

# Modelagem Multifractal utilizando Cascata Multiplicativa com Distribuição Generalizada de Multiplicadores

Flávio Henrique Teles Vieira e Lee Luan Ling

**Resumo**— Neste artigo propõe-se um modelo multifractal de tráfego baseado em uma cascata multiplicativa com distribuição generalizada de multiplicadores (CDGM). Neste modelo, os multiplicadores são obtidos através da estimativa de suas densidades de probabilidade encontradas no tráfego real de redes com o uso dos métodos de Kernel e Aceitação/Rejeição. Análises estatísticas e o comportamento de fila foram realizadas na validação do modelo, verificando seu desempenho em capturar as características dos traços de tráfego em comparação à outros modelos baseados em cascata multiplicativa.

**Palavras-Chave**— Tráfego, Redes, Multifractal, Cascata Multiplicativa.

**Abstract**— In this article we propose a multifractal traffic model that is based on a multiplicative cascade with generalized multiplier distributions. In the proposed model the multipliers are obtained through the estimate of their probability densities encountered in real network traffic by using Kernel and Acceptance/Rejection methods. Statistical analysis and queueing behavior study were accomplished for the model validation, verifying its performance in capturing the traffic trace characteristics in comparison to other multiplicative cascade based models.

**Keywords**— Network, Traffic, Multifractal, Multiplicative Cascade.

## I. 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 são observados em tráfegos de vídeo VBR [1], de LANs Ethernet [2], de redes MAN [3], de Internet [4] e podem ser relacionados ao conceito de auto-similaridade. Enquanto os modelos de tráfego de curta-dependência são significativos por sua simplicidade, eles não capturam as dependências de longa duração presentes nos traços reais de tráfego. Há muitos estudos que revelam a alta variabilidade do tráfego Internet, ou seja, o tráfego é 'bursty' 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]. O parâmetro de Hurst mede o grau de

Flávio Henrique Teles Vieira e Lee Luan Ling, Departamento de Comunicações, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, E-mails: flavio@decom.fee.unicamp.br, lee@decom.fee.unicamp.br. Este trabalho foi parcialmente financiado pela Capes.

auto-similaridade do processo. 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  $H$ . É sabido que diferentes processos com longa dependência com o mesmo valor do parâmetro de Hurst podem gerar comportamentos de fila sensivelmente diferentes [7]. Portanto, modelos com novas características devem ser propostos e estudados.

Para muitos processos de tráfego de rede, a energia em escala dos coeficientes wavelet ou os gráficos variância-tempo normalmente não se comportam linearmente. Muitos destes processos têm comportamento de fragmentos fractais com parâmetro de Hurst variado em diferentes pequenas escalas de tempo [8], ou seja, são multifractais. De fato, o desempenho de fila depende grandemente das irregularidades de tráfego em escalas de tempo pequenas devido à dinâmica complexa das 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. confirmaram esta afirmação, e indicaram que este comportamento pode ter um impacto significativo no desempenho de filas. Em [11], os autores apresentaram o modelo MWM (Multifractal Wavelet Model) mostrando que este modelo prediz melhor o comportamento de traços de tráfego TCP [4] e de Ethernet [2] quando comparado com o movimento Browniano fracionário (fBm). Em [12], Gao e Rubin mostraram que ambos tempo entre chegada de pacotes e tamanho de pacotes podem ser modelados por processos multifractais. Em [13], Krishna et al. propuseram uma cascata multiplicativa em que se assumia a distribuição dos multiplicadores da cascata como sendo gaussiana (VVGM-Variable Variance Gaussian Model) para modelar o tempo entre chegada de pacotes. Além disso, no modelo VVGM, as médias destas distribuições são fixas em 0.5.

Neste trabalho, estendemos o modelo VVGM em termos da distribuição dos multiplicadores da cascata, verificando a melhoria obtida na modelagem do tráfego de redes. Propomos obter as distribuições dos multiplicadores usando o método de Kernel e então o método da Aceitação/Rejeição para a geração dos multiplicadores.

O artigo está organizado da seguinte forma: na seção II, discutimos sobre processos multifractais baseados em cascata multiplicativa. Na seção III, apresentamos o modelo multifractal proposto. Na seção IV, são mostrados os testes realizados para validar o modelo multifractal. Por fim, na seção V, apresentamos as conclusões obtidas.

## II. PROCESSOS MULTIFRACTAIS

O conceito de multifractal foi introduzido por Mandelbrot no contexto de turbulência nos anos 70 [14]. 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 se consideravam 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  $T$  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  $0 \in T$ ,  $[0, 1] \subseteq Q$ .

