

Modelagem Estatística e Análise de Desempenho de Tráfego Multifractal

Jeferson Wilian de Godoy Stênico e Lee Luan Ling

Resumo – Neste artigo propomos o uso de mistura gaussiana na modelagem de traços de tráfego com distribuição de cauda pesada. O modelo é avaliado através da estimativa da probabilidade de perda em sistema de fila com um servidor simples. Uma outra contribuição inédita deste trabalho é a derivação de uma expressão geral para o cálculo de probabilidade de perda de bytes para séries de tráfegos com características multifractais. A eficiência da modelagem estatística e a precisão da estimativa de probabilidade de perda são validadas experimentalmente comparando com outros dois métodos multifractais que considera distribuições de cauda pesada específicas (lognormal e Pareto) amplamente usados na literatura.

Palavras-Chave — *Probabilidade de Perda, Processos Multifractais, Tráfego de Redes.*

Abstract — In this paper we propose the use of Gaussian mixture in modeling traffic traces with heavy tail distribution. The model is evaluated by estimating the loss probability in queuing systems with a simple server. Another novel contribution of this paper is the derivation of a general expression for the loss probability estimation for traffic traces with multifractal characteristics. The efficiency of this statistical modeling and the accuracy of the estimated loss probabilities are experimentally validated by comparing with other two multifractal approaches that consider two specific heavy tail distributions (lognormal and Pareto) widely used in the literature.

Keywords — *Loss Probability, Multifractal Processes, Network Traffic.*

I. INTRODUÇÃO

A partir da publicação do trabalho de Leland et. al. [1] houve uma intensificação das pesquisas sobre tráfego de redes envolvendo a teoria dos fractais. Com o tráfego coletado na rede *Ethernet to Bellcore Morristown Research and Engineering Center*, Leland et. al. verificaram experimentalmente propriedades fractais tais como auto-similaridade e dependência de longa duração.

Essas propriedades com destaques para a dependência de longa duração apresentam uma forte influência no desempenho da rede [2], não sendo adequadamente modeladas por processos Poissonianos ou, mais genericamente, modelos Markovianos.

A distribuição de cauda pesada (*heavy tail*) da duração ou tamanho das sessões ou conexões que originam o tráfego agregado é apontada como causa da característica auto-similar observada [3].

Dessa forma, diferentes modelos matemáticos foram propostos com o objetivo de representar a característica auto-similar constatada no tráfego de redes. Sendo que o modelo

tráfego Browniano fracionário (*fractional Brownian traffic*, fBt) tornou-se amplamente utilizado [3].

Entretanto, observou-se que enquanto em escalas de tempo da ordem de centenas de milissegundos e maiores, o comportamento do tráfego era bem representado por modelos auto-similares; em escalas de tempo da ordem de centenas de milissegundos e menores, as características de tais modelos afastavam-se das apresentadas pelo tráfego real. Tal constatação motivou a busca por modelos de tráfego mais abrangentes, que possibilitassem a uma descrição mais completa do tráfego de redes. Com isso, os processos multifractais surgem como uma generalização dos processos auto-similares, permitindo regularidade e leis de comportamento em escala variantes no tempo e, portanto, proporcionando melhor descrição de processos irregulares.

Tradicionalmente, os problemas de modelagem de tráfego foram resolvidos com base na melhor aproximação das propriedades estatísticas do tráfego, quer ele possua características multi-escala ou não. No entanto, acreditamos firmemente que um método de modelagem robusta de tráfego, deve considerar as características estáticas e dinâmicas da rede. Este trabalho pode ser visto com uma boa alternativa para uma modelagem robusta de tráfego de rede. O critério adotado está baseado em duas medidas: aproximações empíricas das distribuições de chegada do tráfego utilizando do método de misturas de Gaussianas, bastante difundida na literatura [4][5] e avaliação de probabilidade de perda em uma conexão.

O artigo está organizado da seguinte forma: Na Seção II, será apresentada a definição dos processos de tráfegos multifractais, revendo alguns conceitos e analisado algumas características. Na Seção III será feita uma breve descrição do Modelo de Mistura de Gaussianas. Na Seção IV, será mostrada a expressão proposta para a estimativa de probabilidade de perda com base em teoria de filas. Na Seção V, serão mostrados os resultados obtidos. Finalmente na Seção VI, conclui-se.

