando a distribuição uniforme, sendo,  $m_0$  no subintervalo  $[0,1/2]$  e massa igual a  $m_0$  em  $[1/2, 1]$ . Em  $k = 2$ , o intervalo  $[0,1/2]$  é subdividido em  $[0,1/4]$  e  $[1/4,1/2]$  e o mesmo acontece com intervalo  $[1/2,1]$ , obtendo [6]:

$$\begin{aligned} \mu_2[0, 1/4] &= m_0 m_0 & \mu_2[1/4, 1/2] &= m_0 m_1 \\ \mu_2[1/2, 3/4] &= m_1 m_0 & \mu_2[3/4, 1] &= m_1 m_1 \end{aligned}$$

Considere o intervalo diádico  $[t, t + 2^{-k}]$  onde  $t = 0.\eta_1 \dots \eta_k = \sum_{i=1}^k \eta_i 2^{-i}$ . Sejam  $\varphi_0$  e  $\varphi_1$  as frequências relativas de 0's e 1's no desenvolvimento da cascata. A medida  $\mu$  no intervalo diádico é dada por:

$$\mu[t, t + 2^{-k}] = m_0^{k\varphi_0} m_1^{k\varphi_1} \tag{2}$$

Este processo preserva em cada estágio a massa dos intervalos diádicos por isso é chamado de cascata conservativa ou microcanônica. Em cada estágio da cascata os intervalos podem ser divididos em um número  $b \geq 2$  de intervalos de tamanhos iguais. Para  $b > 2$ , o processo resultante é definido como cascata multinomial. Se os multiplicadores usados têm um valor fixo para  $m_0$  e  $b = 2$ , então a cascata multiplicativa é binomial determinística com função de escala:  $\tau(q) = -\log_2(m_0^q + m_1^q) + 1$  [6].



**Figura 1.** Processo de Construção da Cascata Binomial.

## 2.2 Cascatas Multiplicativas

Ao se permitir que os multiplicadores da cascata sejam variáveis aleatórias independentes em  $[0,1]$  com densidade de probabilidade  $f_R(x)$ , obtém-se uma estrutura mais geral do que a determinística em que os multiplicadores são valores fixos (Figura 1). Dessa forma, o processo multifractal obtido  $\{\mu(\Delta t_k)\}_{k=1}^{2^N}$  terá no estágio  $i$  da cascata e no intervalo diádico de comprimento  $\Delta t_k = 2^{-k}$ , que começa em  $t = 0.\eta_1 \dots \eta_k = \sum_{i=1}^k \eta_i 2^{-i}$ , a medida  $\mu$ :

$$\mu(\Delta t_k) = R(\eta_1).R(\eta_1, \eta_2), \dots, R(\eta_1, \dots, \eta_k) \tag{3}$$

onde  $R(\eta_1, \dots, \eta_i)$  é o multiplicador no estágio  $i$  da cascata. Uma vez que os multiplicadores  $R(\eta_1, \dots, \eta_i)$  são i.i.d, pode-se demonstrar que a medida  $\mu$  satisfaz a relação de escala [15]:

$$E(\mu(\Delta t_k)^q) = (E(R)^q)^k = \Delta t_k^{-\log_2 E(R^q)} \quad (4)$$

que define um processo multifractal com função de escala  $\tau(q) = -\log_2 E(R^q)$ .

### 2.3 MWM (Multifractal Wavelet Model)

R. Riedi et al. propuseram o modelo multifractal MWM (Multifractal Wavelet Model) baseado na wavelet de Haar para caracterizar tráfego de redes [12]. O MWM se baseia em uma cascata multiplicativa no domínio wavelet. A transformada wavelet discreta é usada neste modelo devido a sua capacidade de representação multi-escala de sinais [18]. Para se gerar um processo segundo o modelo MWM é preciso aplicar a transformada wavelet discreta (DWT) ao tráfego de rede, calcular os momentos de segunda ordem dos coeficientes wavelet em cada escala, a média e a variância dos coeficiente na escala de maior resolução e calcular  $p_j$ , variável usada para capturar o decaimento de energia dos coeficientes wavelet em escala. O MWM aproxima com eficiência as propriedades do fluxo de tráfego original em relação a distribuição marginal (produz distribuição aproximadamente lognormal) e sua estrutura de correlação; tendo grande destaque na modelagem de tráfego [12][19].