Assim, descrevemos multifractalidade em termos de momentos onde  $\tau(q)$  é a função de escala e  $c(q)$  é o fator de momento de um processo multifractal. Se  $\tau(q)$  é linear em  $q$ , o processo é chamado monofractal, caso contrário, é multifractal. Para processos auto-similares com parâmetro de Hurst  $H$  pode-se mostrar que  $\tau(q) = qH - 1$  e  $c(q) = E(|X(1)|^q)$ .

### A. 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  $m_1$  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]:

$$\begin{aligned} \mu_2[0,1/4] &= m_0m_0 & \mu_2[1/4,1/2] &= m_0m_1 \\ \mu_2[1/2,3/4] &= m_1m_0 & \mu_2[3/4,1] &= m_1m_1 \end{aligned}$$

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 freqüências relativas de 0's e 1's no desenvolvimento da cascata. A medida  $\mu$  no intervalo diádico é dada por:

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

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  $b > 2$  partes de tamanhos iguais, este processo 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].

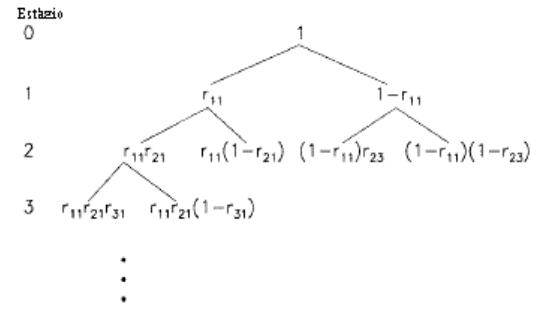


Fig. 1. Processo de Construção da Cascata

### B. Cascatas Multiplicativas

Ao se permitir que os multiplicadores da cascata  $r$  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 terá no estágio  $k$  da cascata 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  $\mu$ :

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

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

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

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

### C. MWM (Multifractal Wavelet Model)

Rudolf Riedi et al. propuseram o modelo multifractal MWM (Multifractal Wavelet Model) baseado na wavelet Haar para caracterizar tráfego de redes [11]. O MWM se baseia em uma cascata multiplicativa no domínio wavelet. A transformada wavelet discreta é usada neste modelo dada sua capacidade de representação multiescala de sinais [15]. Para se gerar um processo segundo o modelo MWM é preciso aplicar a transformada wavelet discreta (DWT) Haar 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 com a 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 [11][16].

### D. Estimação da Densidade de Probabilidade dos Multiplicadores

Sejam os dados de tráfego no estágio  $N$  da cascata,  $X_i^N$ . 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  $(N-j)$ ,  $X_i^{N-j}$  ( $i = 1, \dots, 2^{N-j}$ ), 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 dos multiplicadores pode ser obtida pela equação:

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

para  $i = 1, \dots, 2^{N-j-1}$ . Podemos considerar  $r_j^{(i)}$  como amostras da distribuição de multiplicadores  $f_{R_j}(r)$  no estágio  $j$ . A distribuição dos multiplicadores na escala  $j$  pode ser obtida pelos histogramas de  $r_j^{(i)}$ . O modelo multifractal VVGM é uma cascata multiplicativa que aproxima os histogramas obtidos por gaussianas. A distribuição dos multiplicadores  $f_{R_j}(r)$  neste modelo é gaussiana centrada em 0,5 com variâncias que mudam a cada escala. Essas variâncias são estimadas a partir dos histogramas.

### III. CASCATA MULTIPLICATIVA COM DISTRIBUIÇÃO GENERALIZADA DE MULTIPLICADORES (CDGM)

Nesta seção apresentaremos os métodos que compõem o modelo CDGM e na subseção C a proposta de síntese do modelo CDGM.

