

Um Modelo Multifractal Aplicado à Predição de Tráfego de Redes

Gabriel Rocon Bianchi, Flávio Henrique Teles Vieira e Lee Luan Ling

Resumo—Este trabalho propõe um novo modelo de tráfego multifractal e apresenta um novo preditor de tráfego baseado no modelo proposto. O modelo de tráfego proposto é uma extensão do tráfego Browniano fracionário, um clássico modelo proposto por Norros [5]. Baseado na estrutura de correlação do modelo proposto, um preditor linear de k -passos é implementado. Testes realizados com traços de tráfegos reais demonstram a eficácia do modelo e do preditor propostos.

Palavras-Chave—Redes de comunicação, Internet, tráfego multifractal, predição, análise de desempenho.

Abstract—In this work, we propose a novel multifractal network traffic model and a traffic predictor based on it. This multifractal traffic model extends the notion of the classical fractional Brownian traffic proposed by Norros [5]. Based on the traffic model correlation structure, a discrete-time k -step linear predictor is implemented. Tests carried out on the model as well as on the predictor show their effectiveness.

Index Terms—Network communications, Internet, multifractal traffic, prediction, network performance analysis.

I. INTRODUÇÃO

As modernas redes de comunicação suportam diferentes aplicações, com diferentes requisitos de qualidade de serviço. Para garantir a qualidade de serviço requerida, deve-se empregar estratégias de gerenciamento de tráfego assim como mecanismos de controle. Tais mecanismos de controle devem ser capazes de lidar com tráfegos extremamente complexos. Dessa forma a aplicação de modelos de tráfego mais precisos pode melhorar o desempenho de tais mecanismos de controle.

Trabalhos como o proposto por Leland et al [4] demonstraram que o tráfego de redes exibe propriedades fractais, tais como auto-similaridade e dependência de longo prazo. Modelos auto-similares são capazes de descrever o comportamento dos surtos de tráfego apenas em grandes escalas de tempo (ordem de segundos ou mais). Devido principalmente à sua simplicidade, o movimento Browniano fracionário (fBm) tornou-se o modelo auto-similar mais amplamente empregado [4][5]. Embora modelos auto-similares produzam resultados precisos para determinados cenários, existem outros cenários onde tais modelos falham na representação do comportamento real do tráfego. Na

Gabriel Rocon Bianchi, Flávio Henrique Teles Vieira e Lee Luan Ling, Faculdade de Eng. Elétrica e Computação, Universidade Estadual de Campinas, Campinas, Brasil. E-mails: bianchi@decom.fee.unicamp.br, flavio@decom.fee.unicamp.br, lee@decom.fee.unicamp.br. Este trabalho foi parcialmente financiado pelo CNPq e Ericsson Research (Contrato UNI-20/2320).

realidade, a auto-similaridade e a dependência de longo prazo não são as únicas características fractais do tráfego, e uma descrição global do comportamento do tráfego passa necessariamente por sua caracterização em escalas de tempo pequenas (ordem de centenas de milissegundos ou menos). Recentemente, diversos trabalhos mostraram que o comportamento global do tráfego de redes WAN é convenientemente descrito usando a análise multifractal [1][2][3].

Peltier e Levy Vehel [6] propuseram o movimento Browniano multifractal (mBm), generalizando o fBm com expoente H , para o caso onde H não é mais uma constante, mas sim uma função do tempo. Assim como no fBm, no mBm o valor de $H(t)$ é equivalente ao expoente Hölder, indicando a regularidade do processo em um determinado instante t . É sabido que o fBm é um processo auto-similar que permite convenientemente descrever sinais irregulares que ocorrem em diferentes situações. Entretanto, uma vez que o fBm possui um valor constante de H , a regularidade pontual do fBm é a mesma ao longo de todo o processo, fato este que limita seu campo de aplicação.

