

ESTUDO DE ESTIMAÇÃO DE BANDA EFETIVA PARA TRÁFEGO AUTO-SIMILAR COM VARIÂNCIA INFINITA

Firmiano R. Perlingeiro e Lee L. Ling

Resumo – A Banda Efetiva é um tópico crucial em termos de dimensionamento dos recursos de uma rede, principalmente quando consideramos a questão de “qualidade de serviço”. Este estudo compara a estimação da banda efetiva baseada na aproximação Gaussiana e a não Gaussiana para tráfego agregado, levando em consideração os processos auto-similares com variância finita e infinita. Brevemente descrevemos os métodos propostos por Norros e N. Laskin baseados no movimento Browniano fracionário e no movimento fracionário de Lévy, respectivamente. Experimentalmente comparamos a banda efetiva estimada para ambos os métodos, através da utilização de tráfego real agregado. Para este fim utilizamos conjuntamente métodos analíticos, heurísticos e de simulação para validar os resultados obtidos. Os resultados e conclusões deste estudo confirmam que devemos tratar de forma diferente os processos de tráfego que apresentam variância infinita daqueles com variância finita.

Palavras-chave – Banda efetiva, auto-similaridade, variância infinita, processos Estáveis, modelos de tráfego, CLT e MRA.

Abstract – With up-coming Quality of Service requirements raised by a wide range of communications demands, the effective bandwidth estimation is one of the approaches to satisfy these QoS requirements. This work compares the effective bandwidth estimation based on Gaussian and non-Gaussian approximations, for aggregated traffic, taking the finite and infinite variance cases into account and evaluating its impacts on bandwidth estimation. We briefly describe the Norros approach based on fractional Brownian motion. Also, we present the generalization method proposed by N. Laskin based on fractional Lévy motion. Experimentally we compare the estimated effective bandwidth using real aggregated traffic traces providing benchmarks for both cases. For this end, analytical, heuristical and simulation methods were used jointly to validate this work. This study confirms that we have carefully differentiated the self-similar processes with finite variance from infinite variance cases, in order to estimate its effective bandwidth for aggregated traffic.

Keywords – Effective bandwidth, self-similarity, infinite variance, Stable process, traffic models, CLT and MRA.

1. Introdução

Dentre os diversos métodos de estimação da banda efetiva [2] disponíveis na literatura, a aproximação Gaussiana proposta por Norros [7] pode ser utilizada para

tráfego agregado. O modelo de agregação proposto é baseado no Teorema do Limite Central (CLT), isto é, um processo Gaussiano pode representar um tráfego composto de grande número de fluxos originados por fontes de tráfego independentes.

Em uma rede, a denominada agregação vertical ou superposição de tráfego é produzida através do procedimento de multiplexação. Este procedimento ocorre quando N fluxos de tráfego partilham o mesmo recurso da rede, tais como “buffers” nos comutadores ou roteadores.

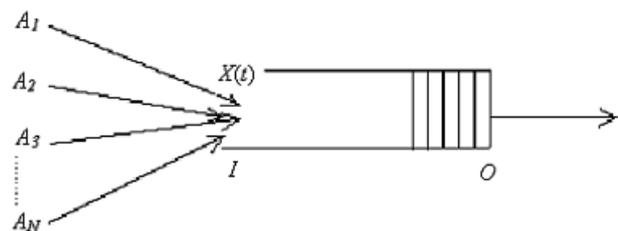


Figura 1: processo de multiplexação

Considere um sistema de “buffer” que está servindo um enlace. $A_i(t)$ representa a quantidade de tráfego do fluxo i na entrada, no intervalo de tempo $[0, t]$, $i = 1, 2, \dots, N$. Assumimos que os N fluxos de tráfego são independentes e que o processo de tráfego $A_i(t)$ é estacionário, isto é, $A_i(t)$ possui incrementos estacionários e $EA_i(t)^2 < \infty$. Com a finalidade de simplificação, o processo $A_i(t)$ é assumido identicamente distribuído [7]. O processo de tráfego agregado verticalmente é denotado por X_t , dado por

$$X_t = X(t) = \sum_{i=1}^N A_i(t) \quad (1)$$

Além das considerações Gaussianas mencionadas, vários estudos de uma grande gama medidas de tráfego e muitos artigos sobre auto-similaridade nos últimos anos têm demonstrado que o tráfego agregado pode ser estatisticamente auto-similar [3][12][13][14]. Contudo, a maioria destes estudos não menciona o caso da auto-similaridade com variância infinita [4]. Em termos de dimensionamento da banda efetiva, além do aspecto da auto-similaridade, devemos considerar se o processo de tráfego apresenta variância finita ou infinita. Geralmente, processos Gaussianos, exponenciais e lognormais são utilizados para modelar os casos de processos auto-similares com variância finita. Os casos que se caracterizam como variância infinita podem ser modelados através de processos mais amplos tais como processos Estáveis e de Pareto.