### 2.4 Estimação da Densidade de Probabilidade dos Multiplicadores

Seja  $X^N$  o processo de tráfego no estágio  $N$  da cascata. A série de tráfego no estágio  $(N - 1)$  da cascata pode ser obtida agregando valores consecutivos do estágio  $N$  em blocos não-sobrepostos de tamanho 2. De forma análoga, dada a série na escala  $(N - j)$ ,  $X_i^{N-j}$  onde  $(i = 1, \dots, 2^{N-j})$ , obtemos os dados na escala  $(N - j - 1)$  pela soma consecutiva dos valores do estágio  $(N - j)$  da seguinte forma:

$$X_i^{N-j-1} = X_{2i-1}^{N-j} + X_{2i}^{N-j} \quad (5)$$

para  $i = 1, \dots, 2^{N-j-1}$ . Este procedimento termina quando a agregação dos valores forma apenas um ponto na última escala da cascata. Uma estimativa  $r_j^{(i)}$  dos multiplicadores da cascata pode ser obtida pela equação [9]:

$$r_j^{(i)} = \frac{X_i^{N-j}}{X_{2i-1}^{N-j-1}} \quad (6)$$

para  $i = 1, \dots, 2^{N-j-1}$ . Podemos considerar  $r_j^{(i)}$  como sendo amostras da distribuição  $f_{R_j}(r)$  dos multiplicadores no estágio  $j$ . A distribuição dos multiplicadores na escala  $j$  pode

ser obtida pelos histogramas de  $r_j^{(i)}$ . O modelo multifractal VVGM por exemplo, é uma cascata multiplicativa que aproxima os histogramas obtidos por gaussianas. Ou seja, neste modelo assume-se que a distribuição dos multiplicadores  $f_{R_j}(r)$  é gaussiana centrada em 0.5 e com variâncias que mudam a cada escala. Essas variâncias são estimadas a partir dos histogramas [24].

### 3 Cascata Multiplicativa com Distribuição Generalizada de Multiplicadores (CDGM)

Nesta seção apresentamos os métodos que compõem o modelo CDGM e na subseção 3.3, o algoritmo de síntese do modelo CDGM.

#### 3.1 Estimação Não-paramétrica de Distribuição de Probabilidade: Método de Kernel

A função de densidade de probabilidade é um conceito fundamental em estatística. Pode-se estimar densidade de probabilidade a partir de dados reais de forma paramétrica, ou seja, assumindo que os dados são oriundos de uma distribuição conhecida. Ou de forma não-paramétrica, a qual é investigada neste trabalho, a fim de encontrar a real distribuição de probabilidade para os multiplicadores da cascata.

O método de Kernel para estimação da função distribuição de probabilidade possui larga aplicação e suas propriedades são bem conhecidas. A não ser pelos histogramas, a estimação por Kernel é provavelmente a mais usada e certamente a mais estudada [18]. Sejam as amostras  $X_1, \dots, X_n$  obtidas com uma distribuição de probabilidade contínua  $f(x)$  que se deseja estimar. O método de Kernel usa uma soma de funções localizadas nas observações para as quais se obterá uma distribuição de probabilidade. Um estimador de  $f(x)$  com Kernel  $K$  é definido por:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (7)$$

onde  $h$  é a largura da janela, também chamada de parâmetro de suavização. Neste método, determina-se o formato do Kernel  $K$  e sua largura  $h$ . Deve-se notar que a escolha de um valor de  $h$  muito pequeno pode resultar em uma distribuição com muito ‘espúrios’, enquanto que se  $h$  é demasiado grande então uma suavização muito intensa pode ocorrer.

A função Kernel  $K$  deve satisfazer a seguinte condição:

$$\int_{-\infty}^{\infty} K(x) dx = 1 \quad (8)$$

Ao se considerar a discrepância entre a densidade de probabilidade estimada  $\hat{f}(x)$  e a

real  $f(x)$ , uma medida comum de ser empregada é o erro quadrático integral médio (EQIM) dado por:

$$EQM(\hat{f}) = \int_{-\infty}^{\infty} E\{\hat{f}(x) - f(x)\}^2 dx + \int_{-\infty}^{\infty} var \hat{f}(x) dx \quad (9)$$

O valor ideal para a largura da janela  $h$  do ponto de vista da minimização do erro quadrático integral médio (EQIM) pode ser dado por [20]:

$$h_{ot} = K_2^{-\frac{2}{5}} \left\{ \int_{-\infty}^{\infty} K(t)^2 dx \right\}^{\frac{1}{5}} \left\{ \int_{-\infty}^{\infty} f''(x)^2 dx \right\}^{-\frac{1}{5}} n^{-\frac{1}{5}} \quad (10)$$

Consideramos o Kernel como sendo gaussiano. Neste caso, o valor ótimo para a largura da janela  $h$  é dado por [20]:

$$h_{ot} = \frac{4}{3} \sigma n^{-\frac{1}{5}} = 1.06 \sigma n^{-\frac{1}{5}} \quad (11)$$

Na Figura 2 são mostrados o histograma obtido usando o método da seção 2.4 e a densidade para os multiplicadores entre os estágios 1 e 2 calculada utilizando o método de Kernel. Nota-se que se supor a distribuição como sendo gaussiana como é feito pelo modelo VVGM pode não ser realista. De fato, como pode ser visto neste caso, a distribuição dos multiplicadores da cascata tende a ser lognormal.

