

Método de Estimação de Parâmetros para Modelagem no Domínio *Wavelet* do Tráfego de Redes de Computadores Usando o Algoritmo de Levenberg-Marquardt

Maykon Renan P. da Silva¹, Flávio Geraldo C. Rocha²

^{1,2}Escola de Engenharia Elétrica, Mecânica e de Computação
Universidade Federal de Goiás (UFG)

maykonrenan@discente.ufg.br, flaviogcr@ufg.br

Abstract. *In this work proposes a multifractal model based on a multiplicative cascade in the wavelet domain, to synthesize network traffic samples. For this purpose, in the proposed model, a parametric modeling based on an exponential function is used for the variance of the multipliers along the stages of the cascade. The exponential function parameters are obtained through the solution of a non-linear system, for this purpose, the Levenberg-Marquardt method is used. The main contribution of the proposed algorithm is to use a fixed and reduced number of parameters to generate network traffic samples that have characteristics such as self-similarity and wide Multifractal Spectrum Width (MSW) similar to the real network traffic traces.*

Resumo. *Este trabalho propõe um modelo multifractal baseado em uma cascata multiplicativa no domínio wavelet, para sintetizar amostras de tráfego de redes. No modelo proposto, utiliza-se uma modelagem paramétrica baseada em uma função exponencial para a variância dos multiplicadores ao longo dos estágios da cascata. Os parâmetros da função exponencial são obtidos através da solução de um sistema não linear, para este fim, utiliza-se o método de Levenberg-Marquardt. A principal contribuição do algoritmo proposto é utilizar um número fixo e reduzido de parâmetros para gerar amostras de tráfego de redes que apresentem características como autossimilaridade e ampla Largura de Espectro Multifractal (MSW) similares as séries reais de tráfego de redes.*

1. Introdução

Os fluxos de redes de computadores podem englobar diferentes protocolos e conteúdos provenientes de diversas tecnologias de acesso ao meio de comunicação. A heterogeneidade dessas redes proporciona um comportamento de tráfego variável, com presença de rajadas difíceis de serem previstas. A melhoria no desempenho das redes de comunicações tem sido alvo de pesquisas constantes, e muitos são os trabalhos propondo modelos para descrição do tráfego de redes [Riedi et al. 1999, Krishna et al. 2003, Millán and Lefranc 2015, Tuberquia-David et al. 2016, Grandemange et al. 2017].

Dentre os modelos existentes para a caracterização do tráfego de redes de comunicações, modelos autossimilares e multifractais têm recebido grande atenção devido aos seus desempenhos em termos de análise e modelagem relacionados ao tráfego real de redes. Os modelos multifractais baseados em cascatas multiplicativas podem ser mais precisos e abrangentes quando aplicados ao controle do tráfego [Feldmann et al.

1998], [Vieira and Lee 2010], já que em situações práticas o grau de autossimilaridade do tráfego é variante com o tempo, o que não ocorre em modelos Gaussianos e monofractais, reduzindo o horizonte de aplicação destes modelos. Estas propriedades de séries temporais, que vão além daquelas descritas pelos fractais, influenciaram o surgimento de propostas baseadas em multifractais. Desta forma, com uma caracterização do tráfego mais detalhada utilizando análise multifractal, espera-se obter melhores estimativas para a modelagem do tráfego de redes de comunicações.

No entanto, os modelos multifractais podem ser complexos e suas aplicações práticas podem ser limitadas pelo número de seus parâmetros, que geralmente é alto e dependente do número de amostras de tráfego de redes que se deseja modelar.

Neste trabalho propõe-se um modelo multifractal baseado em um número reduzido (quando comparado a outros modelos presentes na literatura) e fixo (não dependente do número de amostras de tráfego a ser modelado) de parâmetros.

O restante do artigo está organizado da seguinte forma: a Seção 2 apresenta uma breve descrição da análise multifractal aplicada ao tráfego de redes; a Seção 3 apresenta alguns dos modelos multifractais presentes na literatura; na Seção 4 é proposto o modelo multifractal proposto baseado na caracterização da variância dos multiplicadores de uma cascata multiplicativa no domínio *wavelet*; na Seção 5 é avaliada a eficiência do modelo proposto em comparação a outros modelos presentes na literatura: β MWM (*Multifractal Wavelet Model*) [Riedi et al. 1999], VVGM (*Variable Variance Gaussian Model*) [Krishna et al. 2003] e MFHSW (*MultiFractal Hurst Spectrum Width*) [Tuberquia-David et al. 2016], utilizando dados reais; na Seção 6 são apresentadas as conclusões obtidas.

2. Análise Multifractal do Tráfego de Redes

Fractais são objetos geométricos que possuem uma forma irregular que persiste em várias resoluções de observação [Mandelbrot 1990], como mostra a Figura 1. A maioria dos sinais fractais conhecidos possuem características como a autossimilaridade. O conceito de autossimilaridade está associado aos objetos cuja forma não muda em decorrência da mudança de escala, característica esta que pode influir, por exemplo, no desempenho do tráfego de redes [Vieira and Lee 2010].