Neste estudo, apresentamos os modelos para estimação da banda efetiva para tráfego agregado propostos por Norros [7] e Laskin [9]. O objetivo deste trabalho é testar experimentalmente ambos os métodos de estimação da banda efetiva para tráfego agregado, através de medidas de desempenho usando arquivos de tráfego real.

Adicionalmente, introduzimos um fator de otimização para os casos onde ocorra sobredimensionamento.

Este artigo está assim organizado. Na seção 2 apresentamos uma breve descrição sobre auto-similaridade, incluindo o caso da variância infinita. Na seção 3 apresentamos o método proposto por Norros. Na seção 4 introduzimos o método baseado no movimento fracionário de Lévy (mfL) proposto por Laskin. A seção 5 é dedicada a uma breve descrição sobre MRA (*multiresolution analysis*) e o método de Whittle utilizados na estimação dos parâmetros de auto-similaridade. Na seção 6 são apresentados os resultados experimentais. Finalmente na seção 7 concluímos.

2. Processo Auto-similar com variância infinita

Seja X_t é um processo estocástico contínuo com parâmetro de tempo t . X_t é dito auto-similar com parâmetro de Hurst H , se para qualquer a positivo, este processo na escala de tempo $a^{-H} X_{at}$, é igual ao processo original X_t no senso de suas distribuições. Isto significa que para qualquer seqüência de tempo t_1, \dots, t_k , e qualquer constante positiva a , $a^{-H}(X_{at_1}, X_{at_2}, \dots, X_{at_k})$ possui a mesma distribuição de $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ [1]. Além disso, para $1/2 < H < 1$, o processo X_t apresenta a sua autocorrelação decaindo para zero muito lentamente. O movimento Browniano fracionário (fBm) $\{B_H(t), t \geq 0\}$ é um caso clássico de processo auto-similar considerando que $\{B_H(at), t \geq 0\}$ possui as mesmas distribuições dimensionais finitas de $\{a^H B_H(t), t \geq 0\}$ para todo $a > 0$, com média zero e incrementos estacionários [1][4].

Outro exemplo de processo auto-similar, amplamente utilizado para modelar series temporais com características de dependência de longo prazo é o processo F-ARIMA. De acordo com Taqqu e Teverovsky [4][11] o modelo F-ARIMA (1, d , 1) pode apresentar características de variância finita ou infinita, onde d denota o operador de diferença fracionária deste modelo. A inclusão dos processos auto-similares com variância infinita exige uma análise cuidadosa do comportamento dos parâmetros utilizados para caracterizar as propriedades de dependência de longo prazo dos processos, mais precisamente os parâmetros H e d [1][4].

Seja uma série temporal X_i , com comprimento N , definimos uma correspondente série agregada

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i, \quad k=1,2,\dots,N/m \quad (2)$$

Além disso, assumimos que para N/m e m suficientemente grandes, temos

$$X^{(m)} \stackrel{d}{\sim} m^{H-1} S_n \quad (3)$$

onde $\stackrel{d}{\sim}$ indica igualdade assintótica em distribuição. Aqui S_n representa um processo que depende da distribuição de X , mas não depende de m . O termo “assintoticamente auto-similar” é usado para descrever a equação (3). No caso de variância finita, S_n pode ser representado pelo processo FGN (*Fractional Gaussian Noise*) [1]. O FGN é o processo de incremento do fBm, possuindo média zero e variância finita. Se $1/2 < H < 1$, o processo é positivamente

correlacionado, possuindo dependência de longo prazo. No caso de processos auto-similares com variância infinita, o processo S_n na equação (3), necessita ser representado por um processo mais amplo. Um dos métodos apresentados neste estudo, utiliza os processos α -estáveis [9] que definiremos a seguir.

Considere X_1, X_2, \dots , variáveis aleatórias independentes e identicamente distribuídas com a função de distribuição F e uma seqüência $S_n := X_1 + X_2 + \dots + X_n$, $n = 1, 2, \dots$. Suponha que para as seqüências de constantes $i_n > 0$ e $j_n > 0$ ($n = 1, 2, \dots$), a seqüência $S_n / (i_n j_n)$ converge para uma função distribuição particular. Neste caso, a distribuição limite é chamada distribuição “ α -estável” e a função distribuição F é o domínio de atração da distribuição limite. Este tipo de processo utiliza o parâmetro de Lévy $\alpha \in (1, 2]$ para descrever o grau do peso da cauda de sua distribuição. Considerando um sistema de fila em equilíbrio onde S_n seja o processo de carga e o tamanho da fila representado por X teremos

$$X = \max_{n \geq 0} \{S_n\} \quad (4)$$

Então a cauda da distribuição de probabilidade da equação (5) cai lentamente como uma função de potência, ou seja

$$P\{X > b\} \approx Cb^{-\alpha} \quad \text{com } b \rightarrow \infty \quad (5)$$

b é o espaço de espera disponível no sistema e C uma constante, sendo que para $\alpha < 2$, $\text{Var}\{X\} = \infty$ e para $0 < \alpha \leq 1$, $E\{X\} = \infty$.