Este artigo inova ao propor uma extensão do amplamente usado tráfego Browniano fracionário proposto por Norros [5]. Buscando incorporar características multifractais ao tráfego Browniano fracionário, substitui-se pelo mBm o processo fBm utilizado no modelo original.

A partir do modelo proposto, utilizando-se as fórmulas explícitas de correlação do mBm apresentadas em [7], propõe-se um preditor ótimo no sentido do erro quadrático médio para processos mBm, seguido por sua aplicação em traços de tráfego reais.

II. TRÁFEGO BROWNIANO FRACIONÁRIO

Norros [5] propôs um modelo de processo de ocupação de armazenamento chamado armazenamento Brownian fracionário. Este modelo é alimentado por um processo gaussiano auto-similar denominado tráfego Browniano fracionário definido a seguir.

O tráfego Brownian fracionário $A(s,t)$ em um intervalo de tempo $(s,t]$ é definido da seguinte forma

$$A(s,t) = m(t-s) + \sigma(Z_t - Z_s) \quad s, t \in (-\infty, \infty), s \leq t \quad (1)$$

onde m é a taxa média de entrada, σ é o desvio padrão do processo de chegada de tráfego, e o processo $(Z_t)_{t \in \mathbb{R}}$ é o

movimento Browniano fracionário normalizado. O fBm normalizado é um processo gaussiano de média zero e seu processo de incremento, denominado ruído gaussiano fracionário (fGn) é gaussiano e estacionário.

O tráfego Browniano fracionário é uma maneira simples de incluir, em análises estatísticas de desempenho de redes, as características auto-similares observadas no tráfego. Entretanto, dada sua propriedade gaussiana, em escalas de tempo pequenas o modelo de tráfego Browniano fracionário falha, uma vez que grande parte do tráfego modelado é negativa. Além disso, o fBm é um processo auto-similar onde um único parâmetro H rege tanto a correlação de longa duração quanto a sua regularidade, que possui valor constante (monofractal).

A utilização de modelos que permitam variar o expoente Hölder ao longo de sua trajetória é desejável na representação de processos cuja regularidade varie com o tempo, tais como tráfegos WAN. A fim de superar as limitações relativas à regularidade Hölder do modelo, este trabalho propõe a extensão do tráfego Browniano fracionário, substituindo pelo mBm o processo fBm originalmente empregado.

III. MOVIMENTO BROWNIANO MULTIFRACIONÁRIO

Sejam (X, d_X) e (Y, d_Y) dois espaços métricos. Uma função $f: X \rightarrow Y$ é denominada uma função Hölder com expoente $\beta > 0$, se para cada $x, y \in X$ tal que $d_X(x, y) < 1$ tenha-se:

$$d_Y(f(x), f(y)) \leq c \cdot d_X(x, y)^\beta \quad (2)$$

para alguma constante $c > 0$.

Seja $H: [0, \infty) \rightarrow [a, b] \subset (0, 1)$ uma função Hölder com expoente $\beta > 0$. Para $t \geq 0$, a seguinte função aleatória W é denominada movimento Browniano multifracionário com função $H(t)$:

$$W_{H(t)}(t) = \int_{-\infty}^0 [(t-s)^{H(t)-1/2} - (-s)^{H(t)-1/2}] dB(s) + \int_0^t (t-s)^{H(t)-1/2} dB(s) \quad (3)$$

onde B é o movimento Browniano ordinário.

A partir da definição acima, tem-se que o mBm é um processo gaussiano cujos incrementos em geral não são nem gaussianos nem estacionários. Tem-se também que quando $H(t)=H$ para todo t , então o mBm é apenas o fBm com expoente H .