II. PROCESSOS DE TRÁFEGO MULTIFRACTAL

Definição 1: Um processo estocástico $X(t)$ é chamado multifractal se possui incremento estacionário e satisfaz a seguinte equação:

$$E(|X(t)|^q) = c(q)t^{\tau(q)+1}, \forall t \in T, q \in Q \quad (1)$$

onde T e Q são intervalos de números reais, $\tau(q)$ (função de escala) e $c(q)$ (fator de momento) são funções com o domínio Q . Além disso, assume-se que T e Q possuem comprimentos não nulos, e que $0 \in T, [0,1] \subseteq Q$.

Um processo multifractal é um processo que possui vários expoentes de Hölder [6] em pequenas escalas de tempo, em outras palavras, um processo multifractal $X(t)$ é um processo que possui a seguinte propriedade:

$$X(t) = X(t + \Delta t) - X(t) \sim t^{\alpha(t)} \quad (2)$$

Se existe apenas um único expoente de Hölder, isto é, se $\alpha(t)$ é uma constante $\alpha(t) = \alpha$, o processo é chamado de monofractal com parâmetro de Hurst $H = (\alpha + 1)/2$. Em contraste, a estrutura escala de um processo multifractal é muito mais rica, possuindo múltiplos ou até mesmo infinitos expoentes de Hölder, podendo variar com o tempo.

A estrutura escala pode ser descrito estatisticamente com um espectro multifractal. Para mostrar o espectro multifractal, é considerado um intervalo de tempo normalizado $I = [0, 1]$. Dividindo o intervalo em b^n intervalos segundo o modelo de construção proposto [7], no k -ésimo intervalo tem-se: $I_n^k = [kb^{-n}, (k+1)b^{-n}]$, $k = 0, 1, 2, \dots, b^{-n} - 1$. Dessa forma, o expoente de Hölder aproximado para o k -ésimo intervalo é:

$$\alpha_{b^n}^k = \frac{\log_b |X((k+1)b^{-n}) - X(kb^{-n})|}{-n} \quad (3)$$

O exato expoente de Hölder em $k \in I_{b^n}^k$ é obtido por:

$$\alpha(t) = \lim_{kb^{-n} \rightarrow \infty} \alpha_{b^n}^k \quad (4)$$

Obviamente, $\alpha(t)$ pode possuir muitos valores diferentes no conjunto I . Com isso defini-se:

$$p_n^k(\alpha, \varepsilon) = \begin{cases} 1, & \alpha_{b^n}^k \in (\alpha - \varepsilon, \alpha + \varepsilon) \\ 0, & \text{caso contrário} \end{cases} \quad (5)$$

Seja $N_n(\alpha, \varepsilon)$ o número de $\alpha_{b^n}^k$ tomados dentro $(\alpha - \varepsilon, \alpha + \varepsilon)$. O espectro multifractal é então definido como:

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\log_b N_n(\alpha, \varepsilon)}{n} \quad (6)$$

Dessa forma, $f(\alpha)$ é a frequência de que $\alpha(t)$ tem o valor de α , e isto é formalmente chamado de espectro multifractal de grande desvio. Porém, na prática, um espectro multifractal numericamente mais acessível é o espectro de Legendre. Ele é calculado através dos momentos de $X(t)$. Com isso, para um processo multifractal, tem-se a seguinte relação:

$$E(|X(t)|^q) \sim t^{\inf_{\alpha} (q\alpha - f(\alpha))}, t \rightarrow 0 \quad (7)$$

Comparando a Equação (7) com a Equação (1) da definição de processos multifractais, pode-se considerar que:

$$\tau(q) = \inf_{\alpha} (q\alpha - f(\alpha)) \quad (8)$$

A Equação (8) indica que $\tau(q)$ é a transformada de Legendre de $f(\alpha)$, chamada de função escala do processo, e que pode ser obtido segundo a equação proposta do modelo multifractal proposto [7] por:

$$\tau(q) = \log_b \left(\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+q)}{\Gamma(\alpha)\Gamma(\alpha+\beta+q)} \right) - 1 \quad (9)$$

onde α e β são parâmetros de uma distribuição Beta usada na modelagem da cascata e b define o tipo de divisão da cascata multiplicativa. Para maiores detalhes ver [7].

Definição 2: Seja $X(t)$ a taxa de tráfego no instante t , e $W(t) = \int_0^t X(t)dt$ a carga de tráfego acumulado até o instante t e denota-se por $V(t) = W(t + \Delta t) - W(t)$. Dado $T > 0$, um processo de tráfego acumulativo $W(t)$ é dito ser multifractal com escala de tempo T , se as condições seguintes forem satisfeitas:

- i. $W(t)$ tem incremento estacionário com escala de tempo T , isto é, $V(t, T) = V(t)$.
- ii. $V(T)$ tem distribuição de cauda pesada.