As distribuições “ α -estáveis”, também são denominadas distribuições de Levy-Pareto. No caso de processos auto-similares com variância infinita, S_n pode ser representado pelo LFSN (*Linear Fractional Stable Noise*) [4].

A identificação dos dois casos de auto-similaridade pode ser feita através da seguinte equação

$$H = d + \frac{1}{\alpha} \quad (6)$$

Quando $\alpha = 2$, temos o caso de variância finita. Neste caso a aproximação Gaussiana proposta por Norros é suficiente. No entanto, quando temos $\alpha < 2$, o modelo Gaussiano falha, caracterizando o caso de variância infinita.

3. O método de Norros

O modelo de tráfego proposto por Norros [6] é fundamentado no fBm e é representado por

$$A_t = mt + \sqrt{am} Z_t, \quad t \in (-\infty, \infty) \quad (7)$$

onde Z_t é o fBm normalizado. Note que A_t é um processo com três parâmetros (m, a, H). A taxa média de entrada quantifica o volume de tráfego transportado é denotada por m , sendo $m > 0$. O parâmetro a ($a > 0$) é um coeficiente de variância (*peakedness*). H denota o parâmetro de Hurst e representa a dependência de longo prazo do processo e o grau de intensidade de surto (*burstiness*) [1][6][8]. Neste modelo, a banda efetiva estimada é dada por [2]

$$C = m + \left(k(H) \sqrt{-2 \ln P\{X > b\}} \right)^{1/H} a^{1/(2H)} b^{-(1-H)/H} m^{1/(2H)} \quad (8)$$

onde $k(H) = H^H(1-H)^{1-H}$, b e X são o tamanho e a carga efetiva do “buffer” em bytes, respectivamente [6].

4. O método de Laskin

Este estudo, que aborda a estimação da banda efetiva para processos auto-similares com variância infinita, foi inspirado no artigo chamado “*Fractional Lévy Motion and its Application to Network Traffic Modeling*” de N. Laskin, I. Lambadaris, F. Harmantzis e M. Devetsikiotis [9].

O movimento de Lévy ordinário (oLm) é o movimento simétrico α -estável de Lévy $L_\alpha = \{L_\alpha(t), t \geq 0\}$, $0 < \alpha < 2$. O oLm é um processo estocástico de Markov que começa em zero, possui incrementos estacionários com $H = 1/\alpha$, isto é,

$$L_\alpha(ct) = c^{\frac{1}{\alpha}} L_\alpha(t), \quad t > 0 \quad (9)$$

N. Laskin, Lambadaris, Harmantzis e Devetsikiotis [9] baseados na generalização do movimento Browniano ordinário para o fBm, proposta por Mandelbrot [6] definiram o fLm (*fractional Lévy motion*) que obedece a integral de Riemann-Liouville

$$L_{\alpha,H}(t) = \frac{1}{\Gamma\left(H + \frac{1}{2}\right)} \int_0^t dL_\alpha(\tau)(t - \tau)^{H-1/2} \quad (10)$$

onde $L_\alpha(t)$ é o movimento simétrico α -estável de Lévy (oLm) e $\Gamma(\cdot)$ denota a função gama. O processo de incremento do fLm é um processo de tempo contínuo definido como

$$\{L_{\alpha,H}(t + \tau) - L_{\alpha,H}(t), \tau > 0\} \quad (11)$$

O fLm é um processo auto-similar com incrementos estacionários e parâmetro de Hurst $H = 1/2 + 1/\alpha$. O processo de incremento dado pela equação (11) também é auto-similar com parâmetro H . O modelo de tráfego de Lévy pode ser visto como uma generalização do modelo de tráfego baseado no fBm proposto por Norros [6]. Em tempo contínuo, o processo cumulativo de tráfego $A(t)$, ou seja, a carga total produzida pelas fontes de tráfego no intervalo de tempo $[0, t]$, $t > 0$, pode ser modelada através de

$$A(t) = mt + (\bar{\sigma}m)^{1/\alpha} L_{\alpha,H}(t) \quad (12)$$

onde $m > 0$ é a taxa média de entrada, $\bar{\sigma} > 0$ é um fator de escala que pode ser interpretado como um índice de dispersão ao redor da média, $\alpha \in (1, 2]$ mede a espessura da cauda da distribuição estável, o parâmetro de auto-similaridade $H \in [1/2 - 1/\alpha, 3/2 - 1/\alpha]$ e $L_{\alpha,H}(t)$ é o processo definido em (10). A banda efetiva estimada para o modelo fLm é dada pela equação (13) [9]

