

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

OSCAVO GONZAGA PRATA JÚNIOR

Dissertação apresentada ao Instituto Nacional de Telecomunicações, como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica.

Orientador: PROF. DR. ANILTON SALLES GARCIA

Santa Rita do Sapucaí
2003

Dissertação defendida e aprovada em 29/07/2003, pela comissão julgadora:

Prof. Dr. Anilton Salles Garcia (Orientador) - *DTE-INATEL*

Prof. Dr. Shusaburo Motoyama (membro externo) - *DT-FEEC-UNICAMP*

Prof. Dr. Carlos Alberto Ynoguti (membro interno) - *DCOM-INATEL*

Coordenador do Curso de Mestrado

A DEUS que me deu tudo
que precisei para realizar
esse trabalho. E a minha
amada esposa, que
grandemente me auxiliou.

Agradecimentos

Ao Professor-orientador Anilton Salles Garcia.

Em especial, ao Professor Carlos Alberto Ynoguti, pela sua atenção e paciência, que contribuiu de forma singular para a realização desse trabalho.

À minha esposa Paula, que desde o início dessa jornada não ofereceu resistência à realização de um sonho.

Ao INATEL pela oportunidade oferecida.

E acima de tudo, a Jesus Cristo, o DEUS vivo e senhor da vida.

Índice

Lista de Figuras	vi
Lista de Tabelas	viii
Lista de Abreviaturas e Siglas	ix
Lista de Símbolos	xi
1 Introdução	1
1.1 Prefácio	1
1.2 Histórico	1
1.3 Tratamento Estatístico	2
1.4 Proposta do Trabalho	3
2 Caracterização e Modelagem do Tráfego nas Redes de Pacotes	4
2.1 Introdução	4
2.2 Auto-similaridade	5
2.2.1 A matemática da auto-similaridade	6
2.2.2 Identificando a característica de auto-similaridade em processos estocásticos	7
2.3 Modelos de tráfego	9
2.3.1 Modelo para redes Ethernet	9
2.3.2 Modelos para redes WAN	10
2.3.3 Modelos para redes ATM	10
2.3.4 Modelos para WWW	11
2.4 Conclusões	11
3 Modelagem do Tráfego com Característica Auto-Similar	13
3.1 Introdução	13
3.2 Proposição do Modelo de Tráfego	14

3.2.1	Modelo Matemático do Tráfego	14
3.2.2	Parâmetro de Hurst e a Distribuição de Pareto	15
3.2.3	Validação Teórica do Modelo de Tráfego	21
3.3	Conclusões	22
4	Modelo de fila $G/M/1$ e a Distribuição de Pareto com Parâmetro de Hurst	23
4.1	Introdução	23
4.2	Análise Matemática do Modelo $G/M/1$	23
4.3	Simulação: Modelo $G/M/1$ vs. $M/M/1$	26
4.3.1	Modelo de Simulação	26
4.3.2	Resultados da Simulação	28
4.4	Conclusões	29
5	Conclusão	31
5.1	Contribuições da Dissertação	31
5.2	Proposição para Trabalhos Futuros	32
	Anexos	33
	A Distribuição de Pareto	33
	B Função de Autocorrelação de um Processo Estocástico Estacionário com fdp de Pareto	36
	C Transformada de Laplace da fdp de Pareto	38
	Referências Bibliográficas	41

Lista de Figuras

2.1	Tráfego Ethernet observado em cinco diferentes escalas de tempo. (Figura extraída do artigo “On the Self-Similar Nature of Ethernet Traffic”)	5
2.2	Ilustração do conceito de auto-similaridade do tráfego nas redes de pacotes. Pode-se observar nas diferentes escalas de tempo, (a)–(c), que o comportamento do tráfego de entrada da rede de pacotes do INATEL, apresenta ligeira semelhança.	6
2.3	Em (a) pode-se ver a análise R/S do tráfego de saída da rede de pacotes do INATEL. Em (b) é apresentada a análise R/S do tráfego de entrada da rede de pacotes do INATEL.	9
3.1	Gráfico da variação do parâmetro α , da distribuição de Pareto, em relação aos valores do parâmetro de Hurst – H. Em (a) observa-se o plano formado pela intersecção das faixas de valores dos parâmetros α e H. Em (b) é apresentado o relacionamento entre os valores de α e H.	16
3.2	Comparação entre a função de autocorrelação do modelo FGn, com $H=0.75$, e o função de autocorrelação do Modelo de Pareto considerando apenas a variável α como sendo função do parâmetro H e $\beta > 0$	18
3.3	Relação entre a variável β e o parâmetro de Hurst. Em (a) vemos alguns valores de β em função de H. Em (b) é apresentada a curva obtida através de interpolação dos pontos apresentados em (a). . .	19
3.4	Função densidade de probabilidade de Pareto com parâmetro de Hurst.	20
4.1	Ilustração gráfica da solução da equação (4.4) para diferentes valores de H.	26
4.2	Modelo de simulação empregado para avaliar o comportamento de uma fila $M/M/1$ e $G/M/1$ considerando a mesma média para o tempo entre chegadas das entidades.	27