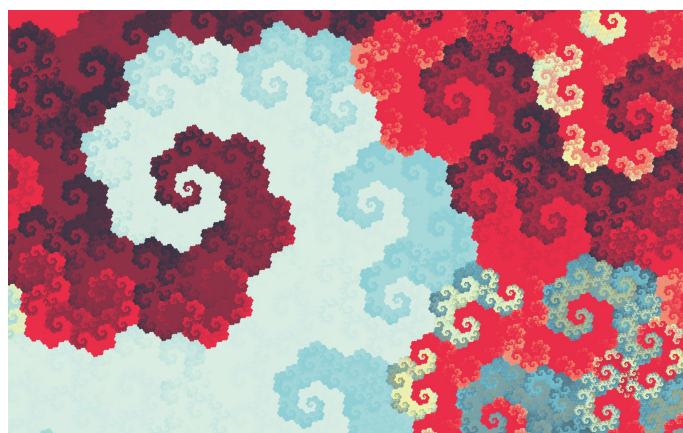


Figura 1. Fractal baseado no conjunto de Mandelbrot [CODERS 2020].

O parâmetro de *Hurst* (H) mede o grau de autossimilaridade do processo com ca-

racterística fractal, onde H está contido no intervalo $(0 < H < 1, 0)$. Convém ressaltar que no caso de processos assintoticamente autossimilares restringindo-se o parâmetro de *Hurst* ao intervalo $(0, 5 < H < 1, 0)$, autossimilaridade implica em longa dependência. Assim, diz-se que processos autossimilares com $H > 0, 5$ apresentam dependência de longa duração; isto significa que a função de autocorrelação (ACF – *Autocorrelation Function*) do processo apresenta decaimento hiperbólico [Abry and Veitch 1998].

A autossimilaridade introduz dificuldades na otimização do desempenho da rede e na garantia de qualidade de serviço (QoS), podendo causar impactos significativos como aumento da latência e da taxa de perda de pacotes. Um conceito relacionado à autossimilaridade é o de distribuição de cauda pesada (*heavy tail*) onde a densidade de probabilidade dos processos de tráfego decai lentamente [Park and Willinger 2000]. Quanto mais o tráfego de uma rede apresenta características de uma distribuição de cauda pesada, mais os valores de parâmetros de QoS dessa rede são degradados, [Rocha 2011].

Assim como os fractais, os multifractais são utilizados em muitas áreas como análise de fenômenos naturais como terremotos [Barman et al. 2015], processamento e decomposição de sinais [Mallat 2008] e modelagem de tráfego de redes de computadores [Riedi et al. 1999], [Krishna et al. 2003], [Tuberquia-David et al. 2016]. Na análise multifractal verifica-se o comportamento em escala de momentos estatísticos dos processos para estimar suas regularidades locais [Feldmann et al. 1998], [Riedi et al. 1999]. Através de ferramentas da análise multifractal algumas propriedades encontradas em processos reais podem ser verificadas. O tráfego de redes, por exemplo, ao ser considerado multifractal significa que possui uma estrutura de forte dependência inerente entre as amostras, com incidência de rajadas em várias escalas [Riedi et al. 1999]. Processos multifractais são definidos a seguir, [Vieira et al. 2012].

Definição 2.1. Um processo estocástico $X(t)$ é multifractal se satisfaz a equação:

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

onde $t \in T$ e $q \in Q$ são intervalos na reta real, $\tau(q)$ e $c(q)$ são funções no domínio Q . Normalmente, assume-se T e Q tenham comprimentos positivos, e que $0 \in T$, $[0, 1] \subseteq Q$.

A **Definição 2.1** permite que “multifractalidade” seja descrita em termos de momentos, onde $\tau(q)$ é a função de escala e $c(q)$ é o fator de momento de um processo multifractal. Caso $\tau(q)$ seja linear em q o processo é dito ser monofractal, caso contrário, é multifractal [Riedi et al. 1999], [Park and Willinger 2000], [Vieira et al. 2012].

3. Trabalhos Relacionados

O modelo multifractal MWM é baseado em uma cascata multiplicativa no domínio *wavelet* [Riedi et al. 1999], tendo sido recentemente utilizado para a modelagem do tráfego de redes [Tuberquia-David et al. 2016]. No MWM, a variável $U_{j,k}$ representa os coeficientes de escala; $W_{j,k}$ representa os coeficientes *wavelets* e $A_{j,k}$ representa os multiplicadores da cascata multiplicativa; o índice j representa o estágio da cascata (também chamado de escala) e k é o instante de tempo dentro da escala j .

Para se gerar um processo segundo o modelo MWM é preciso aplicar a transformada *wavelet* discreta (DWT – *Discrete Wavelet Transform*) ao tráfego de redes, calcular os momentos de segunda ordem dos coeficientes *wavelet* ($W_{j,k}$) em cada escala da cascata, a média e a variância dos coeficientes de escala ($U_{j,k}$) na escala de menor resolução e

calcular ρ_j , variável usada para capturar o decaimento de energia dos coeficientes *wavelet* ($W_{j,k}$) em função da escala j . A Figura 2 ilustra como é construída a cascata multiplicativa no domínio *wavelet* a partir dos parâmetros descritos anteriormente.