A principal característica do processo mBm é que sua regularidade varia no tempo, sendo esta igual a $H(t)$. Esta característica contrasta com o fBm, onde seus expoentes Hölder são constantes e iguais a H . Portanto, enquanto apenas um único número H governa o fBm, para o mBm, uma função $H(t)$ se faz necessária. Esta característica do mBm é bastante útil quando se deseja uma modelagem mais precisa. Para o caso do fBm, uma vez que o parâmetro H governa tanto as altas frequências relacionadas com a regularidade e as baixas frequências ligadas à dependência de longa duração, não é

possível por exemplo, ter ao mesmo tempo um comportamento muito irregular (implicando em H próximo a zero) e dependência de longo prazo (implicando em $H > 1/2$). Portanto, o fBm não é capaz de modelar a contento processos que apresentem simultaneamente estas duas propriedades, tais como traços de tráfego WAN TCP/IP.

A. Ayache e Lévy Vehel [7] obtiveram expressões explícitas para a estrutura de correlação do mBm, mostradas a seguir.

Seja $X(t)$ um processo mBm padrão com função $H(t)$. A auto-correlação de $X(t)$ é dada por:

$$cov_X(t, s) = D(H(t), H(s)) (t^{H(t)+H(s)} + s^{H(t)+H(s)} - |t-s|^{H(t)+H(s)}) \quad (4)$$

onde

$$D(x, y) = \frac{\sqrt{\Gamma(2x+1)\Gamma(2y+1)\sin(\pi x)\sin(\pi y)}}{2\Gamma(x+y+1)\sin(\pi(x+y)/2)} \quad (5)$$

com Γ sendo a função gama.

Percebe-se através de (4) que para o cálculo da auto-correlação é necessário que se conheça a função Hölder $H(t)$. A estimação da função Hölder não é uma tarefa trivial, e quando comparada com a estimação do parâmetro H para o caso do fBm, apresenta maior complexidade. Neste trabalho utilizamos o robusto estimador proposto em [8], o qual permite obter boa precisão para sinais reais.

IV. TRÁFEGO BROWNIANO FRACIONÁRIO ESTENDIDO

O modelo de tráfego proposto neste trabalho consiste na substituição do processo fBm padrão $(Z_t)_{t \in \mathbb{R}}$ apresentado em (1), pelo processo movimento Browniano multifracionário normalizado $(\tilde{W}_t)_{t \in \mathbb{R}}$, mantendo-se os outros parâmetros m e σ inalterados. Sendo o mBm uma generalização do fBm, para uma função $H(t)$ constante, o modelo de tráfego proposto torna-se idêntico ao modelo tráfego Browniano fracionário.

O modelo proposto, quando comparado com o modelo de Norros, é menos parcimonioso, uma vez que ao invés de necessitar de apenas um único parâmetro H , agora se faz necessário uma função Hölder dependente do tempo. Entretanto, o modelo proposto possui a vantagem de ser capaz de capturar não só as estatísticas de segunda ordem (dependência de longo prazo), mas também características multifractais. A. Ayache e Véhel apresentam em [7] uma discussão sobre a característica de dependência de longo prazo presente em processos mBm.

Utilizando-se traços de tráfego reais, foram realizados dois testes que demonstram que o modelo proposto pode efetivamente capturar tanto a característica de dependência de longo prazo quanto as características multifractais. Com este objetivo foram utilizados dois traços de tráfego de redes WAN TCP/IP (dec-pkt-1.tcp e dec-pkt-2.tcp) pertencentes à *Digital Equipment Corporation*, coletados no ponto de acesso

primário da DEC à Internet. Estes traços de tráfego estão disponíveis em <http://ita.lbl.gov/htmlcontrib/DEC-PKT.html>.

Os testes realizados restringiram-se a analisar as primeiras 2048 amostras de tráfego, em uma escala de agregação de 512ms. A escolha da escala de 512ms deveu-se a esta ser a *fBm lower cut-off timescale* para os tráfegos analisados, conforme mostrado em [3]. Erramilli em [3] estabelece que a *fBm lower cut-off timescale* é a escala de tempo de transição entre a aplicabilidade de modelos auto-similares e multifractais. Alguns estudos mostraram também que as características multifractais manifestam-se em escalas de tempo ainda superiores a *fBm lower cut-off timescale* [12]. Além disso, nesta escala de agregação o tráfego é aproximadamente gaussiano, portanto, compatível com o modelo de entrada de tráfego proposto, que também é gaussiano (ver Fig. 1 para o histograma do traço dec-pkt-1.tcp).