Como pode ser observado no item (ii) da definição 2, $V(T)$ deve pertencer a classe de distribuição de cauda pesada [8] porém definir qual distribuição é capaz de descrever melhor a

dinâmica do tráfego multifractal de rede, é uma tarefa trabalhosa, [9]. Neste trabalho, será utilizado de um modelo de mistura de gaussianas (MMG) para descrever as distribuições dos tráfegos reais analisados.

III. MODELO DE MISTURA DE GAUSSIANAS

Um modelo de mistura de gaussianas (MMG) é uma soma ponderada de um número M componentes de densidade de probabilidade. E é dada pela seguinte equação [5]:

$$p(x|\theta) = \sum_{i=1}^M w_i g(x|\mu_i, \Sigma_i) \quad (10)$$

onde x é um vetor D -dimensional dos valores da série que se deseja estimar a densidade, $w_i, i = 1, 2, \dots, M$, são os pesos das misturas, $g(x|\mu_i, \Sigma_i), i = 1, 2, \dots, M$, são as componentes das densidades gaussianas, dada da seguinte forma:

$$g(x|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} e^{-\frac{(x-\mu_i)^2}{2\Sigma_i}} \quad (11)$$

onde, μ_i representa um vetor D -dimensional de médias, e Σ_i é a matriz de covariância $D \times D$ de x , e os pesos das misturas satisfaz a seguinte restrição $\sum_{i=1}^M w_i = 1$.

O modelo de misturas de gaussianas completo é parametrizado pelos vetores de média, covariâncias e pesos, a partir de todas as componentes das densidades. Esses parâmetros serão representados coletivamente pela notação:

$$\theta = \{w_i, \mu_i, \Sigma_i\}, i = 1, 2, \dots, M \quad (12)$$

Existem várias técnicas disponíveis para a estimação dos parâmetros de um MMG [4][5]. Porém o método mais popular é o baseado na estimação de máxima verossimilhança (ML - *Maximum Likelihood*)

a. Estimação de Máxima Verossimilhança

O objetivo da estimativa (ML) é encontrar os parâmetros do modelo que maximizam o MMG para um conjunto de dados. Assumindo $X = \{x_1, x_2, \dots, x_T\}$ como sendo i.i.d. para tornar o problema tratável, pode-se escrever:

$$p(X|\theta) = \prod_{t=1}^T p(x(t)|\theta) \quad (13)$$

No entanto, essa expressão é uma função não linear em relação ao parâmetro θ e fazer a maximização diretamente não é possível. Dessa forma, estimar os parâmetros (ML) pode ser obtido iterativamente usando um caso especial do algoritmo EM (*Expectation-Maximization*).

A ideia básica do algoritmo EM, é iniciar com um modelo inicial θ , e assim estimar um novo modelo $\bar{\theta}$, de tal modo que $p(X|\bar{\theta}) \geq p(X|\theta)$. O novo modelo torna-se então o modelo inicial para a próxima iteração e com isso o processo é repetido até que algum limiar de convergência seja atingido.

Em cada iteração o algoritmo EM, utiliza as seguintes equações na re-estimação dos parâmetros.

$$\bar{w}_i = \frac{1}{T} \sum_{t=1}^T \Pr(i|x(t), \theta) \quad (14)$$

$$\bar{\mu}_i = \frac{\sum_{t=1}^T \Pr(i|x(t), \theta) x(t)}{\sum_{t=1}^T \Pr(i|x(t), \theta)} \quad (15)$$

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^T \Pr(i|x(t), \theta) x(t)^2}{\sum_{t=1}^T \Pr(i|x(t), \theta)} - \bar{\mu}_i^2 \quad (16)$$

A probabilidade a posteriori para o componente i é:

$$\Pr(i|x(t), \theta) = \frac{w_i g(x|\mu_i, \Sigma_i)}{\sum_{k=1}^M w_k g(x|\mu_k, \Sigma_k)} \quad (17)$$

Considerando que $x(t)$ possui a propriedade descrita na Equação (2) então, pode-se escrever as expressões (14),(15),(16) e (17) da seguinte forma:

$$p(X|\theta) = \prod_{t=1}^T p(t^{\tau(q)}|\theta) \quad (18)$$

E conseqüentemente tem-se:

$$\bar{w}_i = \frac{1}{T} \sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta) \quad (19)$$

$$\bar{\mu}_i = \frac{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta) x_t}{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta)} \quad (20)$$