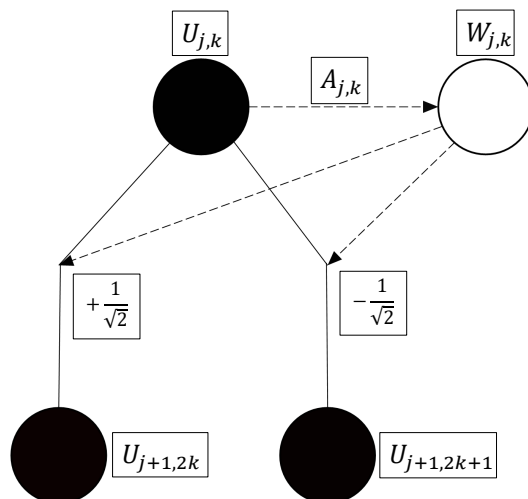


Figura 2. Cascata multiplicativa formada pelos coeficientes de escala $U_{j,k}$, coeficientes *wavelet* $W_{j,k}$ e pelos multiplicadores $A_{j,k}$.

Quando os multiplicadores $A_{j,k}$ podem ser descritos por meio de uma distribuição $\beta(\cdot, \cdot)$, o modelo MWM é chamado de β MWM.

No β MWM o número de parâmetros para gerar uma série sintética de tráfego de redes depende do tamanho da série real de tráfego, ou seja quanto maior o número de amostras da série real de tráfego, maior o número de escalas n necessárias para se construir a cascata multiplicativa e consequentemente maior o comprimento do vetor ρ_j ($j = 0, 1, 2, \dots, n-1$). Assim, o número de parâmetros necessários para gerar um processo de tráfego no β MWM é de $(n + 2)$, onde n é o número de termos do vetor ρ_j e os parâmetros μ_c e σ_c correspondem a média e ao desvio padrão do coeficiente de escala $U_{0,0}$, respectivamente, gerado segundo uma distribuição Gaussiana.

O modelo VVGM [Krishna et al. 2003] baseia-se em uma cascata multiplicativa onde os valores correspondentes aos multiplicadores da cascata são amostras de uma distribuição Gaussiana a cada estágio da cascata, ou seja, nesse modelo assume-se que a distribuição dos multiplicadores é Gaussiana com média (μ) igual a 0.5 e variância (σ_j) que varia a cada escala da cascata, a medida que ela é sintetizada a partir do parâmetro x^N que representa o valor agregado das N amostras de tráfego.

O VVGM atribui um valor fixo de média para os multiplicadores de todas as escalas o que pode comprometer o processo de modelagem, visto que séries reais de tráfego de redes apresentam multiplicadores com média variável a cada escala. Assim como o β MWM, o VVGM requer $(n + 2)$, parâmetros para a síntese de amostras de tráfego de redes. Maiores detalhes sobre o VVGM podem ser obtidos em [Krishna et al. 2003].

O modelo MFHSW [Tuberquia-David et al. 2016] permite ao projetista de redes a configuração a priori dos parâmetros de *Hurst* (H) e de largura do espectro multifractal ($MSW - Multifractal Spectrum Width$). Assim como o β MWM, o MFHSW baseia-se

em uma cascata multiplicativa no domínio *wavelet* e o número de parâmetros necessários para gerar uma série sintética de tráfego está diretamente ligado ao número de amostras da série real de tráfego, além de ser necessário estimar os parâmetros μ_c , σ_c e ρ_j , o MFHSW precisa que o projetista ajuste os parâmetros H e MSW , o que torna o MFHSW um algoritmo que precisa de $(n + 5)$ parâmetros para gerar um processo de tráfego sintético. Não obstante, caso a série sintética obtida não atenda aos parâmetros H e MSW predeterminados, essa série é descartada e o algoritmo é reiniciado, o que melhora a precisão do modelo, mas aumenta consideravelmente o tempo de execução do algoritmo.

4. Modelo Multifractal Proposto

Nesta seção, apresenta-se a **Proposição 1**, e sua respectiva demonstração, que fundamenta o modelo proposto. A **Proposição 1** baseia-se em uma estrutura de cascata multiplicativa no domínio *wavelet* onde a variância dos multiplicadores ao longo das escalas da cascata é modelada por meio de uma função exponencial. Para tanto, na nova abordagem proposta, os valores de ρ_j , que estão relacionados com o decaimento de energia dos coeficientes *wavelets* do modelo β MWM, são calculados de forma paramétrica ao se modelar os momentos de segunda ordem dos multiplicadores da cascata ($\mathbb{E}[A_{(j)}^2]$) a partir do parâmetro de *Hurst* (H). Em seguida apresenta-se dois algoritmos que constituem o modelo proposto: o **Algoritmo 1 (Estimação de Parâmetros)** responsável pela modelagem dos parâmetros do modelo proposto com base na estrutura da cascata multiplicativa no domínio *wavelet*, como mostra a Figura 2; o **Algoritmo 2 (Síntese de Tráfego)** responsável por gerar amostras sintéticas de tráfego de redes com base nos parâmetros estimados pelo **Algoritmo 1**.