#### A. Estimação Não-paramétrica de Distribuição de Probabilidade: Método de Kernel

Estimação de densidade de probabilidade a partir de dados reais pode ser feita de forma paramétrica, ou seja, assumindo que os dados são oriundos de uma distribuição conhecida. Ou de forma não-paramétrica, a qual será investigada neste trabalho, a fim de encontrar uma distribuição mais apropriada 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. Sejam as amostras  $X_1, \dots, X_n$  obtidas com uma distribuição de probabilidade contínua  $f(x)$  a qual queremos estimar. O método de Kernel usa uma soma de funções localizadas nas observações para as quais se obterá uma distribuição de probabilidade. Um estimador 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 ser notado que se um valor de  $h$  muito pequeno é escolhido, pode-se resultar numa 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:

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

O valor ideal para a largura da janela  $h$  do ponto de vista da minimização do erro quadrático integral médio (EQIM) é descrito em [17]. Neste trabalho, usamos Kernel gaussiano, assim, o valor ótimo para a largura da janela  $h$  é dado por [17]:

$$h_{ot} = \frac{4}{3} \sigma n^{-\frac{1}{5}} = 1,06\sigma n^{-\frac{1}{5}} \quad (10)$$

Na Figura 2 são mostrados o histograma obtido usando o método da seção II-D e a densidade para os multiplicadores entre os estágios 1 e 2 utilizando o método de Kernel. Nota-se que se supor a distribuição como sendo gaussiana de média 0,5 como é feito pelo modelo VVGM pode não ser realista. De fato, como neste caso, em que se tende a uma distribuição lognormal. A inserção do método de Kernel aumenta a complexidade computacional do modelo multifractal em  $n^2$ , onde  $n$  é o número de amostras para cálculo da distribuição.

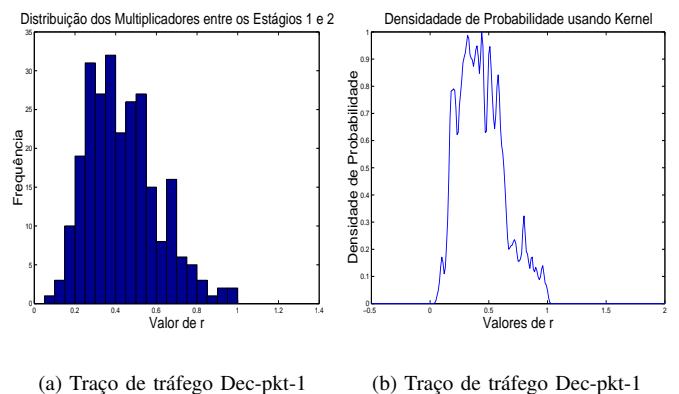


Fig. 2. Distribuição dos Multiplicadores entre os estágios 1 e 2

#### B. Método da Aceitação/ Rejeição

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

$$f(x) \leq cg(x) \quad (11)$$

onde  $c > 1$  é uma constante conhecida que satisfaz a equação (11) 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:

**Algoritmo de Aceitação/ Rejeição:**

Repetir

- 1) Gere amostras  $y$  de  $g(x)$
- 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:

*Teorema 1 (Método da Aceitação/Rejeição):* A variável aleatória  $X$  gerada pelo método de Aceitação/Rejeição possui densidade de probabilidade  $f(x)$ .

A prova deste teorema é trivial usando o fato de que:  $P(X \leq x) = P\{Y \leq x | U \leq f(y)/(cg(y))\}$  com  $u$  e  $y$  sendo independentes.

Quanto mais próxima a função de envoltória  $g(x)$  da função  $f(x)$ , melhor a taxa de aceitação. 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 [18]:

$$c = \sup_x f(x)/g(x) \quad (12)$$

A partir da soma de gaussianas obtida com o método de Kernel, podemos determinar a constante  $c$ :

$$c = \max \hat{f}(x) = 1/(\min(\sigma^2) \cdot n_g \cdot \sqrt{(2\pi)}) \quad (13)$$

onde  $n_g$  é o número de gaussianas usadas. 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 distribuição  $f(x)$ .

**C. Síntese do Processo CDGM**

Nesta seção esboçaremos o método de síntese do processo CDGM. Usamos o valor agregado obtido 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. Primeiramente, após a obtenção da densidade (equação (7)) pelo método de Kernel, usaremos a seguinte densidade para a geração dos multiplicadores em cada estágio  $i = 1, 2, \dots, N$ :

$$\hat{f}_i^*(x) = \frac{1}{n\delta_p} \sum_{p=1}^n K\left(\frac{x-z}{\delta_p}\right) \quad (14)$$

onde a média  $z$  e a variâncias  $\delta_p$ 's são dadas por:

$$z = E\{\hat{f}_i(x)\} \quad p = 1, 2, \dots, n. \quad (15)$$

$$\delta_p = \int_0^1 (x - X_p)^2 \cdot \frac{1}{h} K\left(\frac{x-X_p}{h}\right) dx \quad (16)$$

Estamos usando uma mistura de gaussianas para modelar a real densidade de probabilidade dos multiplicadores. A

equação (14) é usada para evitar um espalhamento muito intenso dos valores que podem ser produzidos pelo método de Aceitação/Rejeição com a variação de  $X_p$  no mesmo estágio.