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta) x_t^2}{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta)} - \bar{\mu}_i^2 \quad (21)$$

$$\Pr(i|t^{\tau(q)}, \theta) = \frac{w_i g(x|\mu_i, \Sigma_i)}{\sum_{k=1}^M w_k g(x|\mu_k, \Sigma_k)} \quad (22)$$

Com isso, baseado na teoria apresentada, pode-se então, descrever a nova abordagem para estimar a probabilidade de perda considerando as características multifractais do tráfego.

IV. ESTIMATIVA DA PROBABILIDADE DE PERDA

Nesse trabalho, é proposto uma abordagem para estimar a probabilidade de perda. A motivação principal da abordagem proposta é criação de uma expressão geral para o cálculo de probabilidade de perda, independente da distribuição de cauda pesada um determinado tráfego tenha. Pode-se agora apresentar a equação proposta para a estimação da probabilidade de perda em um servidor considerando tráfego multifractal de entrada.

Proposição 1 *Seja $T > 0$, um processo de tráfego acumulativo $W(t)$ dito ser multifractal na escala de tempo T , com incremento estacionário na escala de tempo T e apresentando distribuição de cauda pesada, aproximada por misturas de gaussianas. A probabilidade de perda para um servidor com taxa C e tamanho de buffer B é dada pela seguinte expressão:*

$$P_{steady} \left(1 - \frac{\lambda}{C}\right) * \sum_{i=1}^M \left(\int_0^{\infty} \left(\frac{1}{\sqrt{2\pi\bar{\sigma}_i^2}} e^{-\frac{((Ct+B)-(\bar{\mu}_i))^2}{2\bar{\sigma}_i^2}} \right) dt \right) \quad (23)$$

$$\text{onde } \bar{\mu}_i = \frac{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta) x_t}{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta)} \text{ e } \bar{\sigma}_i^2 = \frac{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta) x_t^2}{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta)} - \bar{\mu}_i^2$$

Demonstração

Sejam $Q(t)$ o comprimento da fila em servidor em um tempo t ; $V(t - t_0) = W(t) - W(t - t_0)$ a carga de tráfego acumulado no período $(t - t_0)$ e $O(t - t_0)$, denotando a utilização do tráfego em $(t - t_0)$.

Em uma fila em equilíbrio temos:

$$Q(t_0) + V(t - t_0) = Q(t) + O(t - t_0). \quad (24)$$

Assumindo $V(t) = 0$ e $Q(t) = 0$ em $t = 0$ e $t_0 = 0$, pode-se escrever o comprimento da fila $Q(t)$ como:

$$Q(t) = \max(V(t) - O(t), 0). \quad (25)$$

Seja C a taxa de serviço da fila e $S(t)$ duração total quando os servidores estiverem inativos até o tempo t . Com isso, tem-se que:

$$O(t) = C(t - S(t)). \quad (26)$$

Sendo $Y(t) = V(t) - Ct$ e substituindo (26) em (25) tem-se:

$$Q(t) = \max(Y(t) + \Delta(t), 0), \quad (27)$$

onde $\Delta(t) = CI(t)$. Esse resultado em conjunto com a lei de probabilidade, permite escrever a probabilidade de perda na fila como:

$$\begin{aligned} P_{loss}(t) &= P(Q(t) > q) = P(Y(t) + \Delta(t) > q, Y(t) > q) \\ &\quad + P(Y(t) + \Delta(t) > q, Y(t) \leq q) \\ &= P(Y(t) > q) + P(Y(t) \leq q < Y(t) + \Delta(t)) \end{aligned} \quad (28)$$

ou

$$P_{loss} = P(Y(t) > q) + P(Y(t) \leq q < Y(t) + \Delta(t)) \quad (29)$$

O primeiro termo $P(Y(t) > q)$ da Equação (29) chama-se de probabilidade de perda absoluta (P_{abs}) e o segundo termo $P(Y(t) \leq q < Y(t) + \Delta(t))$ de probabilidade oportuna (P_{opp}), no qual um processo multifractal é governado.