Proposição 1. Uma série temporal de tráfego real de redes $C(t)$ pode ser descrita por um processo sintético $C^{(n)}[k]$ gerado por meio do modelo proposto que é baseado em três parâmetros: média (μ_c) e desvio padrão (σ_c) do coeficiente de escala $U_{0,0}$ e parâmetro de *Hurst* (H). A taxa de distribuição (α), parâmetro de forma (β) e o parâmetro de localização (γ) de uma função exponencial utilizada para modelar a variância dos multiplicadores $A_{(j)}$ são obtidos a partir do valor de H . O processo sintético é definido pela seguinte equação:

$$C^{(n)}[k] \stackrel{d}{=} 2^{-n} G(\mu_c, \sigma_c) \left[1 + \beta \left(\frac{2 - 4^H}{4^H - 4}, \frac{2 - 4^H}{4^H - 4} \right) \right] \times \prod_{j=0}^{n-2} \left[1 + \beta \left(\frac{e^{\left(\frac{j+1-\beta}{\gamma}\right)^2} - \alpha}{2\alpha}, \frac{e^{\left(\frac{j+1-\beta}{\gamma}\right)^2} - \alpha}{2\alpha} \right) \right], \quad (2)$$

onde $k = 0, 1, \dots, 2^j - 1$ é o deslocamento de tempo para $j = 0, 1, \dots, n - 1$ e n é a variável relacionada à escala de tempo mais fina (de maior resolução) de uma cascata multiplicativa no domínio *wavelet*.

Demonstração. O processo para se gerar uma série sintética de tráfego de redes pode ser definido pela seguinte equação [Riedi et al. 1999]:

$$C^{(n)}[k] \stackrel{d}{=} 2^{-n} U_{0,0} \prod_{j=0}^{n-1} [1 + A_{(j)}], \quad (3)$$

onde $U_{0,0} \stackrel{d}{=} G(\mu_c, \sigma_c)$ e $A_{(j)}$ representa $A_{j,k}$, pois os multiplicadores podem ser descritos por uma distribuição β simétrica, onde $A_{(j)} \stackrel{d}{=} \beta(\rho_j, \rho_j)$. Os valores de $\beta(\rho_j, \rho_j)$ estão contidos no intervalo $[-1, 1]$.

A partir do parâmetro de *Hurst* (H) calcula-se o parâmetro ρ_j para $j = n - 1$ (última escala), da seguinte forma [Riedi et al. 1999]:

$$\rho_{n-1} = \frac{2^{2H-1} - 1}{2 - 2^{2H-1}}, \quad (4)$$

onde,

$$\lim_{H \rightarrow 0,5} \rho_{n-1} = 0 \text{ e } \lim_{H \rightarrow 1,0} \rho_{n-1} = \infty,$$

logo, $\forall H \in \mathbb{R}, \exists \epsilon > 0 \mid \lim_{\epsilon \rightarrow 0} H \in [0, 5 + \epsilon; 1, 0 - \epsilon]$.

Pode-se calcular a variância de $A_{(j)}$ em função de ρ_j pela seguinte relação [Riedi et al. 1999]:

$$\begin{aligned} \mathbb{E}[A_{(j)}^2] - \mathbb{E}[A_{(j)}]^2 &= \frac{1}{2\rho_j + 1}, \\ \mathbb{E}[A_{(j)}^2] &= \frac{1}{2\rho_j + 1}, \end{aligned} \quad (5)$$

onde $\mathbb{E}[A_{(j)}] = 0$. Substituindo (4) em (5) pode-se calcular $\mathbb{E}[A_{(j)}^2]$ para a última escala, pela seguinte equação:

$$\mathbb{E}[A_{(n-1)}^2] = \frac{1}{2 \frac{2^{2H-1}-1}{2-2^{2H-1}} + 1} = 2^{2-2H} - 1. \quad (6)$$

Assumindo que para $j = 0$ a cascata multiplicativa na primeira escala é formada por apenas um multiplicador $A_{(0)}$, a variância de apenas um termo é teoricamente zero. Na abordagem proposta assume-se que a variância é infinitesimal e definida por:

$$\mathbb{E}[A_{(0)}^2] = \epsilon. \quad (7)$$

Na prática, o valor de ϵ representa o grau de variabilidade entre os valores agregados ($U_{0,0}$) de tempos entre chegadas de amostras de tráfego de redes. Assumindo que a variância dos multiplicadores pode ser modelada por uma função exponencial de três parâmetros (α , β e γ), obtém-se:

$$\mathbb{E}[A_{(j)}^2] = \alpha e^{-\left(\frac{j+1-\beta}{\gamma}\right)^2}, \text{ para } j = 0, 1, \dots, n - 1. \quad (8)$$

Substituindo (6) e (7) em (8), obtém-se:

$$\begin{cases} \alpha e^{-\left(\frac{n-\beta}{\gamma}\right)^2} = 2^{2-2H} - 1, \\ \alpha e^{-\left(\frac{1-\beta}{\gamma}\right)^2} = \epsilon. \end{cases} \quad (9)$$

Tendo H como parâmetro de entrada, soluciona-se (9) para obter α , β e γ . Entretanto, como (9) não pode ser solucionado diretamente, utiliza-se o método de Levenberg-Marquardt (LM) com o objetivo de encontrar os melhores valores para α , β e γ na parametrização da variância dos multiplicadores.

O método de Levenberg-Marquardt, ao utilizar o critério dos mínimos quadrados não lineares, oferece um método eficiente para essa parametrização.