$$c = m + \left(\frac{M_\alpha}{P\{X > b\}} \right)^{\frac{1}{\alpha\left(\frac{H-1}{2}+1\right)}} \frac{1}{\bar{\sigma}^{\alpha\left(\frac{H-1}{2}+1\right)}} m^{\frac{1}{\alpha\left(\frac{H-1}{2}+1\right)}} b^{\left(\frac{H-\frac{3}{2}+1}{H-\frac{1}{2}+1/\alpha}\right)} \quad (13)$$

onde b é o tamanho do “buffer”, X representa a carga de tráfego no “buffer” e ainda

$$M_\alpha = \frac{\bar{\sigma}}{\alpha\pi} \Gamma(\alpha + 1) \sin \frac{\pi\alpha}{2} \quad (14)$$

5. Estimação dos parâmetros H and d

O método utilizado para avaliar as propriedades auto-similares é baseado na decomposição de *wavelet* diretamente via coeficientes. *Wavelets* constituem uma excelente ferramenta matemática em MRA (*multi resolution analysis*). A transformada *wavelet* discreta (DWT) possui a habilidade de localizar um sinal no tempo e em escala.

Matematicamente, a MRA garante a existência de uma função de escala φ (usada para expressar aproximação) e uma função *wavelet* ψ (essencial para definição dos detalhes). O requisito básico da análise multiresolução (MRA) é a obtenção de subespaços $\{v_j\}_{j \in \mathbb{Z}}$ que satisfazem algumas propriedades [5]. A análise multiresolução de um sinal $x(t)$ pode ser pensada como uma projeção do sinal dentro de cada subespaço v_j .

Resumidamente, podemos dizer que a MRA permite escrever uma aproximação X_j , de um sinal X , na escala j (resolução 2^j), como uma soma de uma aproximação rudimentar X_{j+1} na escala $j + 1$ (com resolução 2^{j+1}). Neste caso, o nível de detalhe é dado por $D_{j+1} = X_j - X_{j+1}$, isto é, a diferença entre as duas aproximações. É possível repetir este procedimento, por exemplo, escrevendo na escala $j + 1$, X_{j+1} , como uma soma de uma aproximação rudimentar X_{j+2} , sendo a diferença $D_{j+2} = X_{j+1} - X_{j+2}$, e assim por diante, então $X_j = X_{j+1} + D_{j+1} = X_{j+2} + D_{j+1} + D_{j+2}$ [18].

Os detalhes entre os subespaços aproximados, espalhados em várias escalas da função escala, podem ser expressos por

$$detail_j(t) = approx_j(t) - approx_{j-1}(t) \quad (15)$$

As funções que espalham as diferenças entre o subespaço v_j são as *wavelets* $\psi_{j,k}(t)$. Como resultado, o subespaço w_j pode ser definido

$$\begin{aligned} v_2 &= v_3 \oplus w_3 \\ v_1 &= v_2 \oplus w_2 \\ v_0 &= v_1 \oplus w_1 \\ v_0 &= v_j \oplus w_j \oplus w_{j-1} \oplus w_{j-2} \dots \oplus w_1 \quad (16) \end{aligned}$$

A equação (16) pode ser obtida diretamente das projeções de x dos subespaços *wavelets*. A teoria MRA mostra que existe uma função ψ_0 , denominada *wavelet* mãe. Detalhes sobre MRA consultar a referência [18].

O método denominado estimador de Whittle foi utilizado para estimar o valor do parâmetro d usando periodograma [5]. Este estimador é semi-paramétrico visto que somente especifica a forma paramétrica da densidade espectral f quando a frequência ν é próxima de zero. O método assume que

$$f(\nu) \sim G(d) |\nu|^{-d}, \quad \text{as } \nu \rightarrow 0 \quad (17)$$

onde $G(d)$ é alguma distribuição de d . O estimador Whittle assume somente o comportamento da densidade espectral acima de alguma frequência $2\pi M/N$. O parâmetro d pode ser estimado

$$R(d) = \log \left(\frac{1}{M} \sum_{j=1}^M \frac{I(v_j)}{(v_j)^{-2d}} \right) - 2d \frac{1}{M} \sum_{j=1}^M \log v_j \quad (18)$$

6. Resultados Experimentais

Os arquivos de tráfego analisados neste estudo foram capturados na rede da Petrobras através de um analisador de dados DA350 da Acterna™, com uma resolução (*time stamp*) de 32 microssegundos. Neste estudo foram testados mais de 100 arquivos de tráfego real agregado. Dentre estes, identificamos 33 que apresentaram comportamento auto-similar, ou seja, $\frac{1}{2} < H < 1$. Para cada arquivo de tráfego auto-similar identificado avaliamos a relação entre os parâmetros H e d . Os resultados desta caracterização de tráfego são apresentados na Tabela 1 [4][11].