Assumindo $Q(t)$ estacionário, e seja $\rho = 1 - \eta = 1 - \lambda/C$, usando o resultado de [10], o segundo termo P_{opp} pode ser escrito como:

$$P_{opp}(t) = P(Y(t) \leq q < Y(t) + \Delta(t)) = \rho \int_0^t f_{V(u)}(v)|_{v=Cu+q} du \quad (30)$$

A probabilidade de perda absoluta P_{abs} é equivalente a:

$$P_{abs}(t) = P(Y(t) > q) = P(V(t) > Ct + q) = \int_{Ct+q}^{\infty} f_{V(t)}(v) dv \quad (31)$$

Então, o comportamento exato da fila é dado por:

$$P_{loss}(t) = \int_{Ct+q}^{\infty} f_{V(t)}(v) dv + \rho \int_0^t f_{V(u)}(v)|_{v=Cu+q} du \quad (32)$$

A primeira integral de (32) pode ser escrita como:

$$P_{abs}(t) = \int_{Ct+q}^{\infty} f_{V(t)}(v) dv = \frac{1}{2} - \frac{1}{2} \operatorname{erf} \left[\frac{\ln(Ct+B) - \bar{\mu}_i(t)}{\sqrt{2\bar{\sigma}_i^2}(t)} \right] \quad (33)$$

Como $P_{abs} \rightarrow 0$ para $t \rightarrow \infty$ então, mostra-se que a probabilidade de perda em regime permanente é dada por:

$$P_{rp}(t) = \lim_{t \rightarrow \infty} P_{loss}(t) = \rho \int_0^t f_{V(u)}(v)|_{v=Cu+q} du \quad (34)$$

Como um processo multifractal apresenta distribuição cauda pesada e pode ser aproximada por misturas de gaussianas, a probabilidade de perda para processos multifractais em uma fila pode ser dada por:

$$P_{steady} \left(1 - \frac{\lambda}{C}\right) * \sum_{i=1}^M \left(\int_0^{\infty} \left(\frac{1}{\sqrt{2\pi\bar{\sigma}_i^2}} e^{-\frac{((Ct+B)-(\bar{\mu}_i))^2}{2\bar{\sigma}_i^2}} \right) dt \right) \quad (35)$$

$$\text{onde } \bar{\mu}_i = \frac{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta) x_t}{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta)} \text{ e } \bar{\sigma}_i^2 = \frac{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta) x_t^2}{\sum_{t=1}^T \Pr(i|t^{\tau(q)}, \theta)} - \bar{\mu}_i^2 \square$$

V. TESTES EXPERIMENTAIS

Para comprovar a eficácia da Equação (35) para o cálculo de probabilidade de perda de *bytes* em uma conexão, simulações com vários traços de tráfego foram realizadas, cujos resultados foram comparados com dois métodos. O primeiro proposto em [11], esse método utilizada de uma aproximação exponencial para a variância com distribuição normal e a entrada do tráfego tendo distribuição lognormal. Para efeito de simplificação, o método será chamado apenas Lognormal. O segundo método proposto em [12], faz uma aproximação exponencial da relação entre a média ao quadrado e a variância em escala de tempo T e a entrada do tráfego possuindo uma distribuição de Pareto. Para efeito de simplificação, esse método será chamado de Pareto.

Para isso, foi utilizado nas simulações traços de tráfegos reais, tráfego TCP/IP denominada *lbl_tcp_3*, retirada de [13], possuindo 1.789.995 amostras, tráfego de vídeo denominado *Desenho dos Simpsons* [14], possuindo 30.334 amostras (vídeo capturado em alta qualidade) e por último um tráfego advindo de uma rede sem fio, coletado da conferência ACM SIGCOMM 2008 [15] e que será denominado por *Sigcomm08*,

possuindo 1.358.782 amostras. Foram consideradas amostras de tráfego em escalas temporais de agregação, onde os traços de tráfego apresentam características multifractais, [16].

Inicialmente são apresentadas nas Figuras 1, 2 e 3 as aproximações das distribuições de cada série de tráfego pelo método de misturas de gaussianas, e pode-se observar que a utilização do MMG é uma alternativa viável na aproximação de distribuições de cauda pesada.