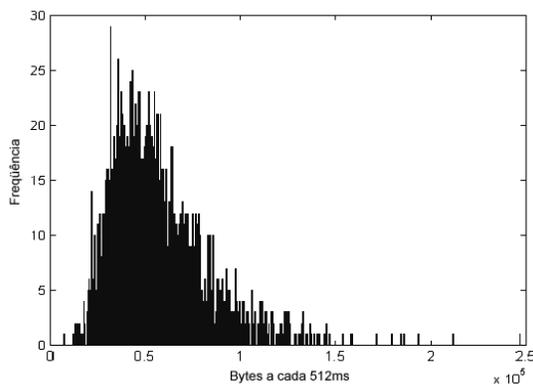


Fig. 1. Histograma do traço dec-pkt-1.tcp agregado em 512ms

A Figura 2 mostra as primeiras 2048 amostras do traço dec-pkt-1.tcp em uma escala de agregação de 512ms (esquerda), e uma realização da síntese do modelo de tráfego proposto (direita) cujos os valores da média e da variância foram obtidos a partir do tráfego real. O método utilizado para a síntese do mBm foi o *Random Midpoints Displacement* [6]. Como pode ser notado na Fig. 3, uma pequena parcela dos dados sintetizados fica abaixo de zero. Assim como para o tráfego Browniano fracionário, a característica gaussiana também limita o uso do método proposto para pequenas escalas de tempo, uma vez que nestas escalas os dados apresentam proporcionalmente baixas médias e elevadas variâncias, conduzindo a um grande número de valores negativos. Uma vez que a característica multifractal manifesta-se principalmente em pequenas escalas de tempo [1][2][3], espera-se que o modelo proposto atinja melhores desempenhos em escalas onde exista um compromisso entre a distribuição gaussiana e a característica multifractal.

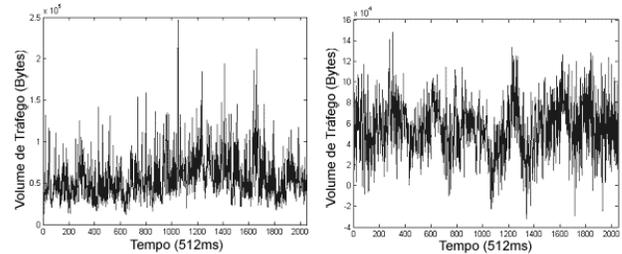


Fig. 2. – Volume de Tráfego: (esquerda) traço dec-pkt-1.tcp, (direita) síntese do modelo de tráfego de entrada proposto.

Objetivando verificar a capacidade do método proposto em capturar as propriedades estatísticas de segunda ordem, estimou-se o parâmetro de Hurst dos dois traços TCP/IP analisados neste trabalho, seguindo-se a comparação dos mesmos com os parâmetros de Hurst obtidos para o tráfego sintetizado através do modelo proposto. A Tabela I mostra os valores dos parâmetros de Hurst estimados tanto para os traços de tráfego real como para os traços sintetizados, obtidos através do robusto estimador Abry-Veitch [11].

Como pode ser observado na Tabela I, o parâmetro de Hurst estimado para os traços sintetizados é ligeiramente diferente do obtido para os traços de tráfego real. Sendo [0,843 0,917] e [0,685 0,760] os intervalos de confiança de 95% para a estimação do parâmetro de Hurst da síntese dos traços dec-pkt-1.tcp e dec-pkt-2.tcp respectivamente, pode-se concluir que para estes dois traços de tráfego, o método proposto pôde capturar sua característica de dependência de longa duração.