A notação adotada para designar os diferentes arquivos de tráfego real analisados neste estudo é a seguinte. Os arquivos de tráfego agregado designados com a letra “S” significam que os arquivos de tráfego foram capturados junto aos servidores de aplicações. Arquivos de tráfego agregado designados pela letra “I”, foram capturados em roteador de tráfego Internet. Os arquivos de tráfego agregado designados através da letra “R” foram capturados em roteador de tráfego IP corporativo. A Figura 2 mostra o cenário de captura dos arquivos de tráfego real analisados neste estudo.

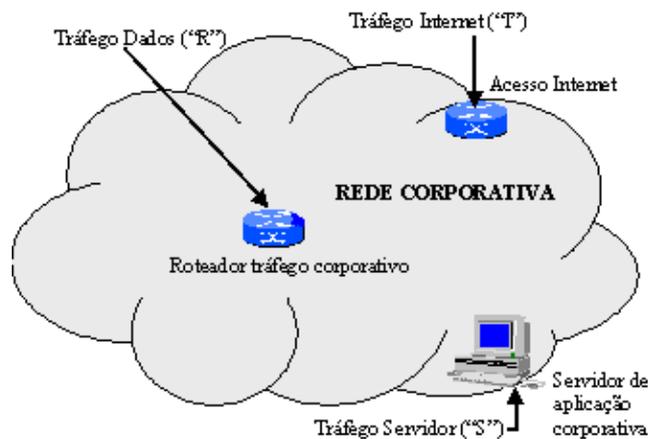


Figura 2: Cenário de captura do tráfego

Através da análise dos parâmetros de auto-similaridade, H e d , apresentados na Tabela 1 e aplicando-se a equação (8), podemos verificar que os arquivos de tráfego agregado do tipo servidor (“S”) não obedecem à relação $H = d + \frac{1}{2}$, isto sugere que o tráfego pode apresentar comportamento auto-similar com variância infinita. Por outro lado, os demais tráfegos agregados como os de internet (“I”) e tráfegos de dados corporativos (“R”) obedeceram à relação $H = d + \frac{1}{2}$, caracterizando-se como os tráfegos agregados auto-similares de variância finita.

Com a finalidade de validar os resultados sobre as características auto-similares dos arquivos de tráfegos reais obtidos e apresentados na Tabela 1, estimamos a banda efetiva para cada um dos arquivos de tráfego aplicando os modelos de Norros (3) e Laskin (13), respectivamente. Adicionalmente, executamos testes de simulação, usando os arquivos de tráfego real, tendo entrada do sistema os seguintes parâmetros: tamanho do “buffer”, perda admissível em bytes (BLR) no valor de 10^{-3} , obtendo-se como resultado da simulação a banda efetiva estimada. O simulador utilizado foi o *SmartBand* [17][18] desenvolvido pelo Projeto UNI-20 da Universidade Estadual de Campinas. Finalmente comparamos os resultados obtidos de estimação com os resultados obtidos em simulação. É esperado que o tráfego agregado auto-similar com variância infinita necessite mais banda efetiva do que o tráfego que apresenta variância finita, devido ao primeiro apresentar um comportamento de surtos mais intenso.

Trace	H	d	$d+1/2$	α
10_7_S_1	0.61	0.37	0.87	1.639344
10_7_S_2	0.71	0.47	0.97	1.408451
10_7_S_3	0.51	0.09	0.51	1.960784
10_7_S_4	0.75	0.09	0.51	1.333333
5_7_S_1	0.46	0.37	0.87	2.173913
5_7_S_2	0.58	0.41	0.97	1.724138
5_7_S_3	0.52	0.36	0.86	1.923077
5_7_S_4	0.49	0.48	0.98	2.040816
5_7_S_5	0.49	0.48	0.98	2.040816
5_7_S_6	0.52	0.42	0.92	1.923077
3_7_I_1	0.71	0.21	0.71	1.408451
3_7_I_2	0.64	0.18	0.68	1.5625
3_7_I_3	0.77	0.18	0.68	1.298701
3_7_I_4	0.60	0.22	0.72	1.666667
3_7_I_5	0.71	0.21	0.71	1.408451
4_7_I_1	0.77	0.20	0.70	1.298701
4_7_I_2	0.74	0.19	0.69	1.351351
4_7_I_3	0.66	0.20	0.70	1.515152
4_7_I_4	0.80	0.21	0.71	1.25
4_7_I_5	0.74	0.19	0.69	1.351351
4_7_I_6	0.68	0.20	0.70	1.470588
4_7_I_7	0.69	0.20	0.70	1.449275
4_7_I_8	0.77	0.21	0.71	1.298701
4_7_I_9	0.81	0.27	0.77	1.234568
4_7_I_10	0.69	0.18	0.68	1.449275
3_7_R_1	0.66	0.17	0.67	1.515152
3_7_R_4	0.67	0.18	0.68	1.492537
3_7_R_7	0.64	0.15	0.65	1.5625
4_7_R_1	0.63	0.19	0.69	1.587302
4_7_R_2	0.66	0.21	0.71	1.515152
4_7_R_3	0.62	0.16	0.66	1.612903
4_7_R_4	0.62	0.21	0.71	1.612903
4_7_R_5	0.64	0.14	0.64	1.5625