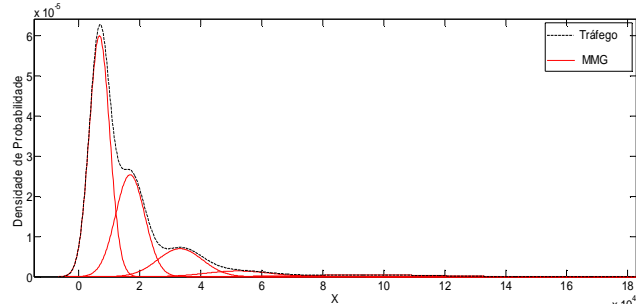


Fig.1: Aproximação da Distribuição do Tráfego *lbl_tcp_3* por Misturas de Gaussianas

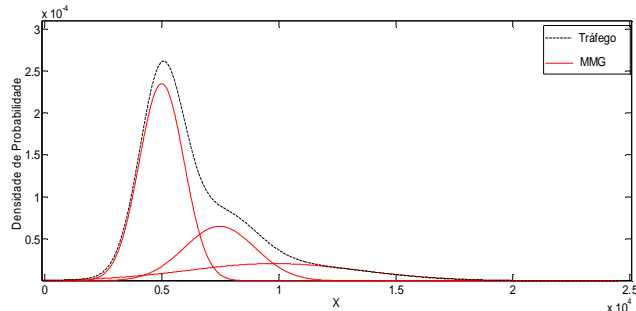


Fig.2: Aproximação da Distribuição do Tráfego de Vídeo (*Desenho do Simpsons*) por Misturas de Gaussianas

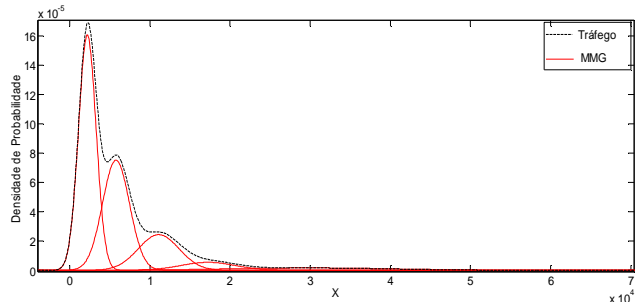


Fig.3: Aproximação da Distribuição do Tráfego de rede sem fio (*Sigcomm08*) por misturas de gaussianas.

Na Tabela I são apresentadas as configurações utilizadas nas simulações para o servidor de rede.

TABELA I CONFIGURAÇÃO DO SISTEMA

Série de Tráfego	Capacidade do Servidor (Bytes/s)	Tamanho do Buffer (Bytes)
<i>lbl_tcp_3</i>	2×10^4	1×10^5
Vídeo	1×10^4	1×10^5
<i>Sigcomm08</i>	1×10^4	1×10^7

Na Tabela II compara as estimativas de probabilidade de perda em bytes para os tráfegos definidos na Tabela I.

- Simulação
- Proposto: aproximação proposta neste artigo;
- Lognormal: a proposta de aproximação exponencial para a variância com distribuição normal e a entrada do tráfego tendo distribuição lognormal [11];
- Pareto: a proposta de aproximação da relação entre a média ao quadrado e a variância em escala de tempo T e a entrada do tráfego possuindo uma distribuição de Pareto [12].

TABELA II PROBABILIDADE DE PERDA

Série de Tráfego	<i>LBL_TCP_3</i>	<i>VIDEO</i>	<i>SIGCOMM08</i>
Simulação	4.70×10^{-3}	3.99×10^{-5}	2.37×10^{-5}
Proposto	4.60×10^{-3}	3.94×10^{-5}	2.34×10^{-5}
Lognormal	4.56×10^{-3}	3.88×10^{-5}	2.30×10^{-5}
Pareto	4.50×10^{-3}	3.87×10^{-5}	2.32×10^{-5}

Nas Figuras 4 a 6 mostram a probabilidade de perda versus o tamanho do buffer, sendo que a abordagem proposta para estimação de probabilidade de perda é dada pela Equação (35). Pode-se notar que apesar dos resultados estarem muito próximos uns dos outros, o desempenho da equação proposta para o cálculo de probabilidade de perda está mais próximo das simulações com os do tráfego real do que os valores obtidos considerando os métodos Lognormal [11], Pareto [12].

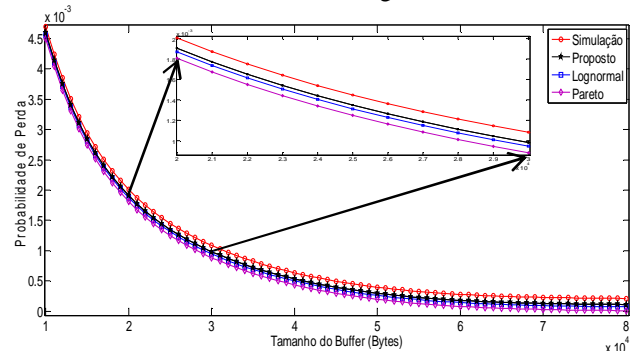


Fig 4 - Probabilidade de Perda versus Tamanho do Buffer para o Traço de Tráfego Internet *lbl_tcp_3*.