O segundo teste realizado busca avaliar se o modelo proposto é capaz de capturar a regularidade Hölder presente nos dados originais. Com este objetivo, calculou-se o espectro multifractal de Legendre [2], obtido através da aplicação da transformada de Legendre à função de partição dos traços em questão. A Figura 3 mostra o espectro multifractal de Legendre para o traço dec-pkt-1.tcp (linha sólida) e o espectro de sua síntese através do modelo proposto (linha tracejada). Embora não mostrado, um comportamento similar ao da Figura 3 foi observado para o traço dec-pkt-2.tcp e sua respectiva síntese através do modelo proposto. Como pode ser observado na Figura 3, para quase toda a faixa de α 's os dois espectros apresentam-se bastante próximos.

Os dois testes realizados apresentam uma clara indicação que o modelo proposto é capaz de capturar tanto as características de dependência longa duração quanto a regularidade Hölder presente nos traços de tráfego analisados.

TABELA I

PARÂMETRO DE HURST ESTIMADO PARA TRAÇOS REAIS E SINTETIZADOS

	dec-pkt-1.tcp		dec-pkt-2.tcp	
	traço real	traço sintético	traço real	traço sintético
Estimador Abry-Veitch	0.835	0.870	0.722	0.757

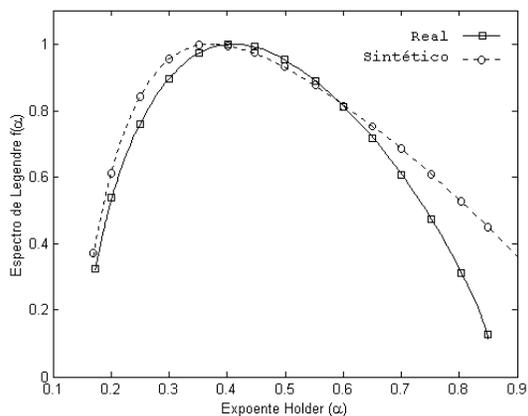


Fig. 3 – Espectro Multifractal de Legendre: (linha sólida) traço de tráfego dec-pkt-1.tcp, (linha tracejada) síntese do traço dec-pkt-1.tcp através do modelo proposto.

Uma vez que o modelo proposto se mostrou capaz de representar as características estatísticas do tráfego, e dado a existência de expressões analíticas para a correlação do processo mBm, propõe-se na próxima seção um novo preditor de tráfego obtido através da utilização de filtros de Wiener.

V. PREDIÇÃO DE TRÁFEGO

O modelo de tráfego proposto é um processo acumulado monotonicamente crescente. Uma vez que o objeto de interesse da predição é o tráfego original e não seu processo acumulado, deve-se voltar as atenções para o processo de incrementos de modelo de tráfego proposto. Como já mencionado anteriormente, geralmente o processo de incrementos do mBm é não estacionário, com exceção para o caso onde se tenha uma função $H(t)$ constante. Portanto, a predição desse tipo de processo requer versões adaptativas de filtros, de forma que o mesmo acompanhe as variações estatísticas que ocorrem quando se opera neste tipo de ambiente.

Seja X um processo mBm e Y seu processo de incrementos. Por definição:

$$cor_Y(t,s) = cor_X(t+1,s+1) - cor_X(t+1,s) - cor_X(t,s+1) + cor_X(t,s) \quad (6)$$

Uma vez que a auto-correlação para processos mBm é dada por (4), utilizando (6) tem-se o valor da auto-correlação do processo de incrementos do mBm.

Propõe-se como preditor linear de k -passos o seguinte preditor Wiener com $p+1$ coeficientes [9]:

$$\hat{x}(n+k) = \sum_{j=0}^p c^o(j)x(n-j) \quad (7)$$

onde:

$\hat{x}(n+k)$ - valor predito de x k -passos a frente;

$x(n-j)$ - j -ésimo valor passado de x ;

c^o - coeficientes ótimos do filtro que minimizam o erro quadrático médio de predição.