Considerando o problema dos mínimos quadrados, dado pela seguinte relação:

$$\min_{\mathbf{p}} |\mathbf{J}_k \mathbf{p}_k + \mathbf{f}_k|, \quad (10)$$

o método de LM para minimização de resíduos baseia-se no seguinte sistema:

$$(\mathbf{J}_k^T \mathbf{J}_k + \lambda_k \mathbf{I}) \mathbf{p}_k = -\mathbf{J}_k^T \mathbf{f}_k, \quad (11)$$

onde \mathbf{f}_k é o vetor de funções não lineares, \mathbf{J}_k é a matriz Jacobiana de \mathbf{f}_k , $\mathbf{I}_{n \times n}$ é a matriz identidade e λ_k é um número real não-negativo denominado parâmetro de Levenberg-Marquardt [Cunha 2000].

Para ajustar os valores $(j, \mathbb{E}[A_{(j)}^2])$ utiliza-se uma função do tipo:

$$y(j; a_1, a_2, a_3) = a_1 e^{-\left(\frac{j+1-a_2}{a_3}\right)^2}. \quad (12)$$

onde a_1 , a_2 e a_3 correspondem aos parâmetros α , β e γ , respectivamente.

Nestes cálculos, toma-se a aproximação inicial $\mathbf{a} = (1, 1, 1)$. Em cada iteração chama-se $\mathbf{a} = \mathbf{a}^k$ e calcula-se:

1. Vetor resíduo;

$$f_j(\mathbf{a}) = \mathbb{E}[A_{(j)}^2] - a_1 e^{-\left(\frac{j+1-a_2}{a_3}\right)^2}, \quad (13)$$

2. Matriz jacobiana. As duas colunas dessa matriz são;

$$J_{j,1} = \frac{\partial f_j}{\partial a_1}(\mathbf{a}) = e^{-\left(\frac{j+1-a_2}{a_3}\right)^2}, j = 0 : n - 1, \quad (14)$$

$$J_{j,2} = \frac{\partial f_j}{\partial a_2}(\mathbf{a}) = \frac{2a_1(j+1-a_2)e^{-\left(\frac{a_2-(j+1)}{a_3}\right)^2}}{a_3^2}, j = 0 : n - 1, \quad (15)$$

$$J_{j,3} = \frac{\partial f_j}{\partial a_3}(\mathbf{a}) = \frac{2a_1(j+1-a_2)^2 e^{-\left(\frac{j+1-a_2}{a_3}\right)^2}}{a_3^3}, j = 0 : n - 1. \quad (16)$$

Com os parâmetros α , β e γ já estimados por meio de (11) pode-se calcular ρ_j para as escalas $j = 0, 1, \dots, n - 1$. Isolando ρ_j em (5), obtém-se:

$$\rho_j = \frac{1}{2} \left(\frac{1}{\mathbb{E}[A_{(j)}^2]} - 1 \right). \quad (17)$$

Substituindo (6) em (17), tem-se ρ_j para a última escala:

$$\rho_{n-1} = \frac{1}{2} \left(\frac{1}{2^{2-2H} - 1} - 1 \right) = \frac{2 - 4^H}{4^H - 4}. \quad (18)$$

Substituindo (8) em (17), tem-se ρ_j para as outras escalas da cascata:

$$\rho_j = \frac{1}{2} \left(\frac{1}{\alpha e^{-\left(\frac{j+1-\beta}{\gamma}\right)^2}} - 1 \right) = \frac{e^{\left(\frac{j+1-\beta}{\gamma}\right)^2} - \alpha}{2\alpha}. \quad (19)$$

Por fim, substituindo (18) e (19) em (3), obtém-se (2). □

No modelo proposto os valores de ρ_j para todas as escalas são calculados de forma paramétrica. Por isso, para gerar uma série sintética de tráfego de redes por meio do proposto são necessários apenas três parâmetros (μ_c , σ_c e H), o que faz do modelo proposto um algoritmo cujo o número de parâmetros não depende da quantidade de amostras de tráfego. O modelo proposto é dividido em dois algoritmos apresentados a seguir:

Algoritmo 1 (Estimação de Parâmetros): Estimação dos parâmetros do modelo proposto a partir de uma série real de tráfego de redes $C(t)$.

Entrada: $C(t)$; Saída: μ_c , σ_c , H .

1. Determinar $U_{j,k}$ e $W_{j,k}$ por meio da DWT (*Discrete Wavelet Transform*), dado que $W_{j,k} = \langle C(t), \psi_{j,k} \rangle$, onde $\psi_{j,k}$ é a *wavelet* de Haar;
2. Utilizar a estrutura da cascata multiplicativa obtida no passo 1 para determinar μ_c e σ_c , média e desvio padrão de $U_{0,0}$, respectivamente;
3. Utilizar $W_{j,k}$ para estimar H por meio do diagrama log-escala (*LD - Logscale Diagram*) proposto por [Abry and Veitch 1998].

Algoritmo 2 (Síntese de Tráfego): Gerar amostras sintéticas de tráfego de redes a partir dos três parâmetros obtidos na saída do **Algoritmo 1**.