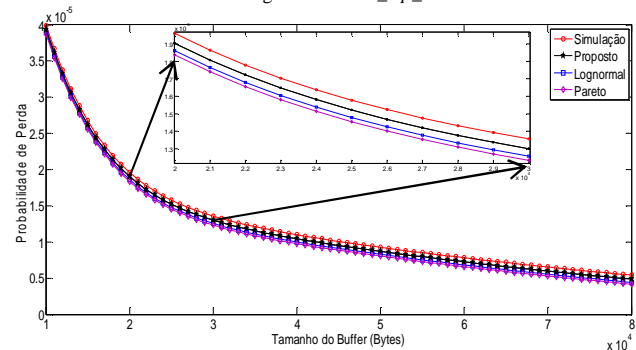


Fig 5 - Probabilidade de Perda versus Tamanho do Buffer para o Traço de Tráfego de Vídeo (*Desenho do Simpson*).

Apresenta-se nas Figuras 7 a 9, a probabilidade de perda em regime permanente em função da capacidade do servidor, comparando novamente com os resultados das abordagens Lognormal [11] e Pareto [12]. E observa-se que a abordagem

proposta consegue melhorar a descrição do tráfego em relação à probabilidade de perda.

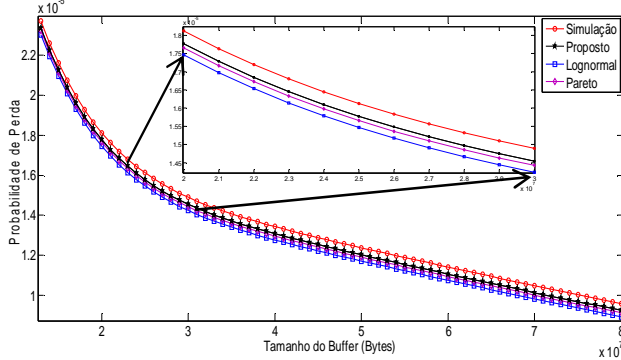


Fig 6 - Probabilidade de Perda versus Tamanho do Buffer para o Traço de Tráfego de Rede Sem Fio (Sigcomm08)

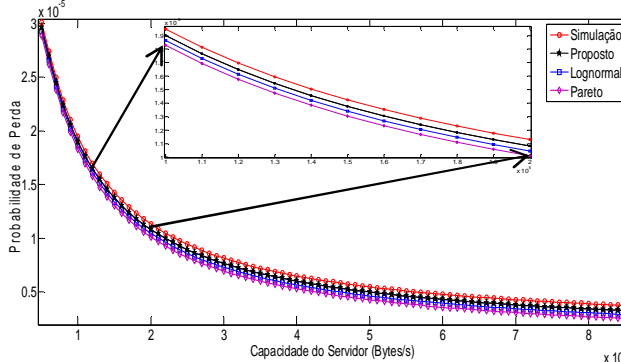


Fig 7 - Probabilidade de Perda versus Capacidade do Servidor para o Traço de Tráfego *lbl_tcp_3*.

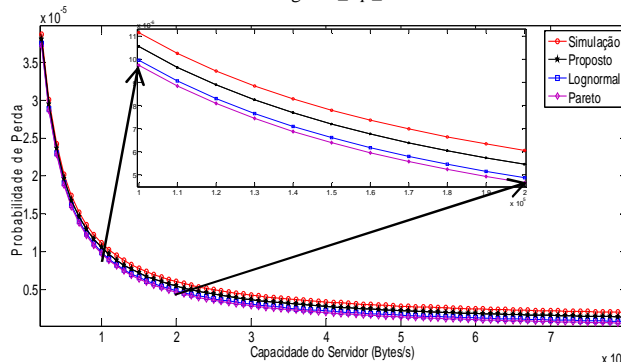


Fig 8 - Probabilidade de Perda versus Capacidade do Servidor para o Traço de Tráfego de Vídeo (*Desenho do Simpsons*).

VI. CONCLUSÃO

Neste trabalho, utilizou-se do método de misturas de gaussianas, muito aplicado na literatura, para fazer aproximações das distribuições de cauda pesada, uma das principais propriedades multifractais do tráfego, a fim de abordar uma solução geral para a estimação de perdas.

