Uma Proposta de Modelagem Matemática para Tratamento de Tráfego com Característica Auto–Similar

Oscavo G. Prata Jr.[1]**, Anilton S. Garcia** [2]

[1]Departamento de Física e Matemática – INATEL
Av. João de Camargo, 510 – 37540-000 Santa Rita do Sapucaí, MG

[2]Departamento de Informática – Universidade Federal do Espírito Santo
Centro Tecnológico – Caixa Postal 01-9044 – 29-60-970 Vitória, ES

`oscavo@inatel.br, anilton@inf.ufes.br`

Abstract. *Few years ago, several studies were done to determine the traffic behaviour in many kinds of nets – ethernet, ISDN, CCSN, ATM, etc.. From these studies on, many models appeared, among them some models stand out: FGn, FBm, Pareto and ON/OFF. This work has the object of presenting a new contribuition to Pareto model´s whenever it´s used to model self–similar traffic. It´s proved that Pareto´s distribution can generate a self–similar behaviour, with the degree of self–similarity give by the Hurst´s parameter, if the variables of distribution shape and location will be appropriately chosen.*

Resumo. *Há poucos anos foram realizados vários estudos para se determinar o comportamento do tráfego em vários tipos de redes – ethernet, ISDN, CCSN, ATM, etc.. A partir desses estudos, surgiram inúmeros modelos entre os quais se destacam: FGn, FBm, Pareto e ON/OFF. Este trabalho apresenta uma nova contribuição ao modelo de Pareto quando empregado em tráfego auto–similar. Demonstra-se que a distribuição de Pareto pode gerar um comportamento auto–similar com o grau dado pelo parâmetro de Hurst, se as variáveis de* ***forma*** *e* ***localização*** *da distribuição forem devidamente escolhidas.*

1. Introdução

Para representar o comportamento de determinada rede, ou mesmo de um de seus elemento, um dos pontos cruciais é desenvolver um modelo matemático capaz de caracterizar o tráfego gerado pelas diversas fontes. Outro aspecto de extrema relevância é conhecer as suas características estatísticas. O tráfego nas redes de pacotes pode apresentar um comportamento auto-similar, diferente do tráfego das redes de telefonia convencionais [Paxson and Floyd, 1995]. A partir do trabalho apresentado por Leland *et. al.* [Leland et al., 1994] o tráfego vem sendo classificado de duas formas: tráfego com dependência de longa duração e tráfego com dependência de curta duração.

Neste trabalho busca-se apresentar uma nova contribuição para a distribuição de Pareto [Gordon, 1995], quando for utilizada para modelar o tempo entre chegadas de entidades (pacotes, células, etc..) em uma rede. Será demonstrado que para a distribuição de Pareto modelar o comportamento auto–similar, com grau dado pelo parâmetro de Hurst, serão consideradas tanto a variável de *forma* como também a variável de *localização*. Para isso, o artigo foi organizado da seguinte forma: a primeira parte traz a proposição do modelo matemático e a inclusão do parâmetro de Hurst e sua validação; a segunda parte traz os resultados obtidos, a partir de simulações computacionais acompanhada das conclusões e comentários pertinentes.

2. Proposição do Modelo Matemático de Tráfego

A nova característica do tráfego, conhecida como dependência de longa duração ou de auto-similaridade, começou a ganhar destaque após as publicações de alguns trabalhos na primeira metade da década de 90 [Leland et al., 1994]. A auto–similaridade do tráfego manifesta-se de diversas formas [Jarferman et al., 1996], no entanto, neste trabalho é utilizado utilizado o conceito apresentado por Mandelbrot [Mandelbrot, 1982] e explorando por Leland [Leland et al., 1994], Fernandez [Fernandez et al., 1998], Cappe [Cappe et al., 2002], e outros para justificar distribuições de probabilidade classificadas de cauda pesada como uma ferramenta de modelagem de tráfego auto-similar.

A distribuição de Pareto pode ser classificada como de cauda pesada se o seu parâmetro de *forma* for adequadamente escolhido [Gordon, 1995]. Porém, não é possível definir o grau de auto–similaridade dada pelo parâmetro de Hurst se a variável de localização não for adequadamente especificada. Partindo desta afirmativa, demonstra-se que a variável de localização deve ser considerada para que a distribuição seja capaz de modelar a auto–similaridade presente no tráfego das redes de pacotes.