É sabido que os coeficientes para o preditor Wiener de k -passos são dados por

$$c^o = R^{-1}r \quad (8)$$

com

$$c^o = [c(0), c(1), \dots, c(p)]^T \quad (9)$$

$$r = [R_x(k), R_x(k+1), \dots, R_x(k+p)]^T \quad (10)$$

$$R = \begin{bmatrix} R_x(0) & R_x(1) & \dots & R_x(p) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(p) & R_x(p-1) & \dots & R_x(0) \end{bmatrix} \quad (11)$$

onde

c^o - vetor de coeficientes ótimos do filtro;

R - matriz de correlação onde os elementos consistem nos valores quadráticos médios das entradas do filtro, assim como as correlações entre essas entradas;

r - vetor de correlação cruzada onde os elementos consistem nas correlações entre o valor predito k -passos a frente e os valores nas entradas do filtro.

No preditor adaptativo proposto, a cada instante, os valores $R_x(\cdot)$ são dados por cor_y (6) onde cor_x são dados por (4).

O erro quadrático médio de predição é dado por

$$\sigma_e^2 = E[(x(n+k) - \hat{x}(n+k))^2] \quad (12)$$

Com o intuito de quantificar a qualidade da predição, adotou-se uma medida de desempenho baseada no erro quadrático médio de predição, normalizado pela variância do processo predito. O erro quadrático médio normalizado (EQMN) e dado então por

$$EQMN = \frac{E[(x(n+k) - \hat{x}(n+k))^2]}{\sigma^2} \quad (13)$$

onde σ^2 é a variância da seqüência $x(n)$.

Com o objetivo de se avaliar o preditor proposto, foram realizadas simulações com os dois traços de tráfego reais mencionados nas seções anteriores. A Figura 4 mostra, para um filtro preditor de 1-passo e com 10 entradas, as séries real e predita compreendidas no intervalo entre as amostras 650 e 700 do traço dec-pkt-1.tcp agregado em uma escala de 512ms.

Utilizando-se um filtro com 10 entradas, são exibidos na Tabela II os valores dos erros quadráticos médios normalizados para a predição de 1-passo dos traços de tráfego analisados neste trabalho.

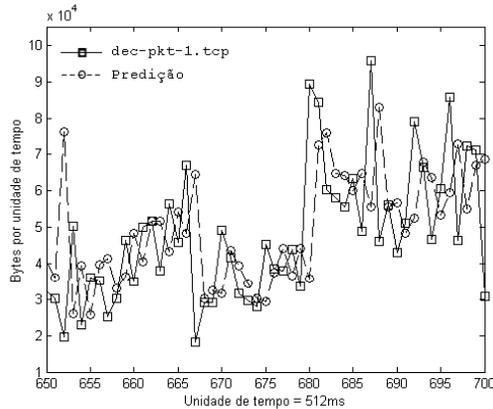


Fig. 4. Predição para amostras do traço dec-pkt-1.tcp: (linha sólida) traço dec-pkt-1.tcp. (linha tracejada) predição de 1-passo.

TABELA II

EQMN OBTIDO PARA OS TRAÇOS DE TRÁFEGO DEC
(ABORDAGEM MULTIFRACTAL)

	dec-pkt-1.tcp	dec-pkt-2.tcp
EQMN	0.6878	0.5635

A fim de comparar os resultados obtidos pelo preditor proposto, um outro preditor foi implementado. Baseado em [10], este preditor utiliza uma abordagem monofractal onde o tráfego é modelado através do fBm, ou seja, considerando-se o modelo tráfego Browniano fracionário. O parâmetro de Hurst empregado neste preditor também foi estimado através do método proposto Abry-Veitch [11]. Tabela III exhibe os valores de EQMN obtidos através da abordagem monofractal.