Entrada: μ_c , σ_c , H ; Saída: $C^{(n)}[k]$.

1. Tendo H como parâmetro de entrada, determinar α , β e γ por meio de (11) (método de Levenberg-Marquardt);
2. Para $j = 0$, gerar $U_{0,0}$ de acordo com uma distribuição Gaussiana com parâmetros μ_c e σ_c , i.e., $G(\mu_c, \sigma_c)$;
3. Utilizar α e γ para calcular ρ_j por meio de (18) e (19) para cada escala j , e gerar k valores aleatórios de acordo com a distribuição $\beta(\rho_j, \rho_j)$ para obter $A_{(j)}$;
4. Utilizar (2) para gerar uma série de amostras sintéticas ($C^{(n)}[k]$) de tráfego de redes.

O modelo proposto diferencia-se dos modelos β MWM, VVGM e MFHSW, por ser eficiente e não utilizar uma quantidade de parâmetros dependente do número de amostras a serem modeladas. O modelo proposto utiliza um número fixo e reduzido de parâmetros, sendo eles: a média (μ_c) e desvio padrão (σ_c) do tráfego agregado e o parâmetro de *Hurst* (H). Por meio desses parâmetros, é possível gerar tráfego sintético utilizando o **Algoritmo 2**.

5. Resultados

Foram conduzidas simulações com o objetivo de avaliar o desempenho do modelo proposto em descrever com acurácia as características das séries reais de tráfego de redes.

São apresentados resultados obtidos utilizando as séries reais *BCpAug89* [Danzig et al. 2000] e *dec-pkt-2* [Danzig et al. 2000] que, apesar de não serem atuais, são clássicas na literatura. Além disso, são apresentados os resultados utilizando a série de tráfego *MultipathTCP* [Coninck et al. 2016]. Essas séries de tráfego de redes apresentam amostras de tempo entre chegadas de pacotes dado em segundos. Resultados similares foram obtidos para outras séries reais de tráfego.

A Figura 3 apresenta a variância dos multiplicadores para as séries reais de tráfego *BCpAug89*, *dec-pkt-2* e *MultipathTCP* e para as séries sintéticas geradas por meio dos modelos β MWM, VVGM, MFHSW e proposto.

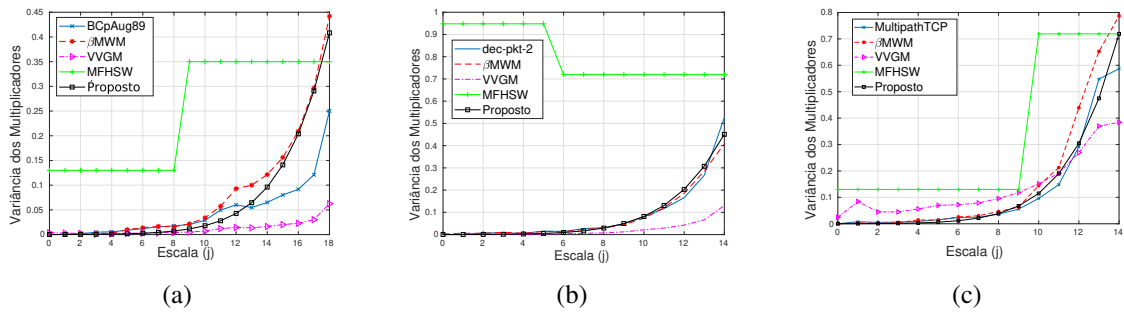


Figura 3. Variância dos multiplicadores ($\mathbb{E}[A_{(j)}^2]$).

A respeito dos resultados para a variância dos multiplicadores, para as séries de tráfego analisadas, o modelo proposto, que utiliza menos parâmetros que os demais modelos, obteve desempenho similar ou superior aos demais, em termos de representação da curva da variância real dos multiplicadores. A Tabela 1 corrobora essa afirmação.

Tabela 1. EQM das curvas da variância dos multiplicadores de cada modelo.

Modelos Multifractais	<i>BCpAug89</i>	<i>dec-pkt-2</i>	<i>MultipathTCP</i>
β MWM	0,0050	0,0021	0,0399
VVGM	0,0155	0,0247	0,0538
MFHSW	0,0346	0,5437	0,0797
Proposto	0,0029	0,0014	0,0016

A Tabela 2 apresenta estatísticas e o erro relativo percentual ($E(\%)$) para as séries reais e sintéticas geradas a partir dos modelos considerados. Os resultados da Tabela 2 foram obtidos com uma média de 50 simulações. Analisando os resultados é possível observar que o modelo proposto apresentou o menor $E(\%)$ para o parâmetro de *Hurst*, além disso apresentou erro apenas ligeiramente maior do que o β MWM para a variância. O VVGM foi o único modelo que apresentou resultados destoantes para a média das amostras das séries reais de tráfego. O modelo MFHSW obteve o pior desempenho em relação à variância das amostras de tráfego de redes, apresentando um alto $E(\%)$, decorrente de sua ineficiência em estimar a variância dos multiplicadores para diferentes escalas, como mostrou a Figura 3 para diferentes séries reais de tráfego.