### 3.2 Método da Aceitação/ Rejeição

Seja  $X$  uma variável cuja distribuição de probabilidade seja  $f(x)$ . Supondo que exista uma função  $g(x)$  tal que:

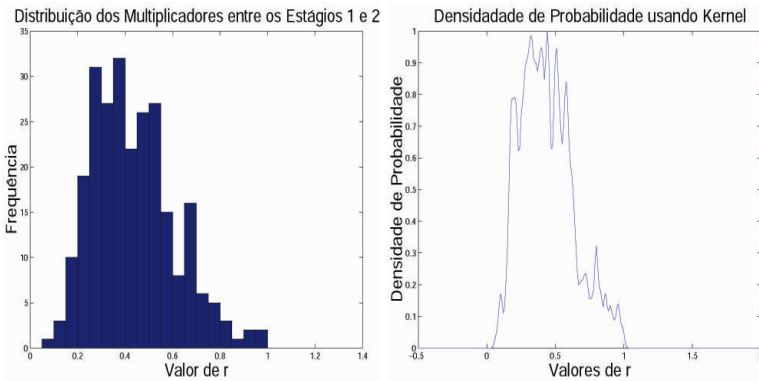
$$f(x) \leq cg(x) \quad (12)$$

onde  $c > 1$  é uma constante conhecida que satisfaz a equação (12) para todo  $x$ . Assim sendo, a função  $q(x) = cg(x)$  constitui uma envoltória de  $f(x)$ . A geração de amostras da variável  $X$  pelo método da Aceitação/Rejeição consiste basicamente da execução do seguinte algoritmo [21]:

#### Algoritmo de Aceitação/ Rejeição:

Repetir

1. Gere amostras  $y$  de  $g(x)$



(a) Traço de tráfego Dec-pkt-1

(b) Traço de tráfego Dec-pkt-1

**Figura 2.** Distribuição dos Multiplicadores entre os estágios 1 e 2

2. Gere amostras  $u$  usando uma distribuição uniforme  $U(0,1)$   
Até que:
3.  $u \leq f(y)/cg(y)$

Faça  $X = y$ .

Pode-se demonstrar que o processo  $X$  obtido com o método da Aceitação/Rejeição possui a densidade de probabilidade requerida, gerando o seguinte teorema [22]:

**Teorema**[Método da Aceitação/Rejeição] A variável aleatória  $X$  gerada pelo método de Aceitação /Rejeição possui densidade de probabilidade  $f(x)$ .

Quanto mais próxima a função de envoltória  $g(x)$  da função  $f(x)$ , melhor a taxa de aceitação de amostras geradas. Para este método ser eficiente, a constante  $c$  deve ser escolhida de tal forma que a taxa de rejeição seja baixa. Uma maneira de se determinar o valor ótimo de  $c$  é utilizar a seguinte equação [21]:

$$c = \sup_x f(x)/g(x) \tag{13}$$

A partir da soma de gaussianas representada pela equação (7) obtida com o método de



Kernel descrito na seção anterior, podemos determinar a constante  $c$  da seguinte forma:

$$c = \max \hat{f}(x) = 1/(\min(\sigma_i^2).n_g.\sqrt{(2\pi)}) \quad (14)$$

onde  $n_g$  é o número de gaussianas usadas para estimação de  $f(x)$ . Além disso, se faz  $g(x)$  igual a distribuição uniforme, ou seja,  $g(x) = 1$  para todo  $x$ , com  $f(x)$  igual a mistura de distribuições gaussianas para gerar amostras com a distribuição desejada  $f(x)$ .

### 3.3 Síntese do Processo CDGM

Nesta seção apresentamos o algoritmo de síntese do processo CDGM. O algoritmo de síntese de tráfego proposto faz uso do valor agregado de tráfego na última escala da cascata, a densidade de probabilidade dos multiplicadores estimada através do método de Kernel e os multiplicadores obtidos em cada estágio com o método da Aceitação/Rejeição. Explícitamente, nós calculamos a distribuição  $\hat{f}_i^*(x)$  de probabilidade dos multiplicadores em cada estágio  $i = 1, 2, \dots, N$  através do método de Kernel pela seguinte equação:

$$\hat{f}_i^*(x) = \frac{1}{n\delta_p^i} \sum_{p=1}^n K\left(\frac{x - z_i}{\delta_p^i}\right) \quad (15)$$

onde a média  $z_i$  e a variância  $\delta_p^i$  são dadas respectivamente por:

$$z_i = E\{\hat{f}_i(x)\} \quad i = 1, 2, \dots, N \quad (16)$$

$$\delta_p^i = \left( \int_0^1 \frac{1}{(h_i)^2} K\left(\frac{x-y}{h_i}\right)^2 \hat{f}_i(y) dy \right) - \left\{ \int_0^1 \frac{1}{h_i} K\left(\frac{x-y}{h_i}\right) \hat{f}_i(y) dy \right\}^2 \quad (17)$$

onde  $p = 1, 2, \dots, n$ .