4.3	Comparação gráfica entre os modelos $G/M/1$ e $M/M/1$ levando-se em consideração o comprimento médio da fila.	29
-----	---	----

Lista de Tabelas

2.1	Exemplos de aplicações suportadas pela tecnologia ATM, classificados de acordo com os tipos de classes de serviços	11
4.1	Equações válidas para um sistema de filas do tipo $G/M/1$ para σ determinado a partir da solução da transformada Laplace–Stieltjes para a distribuição do tempo entre chegadas.	24
4.2	Tempo médio, dado em segundos, gasto na fila $G/M/1$ para diferentes fatores de utilização, considerando-se que a distribuição do tempo entre chegadas é a distribuição de Pareto com parâmetro H	28
4.3	Tempo médio gasto na fila $M/M/1$ para diferentes fatores de utilização, considerando-se que a distribuição do tempo entre chegadas é exponencial com média dada pela distribuição de Pareto com parâmetro H	28

Lista de Abreviaturas e Siglas

AAL	<i>ATM Adaptation Layer</i> - Camada de adaptação ATM
ATM	<i>Asynchronous Transfer Mode</i> - Modo de transferência assíncrono
FBm	<i>Fractional Brownian Motion</i> - Movimento Browniano Fracional
<i>fdc</i>	<i>Função distribuição cumulativa</i>
<i>fdc⁻¹</i>	<i>Função distribuição cumulativa complementar</i>
<i>fdp</i>	<i>Função densidade de probabilidade</i>
FGn	<i>Fractional Gaussian Noise</i> - Ruído Gaussiano Fracional
<i>G/M/1</i>	<i>Sistema de fila com um servidor onde o tempo de serviço e dado por um distribuição exponencial e o intervalo entre chegada das entidades segue uma distribuição genérica.</i>
IDC	<i>Index of Dispersion for Counts</i> - Índice de Dispersão
<i>M/G/1</i>	<i>Sistema de fila com um servidor onde o tempo de serviço e dado por um distribuição genérica e o intervalo entre chegada das entidades segue uma distribuição exponencial.</i>

M/M/1	<i>Sistema de fila com um servidor onde o tempo de serviço e dado por um distribuição exponencial e o intervalo entre chegada das entidades também segue uma distribuição exponencial.</i>
MMPP	<i>Markov Modulated Poisson Process - Processos de Poisson cuja taxa de chegada é uma variável aleatória que é modulada por uma cadeia de Markov de tempo-contínuo</i>
OPNET[®]	<i>Optimun Network Performance - Ferramente de simulação para redes de telecomunicações.</i>
R/S	<i>Rescaled Adjust Range or R/S-statistic - Estatística de Amplitude Reescalonada Ajustada</i>
WAN	<i>Wide-area network - Rede metropolitana</i>
WWW	<i>World Wide Web - Rede mundial de computadores</i>
VBR	<i>Variable Bit Rate - Taxa de bit variável</i>

Lista de Símbolos

X_t	Processo estocástico
n	Número de observações
H	Parâmetro de Hurst
$\rho(\cdot)$	Função de autocorrelação do processo FGn
τ	Atraso da função de autocorrelação
$\mathcal{S}(\cdot)$	Densidade espectral
$Y(\cdot)$	Média da variável aleatória X_i
$L(\cdot)$	Função que varia lentamente quando seu argumento tende para infinito
$\alpha(\cdot)$	Variável da distribuição de Pareto expressa em função do parâmetro de Hurst
$\beta(\cdot)$	Variável da distribuição de Pareto expressa em função do parâmetro de Hurst
$f(\cdot)$	Função arbitrária
$g(\cdot)$	Função arbitrária
$r(\cdot)$	Função de autocorrelação de um processo estocástico auto-similar
α	Variável da distribuição de Pareto conhecida como <i>shape</i>

β	Variável da distribuição de Pareto conhecida como <i>location</i>
μ ou μ_1	Média da distribuição de Pareto
μ_2	Segundo momento da distribuição de Pareto
σ_T^2	Segundo momento central da distribuição de Pareto – Variância
$F_T(\cdot)$	Função de distribuição cumulativa de Pareto
$F_T^{-1}(\cdot)$	Função de distribuição cumulativa complementar de Pareto
$R_T(\cdot)$	Função de autocorrelação de um processo estocástico independente e identicamente distribuído com fdp de Pareto.
$f(\cdot)$	Função densidade de probabilidade da variável aleatória T
$\Gamma(\cdot)$	Função gama
$\Gamma(\cdot, \cdot)$	Função gama incompleta
$\mathcal{L}\{f(\cdot)\}$	Transformada de Laplace de $f(\cdot)$
$\varepsilon(\cdot)$	Erro médio quadrático
s	Variável complexa de Laplace

Resumo

PRATA JR., O.G. Uma Proposta de Modelagem Matemática para Tratamento de Tráfego com Característica Auto-Similar. Santa Rita do Sapucaí, 2003. Instituto Nacional de Telecomunicações.