Para processos multifractais, o espectro geralmente apresenta uma forma parabólica côncava, cuja a largura (*MSW*) é proporcional a variabilidade dos expoentes

Tabela 2. Estatísticas para média, variância e H

<i>BCpAug89</i>	Média (10^{-3})	Variância (10^{-5})	H
Real	3,00	2,90	0,7156
β MWM E(%)	3,00 (0,0)	2,90 (0,0)	0,7275 (1,7)
VVGM E(%)	3,30 (10,0)	1,48 (48,9)	0,7981 (11,5)
MFHSW E(%)	3,00 (0,0)	58,22 (1907,5)	0,7097 (0,8)
Proposto E(%)	3,00 (0,0)	3,12 (7,5)	0,7172 (0,2)
<i>dec-pkt-2</i>	Média (10^{-3})	Variância (10^{-6})	H
Real	2,00	8,43	0,7195
β MWM E(%)	2,00 (0,0)	8,44 (0,1)	0,6963 (3,2)
VVGM E(%)	2,10 (5,0)	9,40 (11,5)	0,7432 (3,3)
MFHSW E(%)	2,00 (0,0)	55,18 (554,5)	0,7271 (1,05)
Proposto E(%)	2,00 (0,0)	8,91 (5,7)	0,7124 (0,9)
<i>MultipathTCP</i>	Média (10^{-2})	Variância (10^{-4})	H
Real	1,11	7,78	0,6092
β MWM E(%)	1,11 (0,0)	7,82 (0,5)	0,6218 (2,1)
VVGM E(%)	1,08 (2,7)	4,99 (35,8)	0,7084 (16,3)
MFHSW E(%)	1,11 (0,0)	54,00 (594,1)	0,6209 (1,9)
Proposto E(%)	1,11 (0,0)	7,79 (0,1)	0,6102 (0,8)

de Hölder [Riedi et al. 1999], [Tuberquia-David et al. 2016]. A Figura 4 apresenta os espectros multifractais para as séries reais de tráfego de redes *BCpAug89*, *dec-pkt-2* e *MultipathTCP*. O modelo proposto apresentou um espectro largo e com uma concavidade similar ao apresentado pelas séries reais. Essa afirmação é corroborada pela Tabela 3, onde o modelo proposto apresentou o menor E(%) para a MSW para as três séries utilizadas. Embora o MFHSW tenha apresentado uma MSW próxima à largura das séries reais, observa-se por meio da Figura 4 que a curva do espectro do modelo MFHSW utilizando a série *dec-pkt-2* está deslocada no eixo dos expoentes de Hölder em relação às curvas obtidas para a série real e para o modelo proposto.

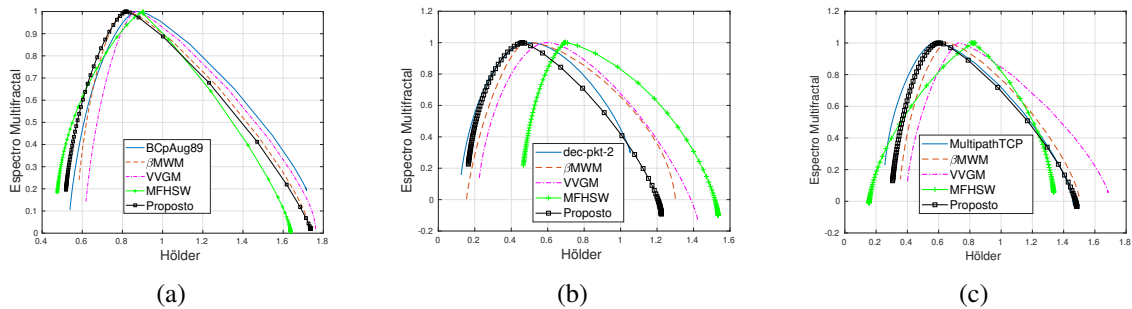


Figura 4. Espectro multifractal.

A Tabela 4 apresenta os parâmetros necessários para gerar amostras sintéticas de tráfego de redes de acordo com os modelos proposto, β MWM, VVGM e MFHSW. Os modelos β MWM, VVGM e MFHSW apresentam um vetor de parâmetros, onde o tamanho do mesmo depende do número de escalas j (que está diretamente relacionada com

Tabela 3. Largura do Espectro Multifractal (MSW).

Modelos Multifractais	BCpAug89	dec-pkt-2	MultipathTCP
Real	1, 1780	0, 9278	1, 2057
β MWM E(%)	1, 1486 (2, 5)	1, 1464 (23, 6)	1, 1420 (5, 3)
VVGM E(%)	1, 1434 (2, 9)	1, 1980 (29, 1)	1, 2845 (6, 5)
MFHSW E(%)	1, 1665 (1, 0)	1, 0685 (15, 2)	1, 1915 (1, 2)
Proposto E(%)	1, 1852 (0, 6)	1, 0588 (14, 1)	1, 1953 (0, 8)

a quantidade de amostras de tráfego de rede a serem modeladas), ou seja, quanto maior o número de escalas maior o vetor e consequentemente maior o número de parâmetros. O modelo proposto diferente dos demais modelos comparados apresenta um número fixo e reduzido de parâmetros, que independe do número de escalas, ou seja, o número de parâmetros independe da quantidade de amostras do tráfego a ser modelado.