Note que estamos usando uma mistura de gaussianas para modelar a real densidade de probabilidade dos multiplicadores. Assim, o algoritmo de síntese do processo CDGM consiste basicamente de:

1. Calcula-se o valor agregado obtido na última escala  $N$ ;
2. Estima-se os multiplicadores usando a equação (6);
3. Utiliza-se o método de Kernel para se estimar a real distribuição dos multiplicadores;
4. Aplica-se o método da Aceitação/Rejeição na distribuição dada pela equação (15) para gerar os multiplicadores em cada estágio  $i$ ;
5. Com os multiplicadores disponíveis, gera-se amostras de um processo multifractal por meio do procedimento descrito na seção 2.1 (cascata multiplicativa).

## 4 Validação do Modelo Multifractal Proposto

Com o objetivo de comparar as características estatísticas do modelo CDGM e do tráfego real, assim como de outros modelos multifractais de tráfego, foram conduzidos alguns testes estatísticos entre os quais: média, variância, função de autocorrelação, momentos de ordem  $q$  e espectro multifractal. Além disso, para verificar a precisão do modelo CDGM em representar tráfego real, realizamos simulações para analisar o tamanho de fila e perda de bytes em um sistema alimentado pelo modelo CDGM. Utilizamos nas simulações traços de tráfego TCP/IP (dec-pkt-1.tcp e dec-pkt-2.tcp) obtidos da Digital Equipment Corporation<sup>3</sup>. Considerou-se amostras de tráfego em uma escala de agregação de 512ms para estes traços de tráfego TCP/IP, escala na qual os traços apresentam características multifractais [10]. Também usamos traços de tráfego Ethernet obtidos da Bellcore<sup>4</sup> que apresentam características auto-similares e multifractais [2]. Todos as simulações e testes foram realizados utilizando o *software* Matlab.

A Tabela 1 mostra algumas estatísticas para a série Bc-Aug da Bellcore. O modelo CDGM possui dentre as cascatas analisadas, parâmetro de Hurst cujo valor mais se aproxima do valor estimado para uma dada série de tráfego real. O Parâmetro de Hurst foi calculado utilizando o método baseado em wavelets descrito em [25].

**Tabela 1.** Média, Variância e Parâmetro de Hurst.

	Média	Variância	P.Hurst
BAug	$1,6282 \cdot 10^5$	$8,4269 \cdot 10^9$	0,8979
CDGM	$1,6282 \cdot 10^5$	$1,4037 \cdot 10^{10}$	0,9123
MWM	$1,6241 \cdot 10^5$	$6,7068^9$	0,7871
VVGM	$1,6282 \cdot 10^5$	$9,2856 \cdot 10^8$	1,1449

### 4.1 Função de Autocorrelação

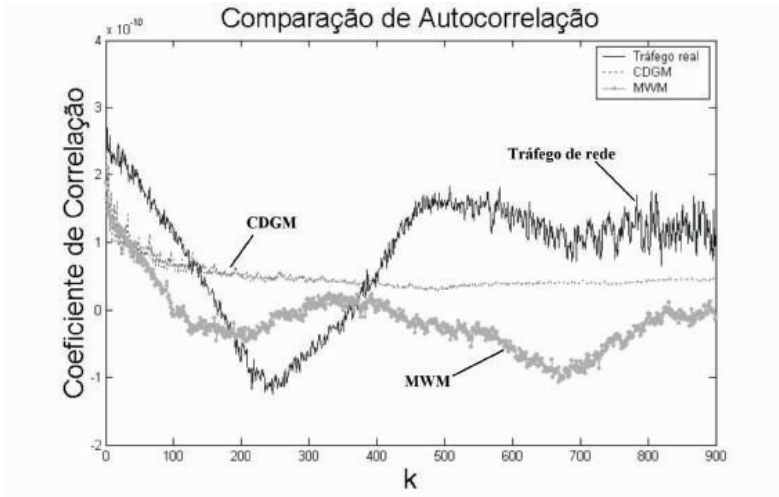
A função de autocorrelação e o coeficiente de correlação refletem as estatísticas de segunda ordem de uma série; dando uma idéia a respeito da longa-dependência nos dados. O coeficiente de correlação pode ser visto como a covariância normalizada de um processo. Seja uma série  $y(t)$  com média  $\mu_t$  e desvio-padrão  $\sigma_t$ , e a mesma série deslocada no tempo  $y(t+k)$  com média  $\mu_{t+k}$  e desvio-padrão  $\sigma_{t+k}$ . O coeficiente de correlação para este processo  $y(t)$  é

<sup>3</sup><http://ita.ee.lbl.gov/html/contrib/DEC-PKT.html>

<sup>4</sup><http://www.acm.org/sigcomm/ITA>

dado por:

$$\rho(k) = \frac{E[(y(t+k) - \mu_{t+k})(y(t) - \mu_t)]}{\sigma_{t+k}\sigma_t} \quad (18)$$



**Figura 3.** Coeficientes de correlação.

A Figura 3 mostra os coeficientes de correlação para o traço de tráfego real Dec-pkt-2, assim como para os traços obtidos segundo os modelos CDGM e MWM. Para maioria dos valores de  $k$  analisados, o modelo CDGM obteve coeficientes de correlação mais próximos aos dos traços reais do que o importante modelo MWM. Resultados semelhantes foram obtidos para todas as séries de tráfego analisadas.