O algoritmo de síntese do processo CDGM consiste basicamente de:

- 1) Calcula-se o valor agregado obtido na última escala;
- 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 (14) para gerar os multiplicadores em cada estágio  $i$ ;
- 5) Com os multiplicadores disponíveis, aplica-se o mesmo procedimento descrito anteriormente para geração de uma cascata binomial (seção II-A).

**IV. VALIDAÇÃO DO MODELO**

Com o objetivo de comparar as características estatísticas do modelo CDGM e do tráfego real, assim como de outros modelos 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 foram realizadas simulações analizando 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>1</sup>. Considerou-se amostras de tráfego em uma escala de agregação de 512ms para os traços de tráfego TCP/IP, escala na qual os traços apresentam forte característica multifractal [10]. Também usamos traços de tráfego Ethernet obtidos da Bellcore<sup>2</sup> que apresentam características auto-similares e multifractais.

A Tabela I mostra algumas estatísticas para a série Bellcore-Aug89. O modelo CDGM possui dentre as cascatas analisadas, parâmetro de Hurst que mais se aproxima do da série de tráfego real a custo de uma variância ligeiramente maior.

TABELA I

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	$1,3141^5$	$6,7068^9$	0,7871
VVGM	$1,6282 \cdot 10^5$	$9,2856 \cdot 10^8$	1,1449

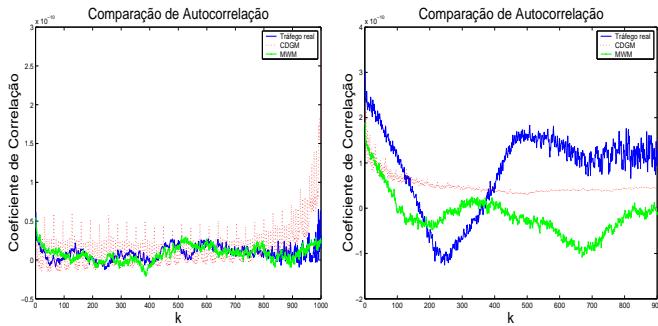
*A. Função de Autocorrelação*

A função de autocorrelação reflete as estatísticas de segunda ordem de uma série, dando uma idéia a respeito da longa-dependência nos dados. Para uma série  $y(t)$  com variância  $\sigma^2$ , o coeficiente de correlação é dado por:

$$\rho = \frac{E[y(t+k)y(t)]}{\sigma^2} \quad (17)$$

<sup>1</sup><http://ita.ee.lbl.gov/html/contrib/DEC-PKT.html>

<sup>2</sup><http://www.acm.org/sigcomm/ITA>



(a) Série Bc-Aug89

(b) Série Dec-pkt-2

Fig. 3. Coeficientes de correlação

A Figura 3 mostra os coeficientes de correlação para os traços de tráfego reais BellcoreAug89 e Dec-pkt-2, assim como para os modelos CDGM e MWM. Para maioria dos valores de  $k$  analisados, o modelo CDGM obteve correlação mais próxima da dos traços reais do que o importante modelo MWM.

#### B. Momentos de Ordem $q$

O tráfego agregado influencia o comportamento de fila nos multiplexadores ao invés de somente conexões individuais. Portanto, é importante investigar os momentos do tráfego agregado. Além do mais, agregação de fluxos ocorre em vários pontos da rede. Seja:

$$X_k^{(m)} = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i^N \quad (18)$$

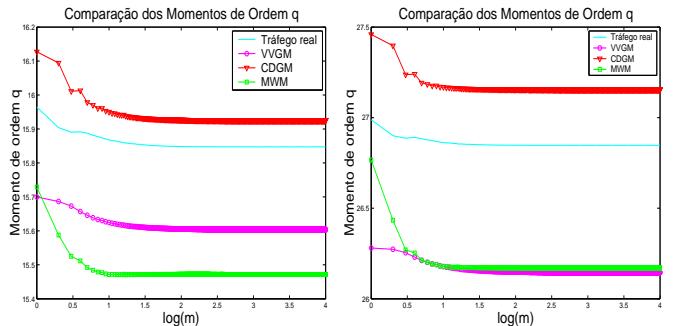
onde  $i = 1, 2, \dots, 2^N$ ,  $k = 1, 2, \dots, L$  e  $L = 2^N/m$  é o total de agregações para um valor fixo de  $m$ . Estimamos o momento de ordem  $q$  do tráfego agregado da seguinte forma:

$$\hat{\mu}^{(m)}(q) = \frac{1}{L} \sum_{k=1}^L |X_k^{(m)}|^q \quad (19)$$

A Figura 4 apresenta a variação de  $\log_{10}(\hat{\mu}^{(m)}(q))$  versus  $\log_{10}(m)$  para  $q = 2$  e  $q = 4$ . Em relação aos outros modelos, o CDGM é o que melhor aproxima os momentos, possuindo momentos até mais elevados do que dos traços reais de tráfego. Isso demonstra que é um modelo não otimista em relação a agregação de traços de tráfego. Cabe lembrar que os modelos monofractais não capturam os momentos de mais alta ordem dos tráfego de redes.

#### C. Espectro Multifractal

Em contraste a outros modelos de tráfego, processos multifractais contêm uma multiplicidade de expoentes de Hölder locais dentro de qualquer intervalo finito. 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



(a)  $q = 2$

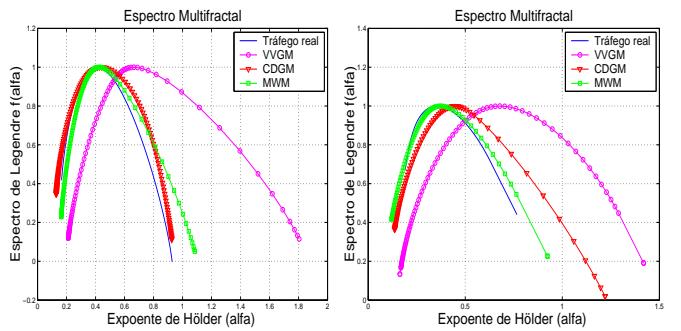
(b)  $q = 4$

Fig. 4. Momentos de Ordem  $q$  para a Série Bc-Aug89

de um processo, ou seja, caracteriza a sua suavidade (rajada) em um certo instante de tempo [19]. A distribuição destes expoentes pode ser representada por uma densidade normalizada chamada espectro multifractal. Em uma interpretação alternativa, o espectro multifractal descreve a dimensão fractal do conjunto de instantes possuindo um dado expoente local [11]. 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:

$$f(\alpha) = \min_q \{q\alpha - \tau(q)\} \quad (20)$$

Através deste método calculamos o espectro multifractal para o processo CDGM. Os espectros multifractais dos traços de tráfego dec-pkt-1 e dec-pkt-2 apresentaram  $\alpha < 1$ , o que indica uma alta incidência de rajadas multiescala. Pode-se observar que o modelo CDGM captura com eficiência o espectro multifractal e por consequência a multifractalidade dos traços de tráfego reais (Figura 5). O modelo VVGM, por sua vez, mostrou um espectro mais amplo porém não fiel ao do tráfego real.



(a) Traço de tráfego Dec-pkt-1

(b) Traço de tráfego Dec-pkt-2

Fig. 5. Espectro multifractal de Legendre

#### D. Testes de Desempenho e Comportamento de Fila

Foi considerado um sistema com um servidor com *buffer* finito sendo alimentado pelo processo CDGM a fim de com-

pararmos seu desempenho em modelar tráfego real. Analisou-se principalmente a ocupação média de bytes no *buffer* e a porcentagem de perda associadas a utilizações do *buffer*. Definimos utilização do *buffer*  $\lambda$  como sendo a razão entre o tempo total de serviço fornecido aos pacotes pelo tempo total de uso do *buffer*. Quanto maior o valor de  $\lambda$  indica que o *buffer* está constantemente sendo usado e consequentemente 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 ótimo em redes reais. A utilização do *buffer* depende da capacidade do servidor, ou seja, da taxa em que os dados são transmitidos.

A Figura 6.a apresenta os resultados para a perda de bytes em função da utilização do *buffer* para um *buffer* finito de tamanho 64Kbytes e a Figura 6.b mostra a ocupação média de bytes no *buffer* versus a utilização do mesmo. Pode-se notar uma porcentagem de perda maior para o modelo CDGM em relação ao MWM para alguns traços de tráfego, no entanto em termos de tamanho médio de fila a diferença é mais suave (Figura 6.b).