Tabela 1: Estimação dos parâmetros H e d

Apesar do modelo proposto por Laskin atender os requisitos de banda efetiva para o tráfego agregado auto-similar de variância infinita, observamos que este modelo pode superestimar a banda efetiva em alguns casos. Diante deste fato, propomos a utilização de um fator de otimização juntamente como o uso deste modelo. Este fator de

otimização varia conforme a intensidade do parâmetro H nas seguintes faixas:

$$f_{op} = \frac{2 EB}{5 \sqrt{b'L}}, \quad \text{if } 0 < H \leq 0,7 \quad (19)$$

$$f_{op} = \frac{2 EB}{75 \sqrt{b'L}}, \quad \text{if } H > 0,7 \quad (20)$$

onde L é o tamanho máximo de surto ou rajada. H é o parâmetro de Hurst. EB é a banda efetiva estimada. b' é o tamanho de "buffer" normalizado ($b' = b/b_0$), onde b_0 denota o tamanho mínimo "buffer" considerado. Neste trabalho, o tamanho mínimo de "buffer" utilizado tem como referência o tamanho máximo de um quadro em uma rede *Ethernet*, que é 1518 bytes.

As Figuras 3, 4, 5, 6 e 7 mostram exemplos de resultados de simulação de perda de pacotes em relação à banda efetiva estimada para arquivos de tráfego agregado que apresentaram comportamento auto-similar com variância infinita. Cada figura mostra a banda efetiva simulada, e as bandas efetivas estimadas pelos métodos de Norros, Laskin e proposto com otimização.