#### 4.2 Momentos de Ordem $q$

O tráfego agregado pode influenciar fortemente o comportamento de fila e o desempenho dos multiplexadores de rede. Além disso, agregação de fluxos com diferentes intensidades ocorre em vários pontos da rede. Portanto, é importante investigar não somente o comportamento de fluxos individuais de tráfego mas também o do tráfego agregado. Cabe lembrar que uma das deficiências dos modelos monofractais é o de não conseguir capturar os momentos de mais alta ordem do tráfego de rede. Nesta seção, são analisados os momentos

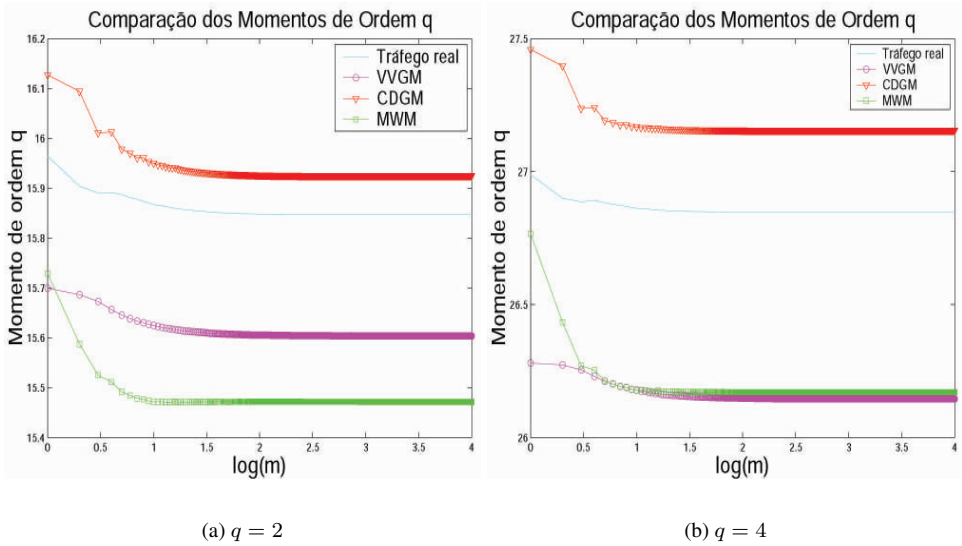


Figura 4. Momentos de Ordem  $q$  para a Série Bc-Aug

do tráfego agregado do CDGM. Seja o processo agregado  $X_k^m$  definido como

$$X_k^m = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i^N, \quad (19)$$

onde  $i = 1, 2, \dots, 2^N$ ,  $k = 1, 2, \dots, L$  e  $L = 2^N/m$  é o número total de agregações para um valor fixo de  $m$ . Estima-se o momento de ordem  $q$  do tráfego agregado da seguinte forma:

$$\hat{\mu}^{(m)}(q) = \frac{1}{L} \sum_{k=1}^L |X_k^m|^q \quad (20)$$

A Figura 4 apresenta os resultados das simulações realizadas para o traço de tráfego dec-pkt-1 na escala de agregação de 512ms, expressando a variação de  $\log_{10}(\hat{\mu}^{(m)}(q))$  versus  $\log_{10}(m)$  para  $q = 2, 3, 4$  e  $5$ . Para os modelos multifractais considerados, as curvas dos momentos de ordem  $q$  do tráfego agregado são obtidas pela média de 100 realizações do processo. Em relação aos outros modelos, o CDGM é o que melhor aproxima os momentos.

Esse fato indica que o modelo proposto pode seguramente reproduzir o comportamento em rajadas do tráfego multimídia nas redes atuais.