Assumiu-se que independente da distribuição de cauda pesada que o tráfego possua, é possível uma aproximação pelo método de misturas de gaussianas. Derivando assim uma equação geral para a caracterização das perdas existentes nos fluxos de tráfego de redes.

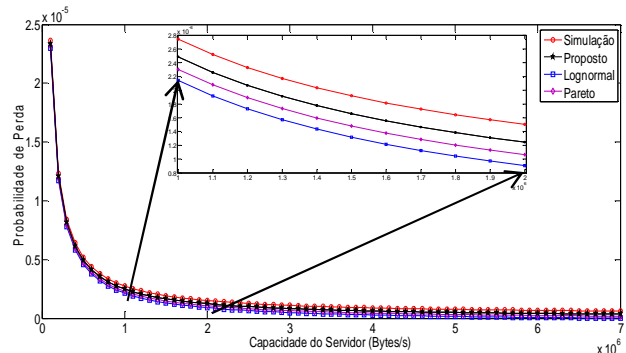


Fig 9 - Probabilidade de Perda versus Capacidade do Servidor para o Traço de Tráfego de Rede Sem Fio (Sigcomm08).

A validação da equação proposta ocorreu através de simulações com séries de tráfego reais. Em relação a alguns modelos existentes na literatura, o modelo Lognormal [11], o modelo Pareto [12] que levam em consideração as características multifractais do tráfego. Os resultados experimentais mostraram a eficiência da abordagem proposta para estimação da probabilidade de perda, conseguindo modelar de forma mais realista e robusta o comportamento de fila no *buffer*, o que a torna uma alternativa analítica promissora no projeto de redes.

REFERÊNCIAS

- [1] Leland W., Taquq M., Willinger W. and Wilson D., On the self-similar nature of Ethernet traffic (extended version), IEEE/ACM Transactions on Networking, v.2, n.1, pp 1-15, Feb 1994.
- [2] Norros, I. A storage Model with Self-Similar Input, Queueing, 16, pp.387-396,1994.
- [3] Park, K. and Willinger, W. Self-Similar Network Traffic and Performance Evaluation. John Wiley and Sons New York, 2000.
- [4] McLachlan, G., ed.: Mixture Models. Marcel Dekker, New York, NY (1988).
- [5] Martinez, W.L; Martinez, A. R. Computational Statistics Handbook with Matlab, Chapman & Hall/CRC, Boca Raton, Florida, 2008.
- [6] Seuret, S. and Gilbert, A.C. Pointwise Hölder Exponent Estimation in Data Network Traffic". In *ITC Specialist Semina*, Monterey,2000.
- [7] Stenico, J. W. G.; Lee, L, L. Modelagem de Processos Multifractais Baseada em uma Nova Cascata Conservativa Multiplicativa, 10/2011, XXIX Simpósio Brasileiro de Telecomunicações - SBRT 11,Vol. 1, pp.1-6, Curitiba, PR, Brasil, 2011.
- [8] Asmussen, S.Ruin Probabilities, World Scientific, Singapura, 2000.
- [9] Stenico, J. W. G.; Lee, L, L. Accurate Second-Order Moment Multifractal Traffic Modelling, April/May 2012, Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos – SBRC 12, Ouro Preto, MG, Brasil, pp. 409- 421, 2012.
- [10] Benes V., General Stochastic Processes in theory of Queues, Reading, MA: Addison Wesley,1963.
- [11] Stenico, J.W.G. and Ling, L.L. A Multifractal Based Dynamic Bandwidth Allocation Approach for Network Traffic Flows IEEE International Conference on Communications (ICC), 23-27 May 2010, pages 1 – 6.
- [12] Stenico, J.W.G and Ling, L.L. A Control Admission Scheme for Pareto Arrivals with Multi-Scale Characteristics In: Proceedings of The International Workshop on Telecommunications - IWT 2011, pp. 220-224, May – 2011, Rio de Janeiro – Brazil.
- [13] <http://ita.ee.lbl.gov/html/traces.html>.
- [14] <http://www.cs.columbia.edu/~hgs/internet/traces.html>
- [15] <http://crawdad.cs.dartmouth.edu/umd/sigcomm2008>
- [16] Riedi, R.H.; Crouse, M.S.; Ribeiro V.J.; and Baraniuk, R.G. A Multifractal Wavelet Model with Application to Network Traffic. IEEE Transactions on Information Theory, 1999.