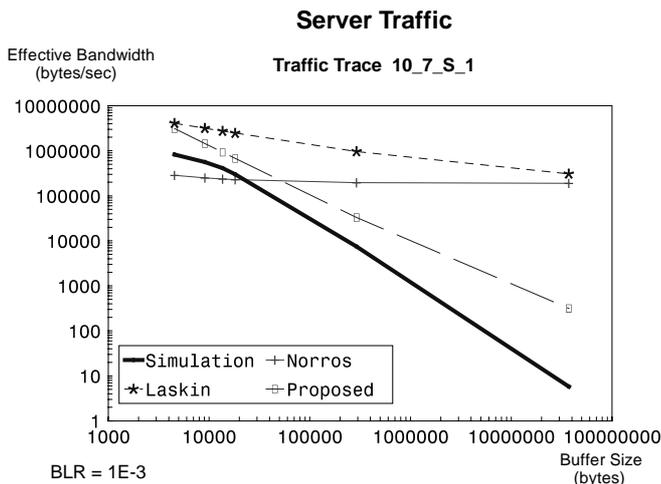


Figura 3: Arquivo Tráfego Servidor 10_7_S_1

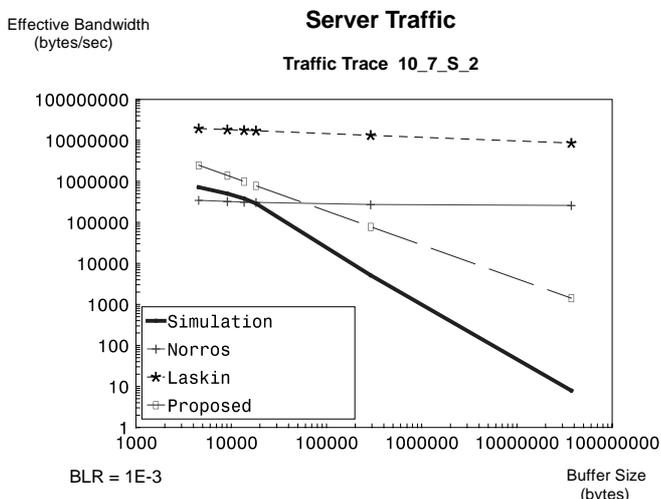


Figura 4: Arquivo Tráfego Servidor 10_7_S_2

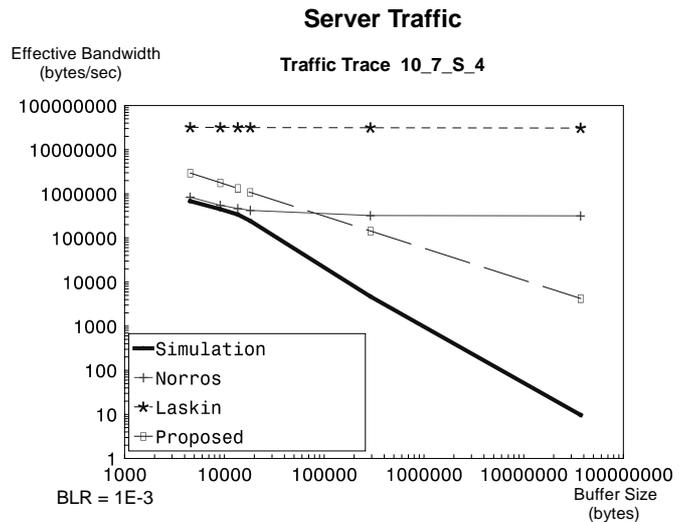


Figure 5: Arquivo Tráfego Servidor 10_7_S_4

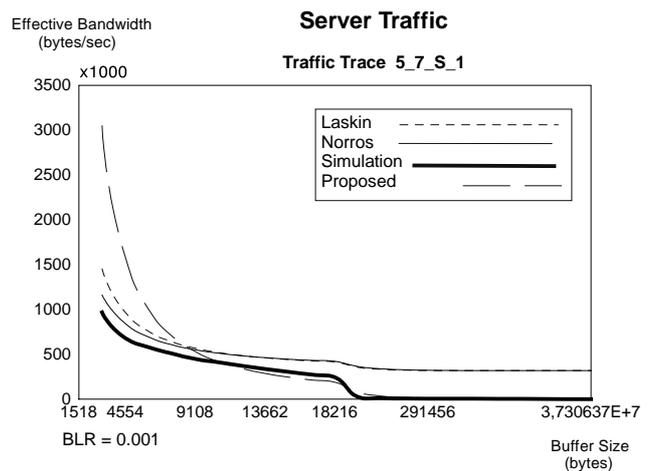


Figura 6: Arquivo Tráfego Servidor 5_7_S_1

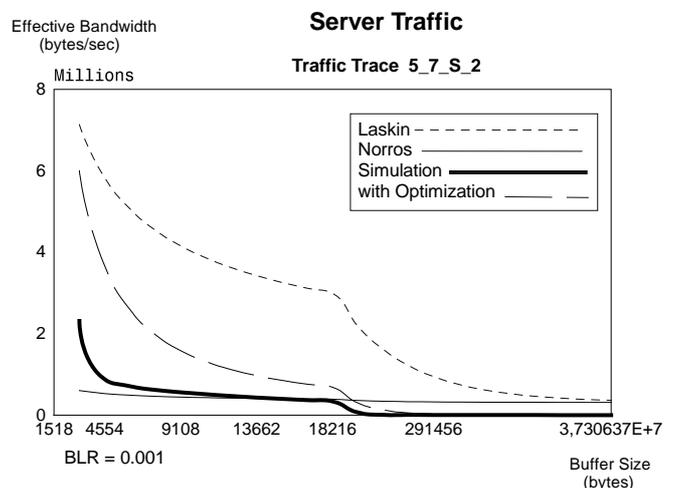


Figura 7: Arquivo Tráfego Servidor 5_7_S_2

7. Conclusões

Dentre os três tipos de tráfegos agregados analisados neste estudo, sendo: internet (“I”), dados corporativos (“R”) e servidor (“S”), apresentados na Tabela 1, somente o tráfego de servidor (“S”) apresentou a característica auto-similar com variância infinita. A possível explicação para este fato, é a seguinte. Conforme mostrado na Figura 2, os arquivos de tráfego de internet (“I”) e dados corporativos (“R”) foram capturados nos elementos de rede como roteadores no meio da nuvem (da rede corporativa). Este tipo de tráfego está menos sujeito a influência do nível de aplicação. Por outro lado, o tráfego de servidor (“S”) foi capturado no acesso a rede (no ponto de rede do servidor). Neste ponto, é possível que a influência dos controles de fluxos da rede e do próprio TCP sejam maiores. Este fato pode ser uma das causas deste tipo de tráfego apresentar um comportamento auto-similar com variância infinita.

Através da análise das Figuras 3, 4, 5, 6 e 7, podemos observar que para este tipo de tráfego agregado a aproximação Gaussiana não funciona plenamente em algumas faixas de tamanho de “buffer” causando perda de informação. Nas Figuras 3, 4 e 7 observa-se a falha da aproximação Gaussiana que subestima a banda efetiva para tamanhos de “buffer” menores que 30.000 bytes, 20.000 bytes e 12.000 bytes, respectivamente. Por outro lado, o modelo proposto por Laskin, baseado na aproximação α -estável, atendeu adequadamente a estimação da banda efetiva quando comparada com os testes de simulação. Contudo, o modelo de Laskin pode em alguns casos superestimar a banda efetiva necessária para atendimento de um determinado tráfego. Nas Figuras 3, 4, 5 e 7 verifica-se este sobredimensionamento, necessitando a aplicação do fator de otimização proposto.