### 4.3 Espectro Multifractal

Em contraste a outros modelos de tráfego, processos multifractais contêm uma multiplicidade de expoentes de Hölder locais dentro de qualquer intervalo finito [6]. Os expoentes de Hölder descrevem as características em escala locais de um processo em um determinado ponto no tempo. O conceito de expoente de Hölder está relacionado com a singularidade local de um processo, ou seja, caracteriza a sua suavidade (quantidade de rajadas) em um certo instante de tempo [23]. A distribuição destes expoentes pode ser representada por uma densidade normalizada chamada espectro multifractal. Em outras palavras, o espectro multifractal descreve a dimensão fractal do conjunto de instantes possuindo um dado expoente local [12]. O espectro multifractal  $f(\alpha)$  de um processo  $X(t)$  pode ser visto como a transformada de Legendre de  $\tau(q)$  (função de escala definida anteriormente) pela relação:

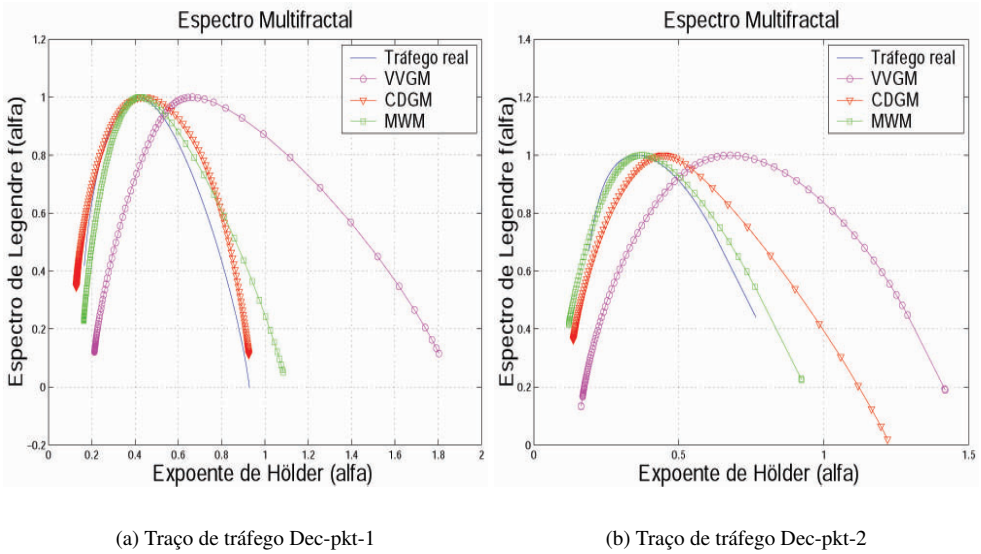
$$f(\alpha) = \min_q \{q\alpha - \tau(q)\} \quad (21)$$

Através da transformada de Legendre calculamos o espectro multifractal para os processos gerados segundo os modelos CDGM, MWM e VVGM. Conforme pode ser visto na Figura 5, os espectros multifractais dos traços de tráfego dec-pkt-1 e dec-pkt-2 apresentam  $\alpha < 1$ , o que indica uma alta incidência de rajadas multiescala. Pode-se observar que o modelo CDGM captura com eficiência o espectro multifractal e por conseqüência a ‘multifractalidade’ dos traços de tráfego reais. O modelo VVGM, por sua vez, mostrou um espectro mais largo porém menos fiel ao do tráfego real.

### 4.4 Testes de Desempenho e Verificação do Comportamento de Fila

Consideramos como modelo de simulação do enlace, um servidor com *buffer* finito sendo alimentado pelo processo CDGM a fim de avaliarmos o desempenho do modelo proposto em descrever o comportamento de fila no *buffer* para traços reais de tráfego. Nesta seção, analisamos principalmente a ocupação média de bytes no *buffer* e a porcentagem de perda associadas a utilizações do *buffer*.

A utilização do *buffer*  $\lambda$  é definida como sendo a razão entre o tempo total de serviço fornecido aos pacotes pelo tempo total de uso do *buffer*. Um valor de  $\lambda$  próximo de 1 indica que o *buffer* está constantemente sendo usado e conseqüentemente há uma maior probabilidade de descarte de dados. Um valor abaixo de 0.4, significa que o *buffer* está sendo subutilizado, um valor entre 0.6 e 0.7 é considerado ideal em redes reais. A utilização do *buffer* depende da capacidade do servidor, ou seja, da taxa em que os dados são transmitidos.



**Figura 5.** Espectro multifractal de Legendre

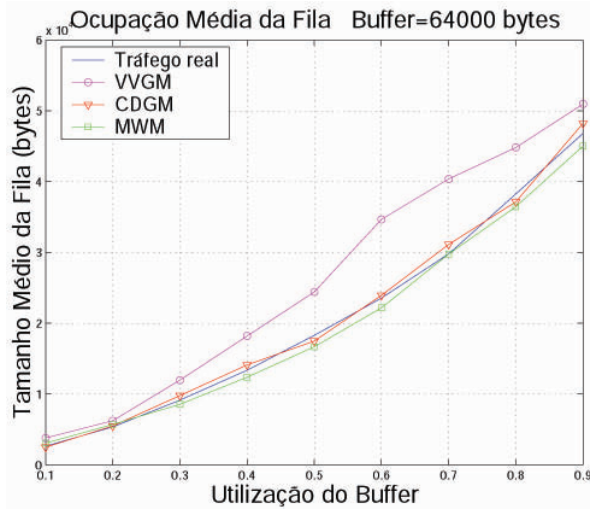
A Figura 6 mostra a ocupação média de bytes para um *buffer* finito de tamanho 64Kbytes versus a utilização do mesmo. Os resultados relacionados a ocupação média no *buffer* indicam que o modelo proposto descreve com precisão o comportamento de fila de fluxos reais de tráfego.