Mandelbrot [Mandelbrot, 1982] estabelece que um modelo probabilístico de cauda pesada é caracterizado por uma variável aleatória T, com a função distribuição de probabilidade complementar da forma $P[|T| \geq t] \approx \frac{L(t)}{t^\alpha}$, com variância infinita, $0 < \alpha < 2$ e $L(t)$ sendo uma função que varia lentamente quando $t \rightarrow \infty$.

A distribuição de Pareto,

$$f_T(t) = \begin{cases} \dfrac{\alpha}{\beta}\left(1 + \dfrac{t}{\beta}\right)^{-\alpha-1} & t, \beta \in \mathbb{R}_*^+, \alpha \in (0,2) \\ 0 & t < 0 \end{cases}, \qquad (1)$$

pode apresentar as características equivalentes à proposta por Mandelbrot se o parâmetro α for adequadamente escolhido, como pode ser visto nos trabalhos de Gordon [Gordon, 1995] e Fernandez [Fernandez et al., 1998].

Em face do exposto, demostra-se que a distribuição de Pareto pode ser considerada um modelo auto–similar com o grau determinado pelo parâmetro de Hurst, se a variável de localização (β) assumir uma faixa de valores restritos, a ser obtida por um método descrito neste trabalho. De acordo com [Gordon, 1995] a variável α da distribuição de Pareto é uma função cujo argumento é o parâmetro de Hurst (H), ou seja, $\alpha(H) = 3 - 2H$ com $0.5 < H < 1$. Para identificar a faixa de valores válidos para a variável de localização da distribuição de Pareto, seja a função de autocorrelação de um processo estocástico FGn (Ruído Gaussiano Fracional), que modela o comportamento auto–similar do tráfego [Hlavacs et al., 1999]. A função de auto–correlação desse processo estocástico é

$$\rho(\tau) = \frac{1}{2}\left(|\tau+1|^{2H} - 2|\tau|^{2H} + |\tau-1|^{2H}\right),\ \tau = 1, 2, 3, \ldots \qquad (2)$$

Além disso, será considerada a estatística de segunda ordem de um processo estocástico estacionário dada por

$$R_T(\tau) = \frac{\beta^{\alpha(H)+1}}{(1-\alpha(H))}\left[\frac{\alpha(H)(\tau+\beta)-(1-\alpha(H))\beta}{(1-\alpha(H))(\tau+\beta)^{\alpha(H)}}\right], \tag{3}$$

onde $0,5 < H < 1$ e $\beta > 0$, que é a função de autocorrelação de um processo estocástico estacionário dado pela distribuição de Pareto, apresentada em (1). Se for feita uma comparação gráfica entre (2) e (3) para um mesmo valor de H e diferentes valores de β nota-se que β também deve ser considerado para que a distribuição de probabilidade apresentada em (1) seja capaz de modelar tráfego auto–similar, onde o grau é dado parâmetro de Hurst.

Para determinar o valor de β na distribuição de Pareto em (3), é tomado o erro médio quadrático entre (2) e (3), com o objetivo de identificar o valor de β em função de H. A partir dessa análise é possível demonstrar que β e H apresentam o seguinte relacionamento,

$$\beta(H) \cong -3.3181H^2 + 4.4567H - 1.141. \tag{4}$$

Substituindo-se (4) e os valores válidos para α em (3) esse processo estocástico apresenta um decaimento hiperbólico que é uma característica de dependência de longa duração, e assim pode-se validar o modelo teoricamente.

3. Simulação do Modelo de Pareto com Parâmetro de Hurst

Na seção anterior, foi apresentada a proposição do modelo matemático para descrever o comportamento auto-similar do tráfego, assim como a sua validação teórica. A partir (1) com a variável de localização descrita por (4) e a de forma por $\alpha = 3-2H$, foi construído uma fonte de tráfego que gera pacotes (*bits/seg*) com tamanho dado por uma distribuição exponencial e o tempo entre a emissão desses pacotes dado pela distribuição de Pareto com parametro de Hurst, descrita na seção anterior.

A verificação da auto-similaridade da fonte simulada foi feita, utilizando-se o teste da Variância–Tempo [Hlavacs et al., 1999, Jarferman et al., 1996], e os resultados são apresentados na Figura-1.