Em termos práticos, este estudo confirma as conclusões obtidas por Taquq e Teverovsky [4][11] que recomendam uma análise criteriosa para diferenciar os processos auto-similares que apresentam variância infinita daqueles com variância finita. Além disso, testamos os modelos propostos por N. Laskin, I. Lambadaris, F. Harmantzis e M. Devetsikiotis [9], para estimação da banda efetiva baseado na aproximação no processo α -estável de Lévy e o modelo proposto por Norros [6] baseado na aproximação Gaussiana. Observamos que o modelo proposto por Norros para a estimação da banda efetiva para tráfego agregado auto-similar, satisfaz o requisito de QoS banda efetiva quando o tráfego apresenta comportamento auto-similar com variância finita, que é o caso dos arquivos de tráfego do tipo internet e (“I”) e dados corporativo (“R”), conforme listados na Tabela 1. Mas este modelo não se mantém para todos os casos de processos auto-similares com variância infinita analisados neste estudo.

Baseado na caracterização do tráfego auto-similar, nas análises das estimativas de banda efetiva obtidas pelos diferentes métodos e nos testes de simulação, concluímos que o modelo proposto por Laskin com simultâneo uso do fator de otimização se mostra mais realista para aplicação em tráfego agregado auto-similar com variância infinita.

8. Referências Bibliográficas

- [1] Jan Beran, “*Statistics for Long-Memory Process*”, Chapman & Hall, 1994.
- [2] Frank Kelly, “*Notes on Effective Bandwidth*”, University of Cambridge, 1995.
- [3] Jonathan L. Wang, “*Analysis, Modeling and Management of Self-Similar Traffic*”, SBT/IEEE/ITS 1998, São Paulo, Brazil.
- [4] M. S. Taquq and V. Teverovsky “*On Estimating the Intensity of Long-Range Dependence in Finite and Infinite Variance*”, Book Statistical Techniques for Analyzing Heavy Tailed Distributions, R. Feldman and M. S. Taquq, editors. Birkhauser, Boston.
- [5] M. S. Taquq and V. Teverovsky “*Semi-parametric Graphical Estimation Techniques for Long-range Dependence*”, Athens Conference on Applied Probability and Time Series Analysis, Vol. 2, pp. 420-432, New York, 1996. Springer-Verlag. Lecture Notes in Statistics, 115.
- [6] Ilkka Norros, “*On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks*”, IEEE Journal On Selected Areas in Communications, Vol.13, August 1995.
- [7] Ilkka Norros, Parag Pruthi, “*On the Applicability of Gaussian Traffic Models*”, Thirteenth Nordic Teletraffic Seminar, 1996.
- [8] Ilkka Norros, “*A Storage Model with Self-Similar Input*”, Queueing Systems, 1994.
- [9] N.Laskin, I.Lambadaris, F.C.Harmantzis and M.Devetsikiotis, “*Fractional Lévy Motion and its Application to Network Traffic Modeling*”, ITC17, Salvador, Brazil, December 2001.
- [10] Fotios C.Harmantzis, Dimitrios Hatzinakos and Irene Katzela, “*Tail Probabilities for the Multiplexing of Fractional α -Stable Broadband Traffic*”, ICC, Helsinki, Finland, June 2001.
- [11] Biplad Sikdar and Kenneth S. Vastola, “*On the Convergence of MMPP and fractional ARIMA Process with Long-Range Dependence to Fractional Brownian Motion*”, Proceedings of the 34th CISS, Princeton, NJ, 2000.
- [12] Ashok Erramilli, Jonathan L. Wang, “*Monitoring Packet Traffic Level*”, IEEE Globecom, 1994.
- [13] Javier Aracil, Richard Edell, Pravin Varaiya, “*A Phenomenological Approach to Internet Traffic Self-Similarity*”, UCB, 1996.
- [14] J.Neilson, P.R.Morin. “*The Impact of Self-Similarity on Network Analysis*”, Carleton University, School of Computer Science, Ottawa, Canada, 1995.
- [15] Judith L.Jerkins, Jonathan L. Wang, “*A measurement Analysis of ATM Cell-Level Aggregate Traffic*”, IEEE Globecom, November 1997.
- [16] W.E. Leland, W.Willinger, D.V.Wilson, “*On The Self-Similar Nature of Ethernet Traffic (Extended Version)*”, IEEE, pp 1-15, 1994.
- [17] Relatório 1, Projeto Ericsson UNI-20, Universidade Estadual de Campinas, Brasil, Março 2001.
- [18] Relatório 2, Projeto Ericsson UNI-20, Universidade Estadual de Campinas, Brasil, Setembro 2001.
- [19] Relatório 3, Projeto Ericsson UNI-20, Universidade Estadual de Campinas, Brasil, Março 2002.
- [20] Relatório 4, Projeto Ericsson UNI-20, Universidade Estadual de Campinas, Brasil, Novembro 2002.