TABELA III

EQMN OBTIDO PARA OS TRAÇOS DE TRÁFEGO DEC
(ABORDAGEM MONOFRACTAL)

	dec-pkt-1.tcp	dec-pkt-2.tcp
	H=0.835	H=0.722
EQMN	0.7200	0.6128

Comparando os resultados apresentados nas Tabelas II e III observa-se que o preditor baseado nas características multifractais melhora os valores dos EQMN's. O modelo de tráfego proposto consegue modelar melhor as características do tráfego, e uma vez que o preditor possui uma construção adaptativa baseada nesta refinada modelagem, obtém-se EQMN's menores. Em [13] o preditor proposto é aplicado a um esquema de alocação dinâmica de taxa.

VI. CONCLUSÃO

Este trabalho apresentou um novo modelamento de tráfego, estendendo o clássico modelo tráfego Browniano fracionário proposto por Norros [5]. Substitui-se o processo fBm usado no tráfego Browniano fracionário pelo ainda mais completo movimento Browniano multifracionário. O modelo proposto mostrou-se capaz de capturar tanto as correlações de longa duração quando a regularidade Hölder das séries de tráfego reais analisadas. Este trabalho também propõe um novo preditor de tráfego baseado nas características multifractais do tráfego. Um preditor Wiener de k -passos foi projetado assumindo que o tráfego apresenta a mesma estrutura de correlação do modelo mBm. Uma investigação experimental mostrou a eficácia do modelo e do preditor proposto em acompanhar a alta variabilidade dos tráfegos reais analisados.

REFERENCIAS

- [1] A. Gilbert, A. Friedman e W. Willinger, "Data Networks as cascades: Explaining the multifractal nature of Internet WAN traffic", in *Proceedings of ACM Sigcomm*, pp.42-55, 1998.
- [2] R. H. Riedi e J. Levy Véhel, "TCP Traffic is multifractal: A numerical study", *INRIA Research Report No. 3129*, 1997.
- [3] A. Erramilli, O. Narayan, A. Neidhart e I. Sanjeev. "Performance impacts of multiscaling in area TCP/IP traffic". In *Proceedings of INFOCOM 2000*, pp.352-359, 2000.
- [4] W. Leland, M. Taqqu, W. Willinger, e D. Wilson, "On the self-similar nature of Ethernet Traffic (extended version)". *IEEE/ACM Transactions on Networking*, pp.1-15, Fevereiro 1994.
- [5] L. Norros, "A storage model with self-similar input". *Queueing Systems*, v.16, pp.387-396, 1994.
- [6] Peltier, R. e J. L. Véhel, "Multifractional Brownian motion: definition and preliminary results". *INRIA Research Report No. 2645*, 1995.
- [7] A. Ayache, S. C. e Véhel, J. L. "The covariance structure of multifractional Brownian motion, with application to long range dependence", in *Proceedings of ICASSP'00*, 2000.
- [8] A. Benassi, S. Cohen e J. Istas, "Identifying the multifractional function of a Gaussian process", *Stat. and Prob. Letter* 39, 377-345, 1997.
- [9] Haykin, Simon S., "Modern filters", Macmillan Publishing Company, 1989.
- [10] G. A. Hirchoren, D. S. Arantes. "Predictors for the Discrete Time Fractional Gaussian Process", in *Proceedings of SBT/IEEE International Telecommunications Symposium ITS '98*, v.1, pp.49-53, 1998.
- [11] P. Abry e D. Veitch, "Wavelet Analysis of Long Range Dependent Traffic". *IEEE Transactions on Information Theory*, v.4, n.1 pp.1-15, 1998.
- [12] S. Molnár e T. D. Dang, "Scaling analysis of IP traffic components" ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management, Monterey, CA, EUA, Setembro 2000.
- [13] G. R. Bianchi, F. H. T. Vieira e L. L. Ling, "Predictive Dynamic Bandwidth Allocation based on Multifractal Traffic Characteristic", in *Proceedings of SAPIR/ICT 2004*. Aceito para publicação.