4. Conclusões

Neste trabalho buscou-se investigar qual deve ser o relacionamento entre as variáveis de *forma* e *localização* da distribuição de Pareto com o parâmetro de Hurst, para que essa distribuição possa ser empregada na modelagem de tráfego auto–similar com maior exatidão. Através do desenvolvimento matemático feito, demonstrou-se que o modelo de Pareto com o parâmetro de Hurst descreve de forma apropriada a auto–similaridade do tráfego presente nas redes de pacotes, o que caracteriza um ganho adicional a distribuição de Pareto como um modelo auto-similar, apresentada na literatura mencionada. Outro ganho em relação aos modelos mencionados, é a necessidade de se especificar somente o parâmetro de Hurst para utilização da distribuição de Pareto, o que não se aplica por exemplo ao modelo FGn onde é necessário especificar o parâmetro de Hurst bem como a média e a variância [Neto et al., 2000].

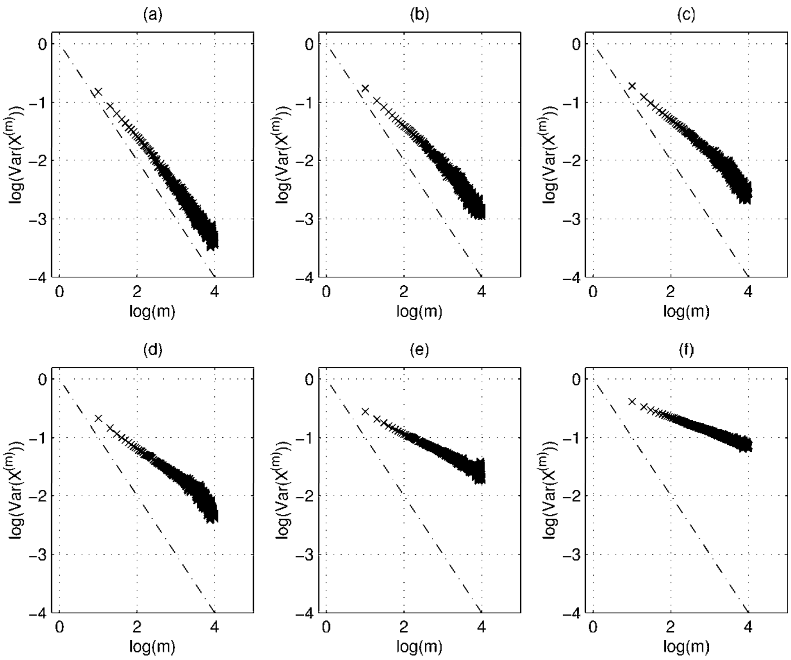

Figura 1: Análise Variância-Tempo. A linha pontilhada indica H=0.5 e as figuras de (a)–(f) apresentam o comportamento da fonte para os valores de H iguais a: 0,55, 0,65, 0,70, 0,75, 0,85 e 0,95, respectivamente.

Referências

Cappe, E. M., Pesquet, J. C., Petropolu, A., and Yang, X. (2002). Long-range dependence and heavy tail modeling for teletraffic data. *IEEE Signal Processing Magazine*, 19(3):14–27.

Fernandez, J. M., Huebener, F., and Liu, D. (1998). Queueing performance comparison for traffic models for internet traffic. In *IEEE GlobeCom*, Sydney.

Gordon, J. (1995). Pareto process as a model of self-similar packet traffic. In *Proc. IEEE Globecom'95*, pages 2232–2236.

Hlavacs, H., Kotsis, G., and Steinkellner, C. (1999). Traffic source modeling. Technical Report TR-991001, Institute for Appl. Comp. Science and Inf. Systems, University of Vienna.

Jarferman, D. L., Melamed, B., and Willinger, W. (1996). *Stochastic Modeling of Traffic Processes*. CRC Press, in frontiers in queueing: models, methods and problems edition.

Leland, W., Taqqu, M., Willinger, W., and Wilson, D. (1994). On the self-similarity nature of ethernet traffic (extended version). *IEEE/ACM Trans. on Networking*, 2(1):1–15.

Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. W. H. Freeman, San Francisco.

Neto, E. L. A., Alberti, A. M., and Mendes, L. S. (2000). Aplicação de modelos multifractais na simulação de redes ATM. *XVIII Simpósio Brasileiro de Telecomunicações*.

Paxson, V. and Floyd, S. (1995). Wide-area traffic: The failure of poisson modeling. *IEEE/ACM Transactions on Networking*, 3(3).