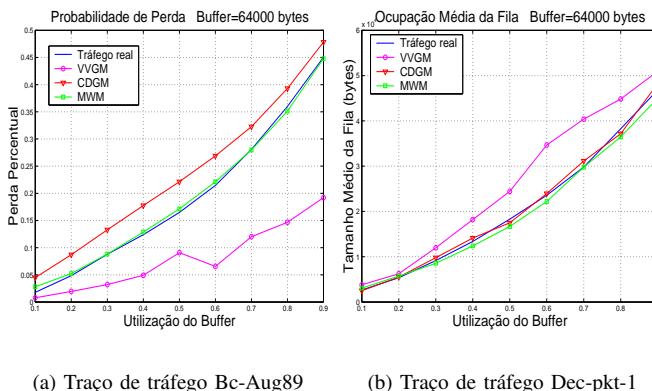
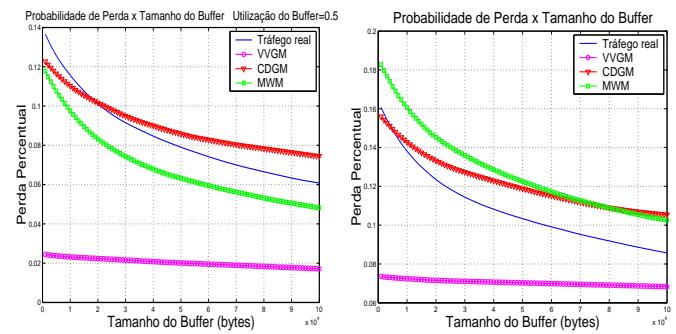


Fig. 6. Probabilidade de perda de bytes e ocupação média do buffer

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 e nota-se que como o CDGM possui média mais próxima da do traço dec-pkt-1 do que o MWM, apresenta melhores resultados.

## V. CONCLUSÕES

Pode-se concluir que o modelo multifractal proposto consegue capturar com eficiência as características do tráfego de redes. Nota-se também que ao se admitir uma densidade de probabilidade gaussiana para os multiplicadores pode-se levar a imprecisões na modelagem, como foi observado nas análises utilizando VVGM. Com o modelo CDGM propomos uma maior generalização desta distribuição com a média da mesma variando em cada estágio (mesmo que pequena a variação), diferente do que ocorre no VVGM. As simulações revelam que os resultados do modelo CDGM são comparáveis e em alguns casos superiores ao do modelo MWM e superiores ao do VVGM.



(a) Traço de tráfego Dec-pkt-1 (b) Traço de tráfego Dec-pkt-1

Fig. 7. Probabilidade de perda de bytes versus tamanho do buffer

## REFERÊNCIAS

- [1] F. H. P. Fitzek e M. Reisslein. "MPEG-4 and H.263 video traces for network performance evaluation". IEEE Trans. Network, v.15, n.6, p.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, v.2, n.1, p.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, p.977-984, Abril, 1995.
- [4] V. Paxson e S. Floyd. "Wide-area traffic: The failure of Poisson modeling". IEEE/ACM Transactions on Networking, v.3, n.3, p.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., v.4, n.2, p.209-223, 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, v.7, n.5, p.629-640, Outubro, 1999.
- [8] Q. Li e D. Mills. "Investigating the Scaling Behavior, Crossover and Anti-persistence of Internet Packet Delay Dynamics", Globecom'99, p.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., p.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, p.352-359, Março, 2000.
- [11] 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 v.45, n.3, p.992-1018, Abril, 1999.
- [12] J. Gao and I. Rubin. "Multiplicative multifractal modelling of long-range-dependent network traffic". Int. J. Commun. Syst., v.14, p.783-801, 2001.
- [13] Murali Krishna et al. "Global scaling exponent for variable variance gaussian multiplicative (VVGM) multifractal cascades". Proc. SPCOM'01, p.19-25, 2001.
- [14] B. B. Mandelbrot, A. Fisher, e L. Calvet. "A Multifractal Model of Asset Return". Yale University, 1997, <http://ideas.repec.org/p/cwl/cwldpp/1164.html>.
- [15] C. K. Chui. "An Introduction to Wavelets". San Diego: Academic, 1992.
- [16] 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, vol. 2, p. 1026-1035, 2000.
- [17] B. Silverman. "Density Estimation for Statistics and Data Analysis". Chapman and Hall, London, 1986.
- [18] K. Monahan. "Numerical Methods of Statistics". Cambridge University Press, NY.
- [19] Z. Zhang, V. Ribeiro, S. Moon, and C. Diot. "Small-Time Scaling behaviors of internet backbone traffic: An Empirical Study," IEEE Infocom, San Francisco, Março, 2003.