Em seguida, analisamos o desempenho do modelo CDGM em estimar a probabilidade de perda de bytes em um enlace de rede através dos processos de tráfego gerados segundo este modelo. Dado um tamanho de *buffer*  $x$ , estimou-se a probabilidade de perda de *bytes*  $P(Q > x)$ , sendo  $Q$  o processo de tamanho da fila no *buffer*, pela seguinte expressão [26]:

$$P(Q > x) \cong \frac{N_x}{N_t} \tag{22}$$

onde  $N_x$  corresponde ao número de *bytes* descartados que não podem ser armazenados no *buffer* de tamanho  $x$  e  $N_t$  é o número total de *bytes* a ser atendido.

A Figura 7.a mostra a relação entre a porcentagem de perda de bytes e o tamanho do *buffer* para o traço de tráfego dec-pkt-1. Na Figura 7.b alocamos como taxa, a média do traço de tráfego dec-pkt-1. Nota-se que o modelo CDGM apresenta melhores resultados, principal-



**Figura 6.** Ocupação média do *buffer* para o traço de tráfego dec-pkt-1.

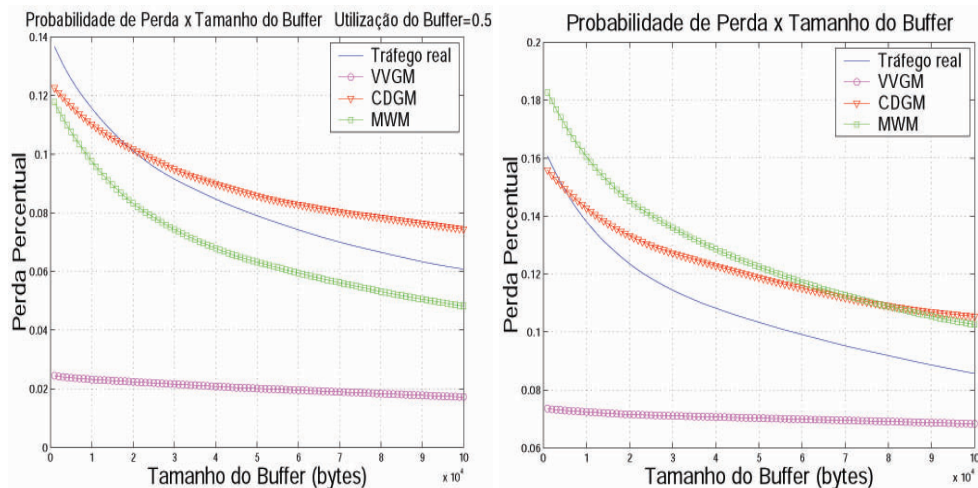
mente para *buffers* de tamanhos menores. Ou seja, obtém-se estimativas mais precisas para a probabilidade de perda para tráfego de redes ao incorporar informações adequadas sobre os multiplicadores no procedimento de geração de processo multifractal baseado em cascata.

## 5 Conclusões

Pode-se concluir que o modelo multifractal proposto consegue capturar com eficiência as características do tráfego de redes. Essa afirmação é confirmada pelos testes realizados.

Notou-se que ao se admitir uma densidade de probabilidade gaussiana para os multiplicadores, uma maior imprecisão na modelagem pode ser obtida. Neste trabalho, propomos uma generalização da distribuição dos multiplicadores com a média da mesma variando em cada estágio; diferente do que ocorre no VVGM. Esta característica singular do modelo proposto permite que um bom desempenho de modelagem seja obtido.

O modelo MWM é outro modelo multifractal comparado e que tenta capturar as características do tráfego de redes através da modelagem do decaimento em escala da energia dos coeficientes wavelet dos processos de tráfego. As simulações revelam que os resultados do modelo CDGM são comparáveis e em alguns casos superiores ao do modelo MWM.



(a) Traço de tráfego Dec-pkt-1

(b) Traço de tráfego Dec-pkt-1

**Figura 7.** Probabilidade de perda de bytes versus tamanho do *buffer*

Alguns trabalhos afirmam que o tráfego de *backbone* Internet pode ser monofractal em algumas escalas de tempo [27]. Nestes casos, o modelo proposto também se aplica pois como modelo multifractal engloba o caso monofractal em sua estrutura. De fato, verificamos uma modelagem adequada para séries monofractais (por exemplo para a série dec-pkt-2).

Devido as características constatadas, conclui-se que o modelo CDGM é uma ferramenta importante para análise de tráfego de redes multimídia. Assim, o modelo proposto pode ser usado em várias aplicações da engenharia de tráfego de redes, como controle de admissão de fluxos, provisão de qualidade de serviço (QoS), entre outras.