Tabela 4. Parâmetros de entrada para a síntese de tráfego de redes de cada Modelo Multifractal.

Modelos	Parâmetros
β MWM	μ_c, σ_c, ρ_j
VVGM	x^N, μ e σ_j
MFHSW	$\mu_c, \sigma_c, \rho_j, H, MSW$
Proposto	μ_c, σ_c, H

6. Conclusões

Neste trabalho, foi proposto um modelo para caracterização do tráfego de redes com características multifractais. O modelo utiliza uma função exponencial para obter de forma paramétrica a variância dos multiplicadores ao longo das escalas de uma cascata multiplicativa no domínio *wavelet*. Os parâmetros da função exponencial são obtidos através da solução de um sistema não linear, para este fim, utiliza-se o método de Levenberg-Marquardt. O modelo proposto foi dividido em dois algoritmos, o **Algoritmo 1 (Estimação de Parâmetros)** é responsável pela obtenção dos parâmetros do modelo e o **Algoritmo 2 (Síntese de Tráfego)** é responsável por gerar amostras de tráfego sintético de rede a partir dos parâmetros obtidos no **Algoritmo 1**.

Os resultados obtidos por meio de simulações utilizando séries reais de tráfego de redes mostraram que o modelo proposto é eficiente em descrever estatísticas de primeira e segunda ordem (média e variância, respectivamente), parâmetro de *Hurst* e também a largura do espectro multifractal (*MSW*). O modelo proposto foi comparado aos modelos β MWM, VVGM e MFHSW. Ao contrário dos demais modelos comparados, independente do número de amostras de tráfego que se deseja modelar, o modelo proposto utiliza um número fixo e reduzido de parâmetros (μ_c, σ_c e H). Essas características tornam o modelo eficiente, mais simples de ser compreendido e mais fácil de ter seus parâmetros armazenados do que os demais modelos comparados.

Referências

Abry, P. and Veitch, D. (1998). A wavelet analysis of long-range dependent traffic. *IEEE Trans. on Information Theory*, 44:2–15.

- Barman, C., Chaudhuri, H., Deb, A., and Sinha, B. (2015). The essence of multifractal detrended fluctuation technique to explore the dynamics of soil radon precursor for earthquakes. *Natural Hazards*, **78**(2):855–877.
- CODERS, A. (2020). Fractals. <https://wall.alphacoders.com>. Último acesso em 15-04-2020.
- Coninck, Q. D., Baerts, M., Hesmans, B., and Bonaventura, O. (2016). Crawdad dataset. https://crawdad.org/uclouvain/mptcp_smartphone/20160304. Último acesso em 12-04-2020.
- Cunha, M. C. C. (2000). *Métodos Numéricos*. Editora Unicamp, São Paulo, BRA.
- Danzig, P., Mogul, J., Paxson, V., and Schwartz, M. (2000). The internet traffic archive. <http://ita.ee.lbl.gov/>. Último acesso em 12-04-2020.
- Feldmann, A., Gilbert, A. C., and Willinger, W. (1998). Data networks as cascades: Investigating the multifractal nature of internet wan traffic. *ACM SIGCOMM Computer Communication Review*, **28**(4):42–55.
- Grandemange, Q., Bhujwala, Y., Gilson, M., Ferveur, O., and Gnaedinger, E. (2017). An as-level approach to network traffic analysis and modelling. *IEEE International Conference on Communications (ICC)*, pages 1–6.
- Krishna, M. P., Gadre, V., and Desai, U. B. (2003). *Multifractal based network traffic modeling*. Kluwer Academic Publishers.
- Mallat, S. (2008). *A wavelet tour of signal processing*. Academic Press, 3^a edição edition.
- Mandelbrot, B. B. (1990). Limit log normal multifractal measures. *In Frontiers of Physics: Landau Memorial Conference*, (163):309–340.
- Millán, G. and Lefranc, G. (2015). Simple technique of multifractal traffic modeling. *In 2015 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*, pages 329–332.
- Park, K. and Willinger, W. (2000). *Self-Similar Network Traffic and Performance Evaluation*. John Wiley Sons, Inc. New York, NY, USA, 1st edition.
- Riedi, R., Crouse, M., Ribeiro, V., and Baraniuk, R. (1999). A multifractal wavelet model with application to network traffic. *IEEE Transaction on Information Theory*, **45**(3):992–1018.
- Rocha, F. G. C. (2011). Modelagem multifractal aplicada à estimação de probabilidade de perda de dados e ao controle de fluxos de tráfego em redes. *UFG, Goiânia, GO*.
- Tuberquia-David, M., Vela-Vargas, F., López-Chávez, H., and Hernández, C. (2016). A multifractal wavelet model for the generation of long-range dependency traffic traces with adjustable parameters. *Expert Systems With Applications*, **62**(2016):373–384.
- Vieira, F. H. T. and Lee, L. L. (2010). An admission control approach for multifractal network traffic flows using effective envelopes. *International Journal of Electronics and Communications*, **64**(7):629–639.
- Vieira, F. H. T., Rocha, F. G. C., and dos Santos Junior, J. A. (2012). Loss probability estimation and control for ofdm/tdma wireless systems considering multifractal traffic characteristics. *Computer Communications*, **35**(2):263–271.