## Referências

- [1] F. H. P. Fitzek e M. Reisslein. MPEG-4 and H.263 video traces for network performance evaluation. *IEEE Network*, 15(6):40-54, Nov/Dez 2001.
- [2] W. E. Leland, M. S. Taqqu, W. Willinger, e D. V. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1-15, 1994.



- [3] R. G. Addie, M. Zukerman, e T. D. Neame. Fractal traffic: Measurements, modelling and performance evaluation. Proceedings of IEEE Infocom '95, pp. 977-984, Abril 1995.
- [4] V. Paxson e S. Floyd. Wide-area traffic: The failure of Poisson modeling. IEEE/ACM Transactions on Networking, 3(3):226-244, 1995.
- [5] A. Erramilli, O. Narayan e W. Willinger. Experimental Queueing Analysis with Long-Range Dependent Packet Traffic, IEEE/ACM Trans. on Net., 4(2), Abril, 1996.
- [6] K.Park e W.Willinger. Self-similar Network Traffic and Performance Evaluation, John Wiley & Sons, New York, 2000.
- [7] M. Grossglauser e J.-C. Bolot. On the relevance of long-range dependence in network traffic. IEEE/ACM Transactions on Networking, 7(5):629-640, Outubro, 1999.
- [8] Q. Li e D. Mills. Investigating the Scaling Behavior, Crossover and Anti-persistence of Internet Packet Delay Dynamics, Globecom 3,1843-1852,1999.
- [9] A. Feldmann, A. C. Gilbert, e W.Willinger. Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic, Proc. of the ACM/SIGCOMM'98, Vancouver, B.C., pp.25-38, 1998.
- [10] A. Erramilli, O. Narayan, A. Neidhardt, e I. Saniee. Performance impacts of multi-scaling in wide area TCP/IP traffic. Proceedings of IEEE Infocom 2000, Março, 2000.
- [11] J. Lévy Véhel e R. Riedi. Fractional Brownian motion and data traffic modelling: The other end of the spectrum. In J. Lévy Véhel, E. Lutton, and C. Tricot, editors, Fractals in Engineering. Springer, 1997.
- [12] R. H. Riedi, M. S. Crouse, V.J. Ribeiro, e R.G. Baraniuk. A Multifractal Wavelet Model with Application to Network Traffic, IEEE Trans. on Information Theory 45(3), Abril, 1999.
- [13] J. Gao e I. Rubin. Multifractal analysis and modeling of long range-dependent traffic. Proceedings of ICC'99, Junho, 1999.
- [14] M. Krishna et al. Global scaling exponent for variable variance gaussian multiplicative (VVGMM) multifractal cascades. Proc. of the Sixth Biennial Conference on Signal Processing and Communications (SPCOM'01), pp.19-25, Julho, 2001.
- [15] D. Zou, X. Zhang, W. Wang. Multi-service traffic models of heterogeneous wireless communication networks. Proc. of the 7th World Congress on Intelligent Control and Automation, pp.495-498, Junho, 2008.

- [16] Y. Xu. A Network Traffic Model Based on Fractal. Proc. of International Conference on Wireless Communications, Networking and Mobile Computing, Set., 2007.
- [17] B. B. Mandelbrot, A. Fisher, e L. Calvet. A Multifractal Model of Asset Return. Yale University, 1997.
- [18] C. K. Chui. An Introduction to Wavelets. San Diego: Academic, 1992.
- [19] V.J.Ribeiro, R.H.Riedi, M.S.Crouse e R.G.Baraniuk. Multiscale queueing analysis of long-range dependent traffic, Proc. IEEE Infocom, Março, 2000.
- [20] B. Silverman. Density Estimation for Statistics and Data Analysis. Chapman and Hall, London, 1986.
- [21] K. Monahan. Numerical Methods of Statistics. Cambridge University Press, NY, 1998.
- [22] L. Devroye. Non-Uniform Random Variate Generation. Springer- Verlag, New York, 1986.
- [23] Z. Zhang, V. Ribeiro, S. Moon e C. Diot. Small-Time Scaling behaviors of internet backbone traffic: An Empirical Study, IEEE Infocom, San Francisco, Março, 2003.
- [24] M. P. Krishna, V. M. Gadre e U. B. Dessay. Multifractal based network traffic modeling. Kluwer Academic Publishers, 2003.
- [25] D. Veitch e P. Abry. A wavelet based joint estimator for the parameters of LRD. IEEE Trans. Info. Theory. 45(3), 1999.
- [26] D. A. Rolls, G. Michailidis e F. Hernández-Campos. Queueing analysis of network traffic: methodology and visualization tools. Computer Networks 48, pp. 447-473, Janeiro, 2005.
- [27] D. Veitch, N. Hohn e P. Abry. Multifractality in TCP/IP Traffic: The case Against. Computer Networks 48. pp. 293-313, Janeiro, 2005.