Os modelos de tráfego aplicados no passado para o dimensionamento, análise de desempenho e otimização das redes de telefonia convencional, tornaram-se ineficientes devido à mudança na natureza estatística do tráfego das redes de dados, fomentando assim, na comunidade acadêmica/científica, uma constante busca por modelos que sejam capazes de tratar essa nova característica estatística do tráfego. Essa dissertação apresenta uma proposta de modelo matemático capaz de representar a característica auto-similar do tráfego das atuais redes de telecomunicações. O modelo proposto utiliza a distribuição de Pareto com o parâmetro de Hurst para caracterizar os tempos entre chegadas de pacotes na rede. Como consequência do desenvolvimento do modelo é mostrada a importância da variável de localização da distribuição de Pareto para que essa possa modelar tráfego auto-similar. É feita também, uma análise utilizando-se o modelo desenvolvido dentro da teoria clássica de filas.

Palavras-chave: Modelagem de tráfego, auto-similaridade, $G/M/1$, $M/M/1$, caracterização do tráfego, desempenho, distribuição de Pareto.

Abstract

The traffic models applied in the past for design and performance analysis of the telephony networks became inefficient due to the change in the statistical nature of data traffic. This instigate the scientific/academic community to looks for models that are capable to treat that new characteristic statistics of the traffic. This document presents a proposal of mathematical model capable to represent the self-similar feature of the traffic of the current networks. The intended model uses the Pareto distribution with Hurst parameter to characterize the package interarrival time. Consequently, it is presented the relevant for variable of location for Pareto distribution modelling self-similar traffic. We also present a study using the proposed model for classic queue theory .

Capítulo 1

Introdução

1.1 Prefácio

As antigas redes de telecomunicações eram projetadas para acomodar um único tipo de tráfego – o de voz. Em decorrência disso, inúmeros estudos foram realizados com o intuito de determinar métodos para o dimensionamento, análise e otimização dessas redes [1].

No entanto, o cenário tecnológico vem sofrendo constantes alterações e consequentemente as redes de telecomunicações também são influenciadas por essas mudanças. Em decorrência disso, as modernas redes de telecomunicações, não mais são utilizadas unicamente para acomodar o tráfego de voz. Estas redes vêm sendo utilizadas para o transporte dos mais diversos tipos de informação, ou seja, voz, vídeo, dados, multimídia, etc..

Com essas novas aplicações as redes não são mais estudadas utilizando-se os modelos desenvolvidos no passado [1, 2, 3]. Essa constante evolução, traz um grande desafio aos estudiosos desse assunto, tendo em vista que não existe ainda um modelo concreto e definitivo para tratar o tráfego das atuais redes, fomentando assim a constante busca de modelos que sejam cada vez mais adequados para a modelagem, análise de desempenho e otimização dessas redes.

1.2 Histórico

Os primeiros estudos de tráfego [1] começaram a ser realizados no início do século passado por Malcolm C. Rorty. Em 1909 o engenheiro A. K. Erlang publicou dois teoremas conhecidos até hoje que são Erlang B e Erlang C [1].

Nos anos 30 três modelos foram aceitos na comunidade de comunicação, que foram, o modelo de Poisson, empregado no uso do dimensionamento de facili-

dades, o modelo Erlang B, usado para medir a taxa de bloqueio de requisições em sistemas sem fila de espera e o modelo Erlang C, usado para medir a taxa de bloqueio de requisições em sistemas com fila de espera.

Esse modelos eram aplicados apenas no dimensionamento das rotas diretas. Porém, modelos para o dimensionamento de caminhos secundários ainda não eram considerados. No meado da década de 50, Roger Wilkinson apresenta uma solução para esse problema, todavia sem obter resultados numéricos pela falta de recursos computacionais [1]. No entanto, em 1980, Henry Jacobsen publicou os resultados da solução apresentada por Wilkinson [1].

Em paralelo com o desenvolvimento da telefonia surgiu nos anos 70 a ARPANET [4]. Desde então, o avanço tecnológico contribuiu para que as redes de telecomunicações atingissem o estado dos dias de hoje. Em conjunto com esse desenvolvimento, surgiram também novos problemas no campo do dimensionamento, análise e otimização dessas redes.

Como um dos primeiros estudos apresentados como o objetivo de analisar o desempenho das redes de pacotes pode-se citar o trabalho apresentado por Kleinrock [5], em 1975. No decorrer dos anos, muitos outros trabalhos foram feitos, porém em 1994 surge o trabalho “*On the self-similar nature of Ethernet traffic*” [6] que colocaria os modelos tradicionais, ou seja modelos Markovianos, em condições de questionamento, tendo em vista que o tráfego nas redes de pacotes apresentava um nova natureza estatística, chamada de auto-similar.

Após os trabalhos apresentados por Leland *et. al.* [6], inúmeros outros trabalhos surgiram com intuito de investigar a natureza auto-similar do tráfego das redes de pacotes. Dentre eles podemos citar Beran *et. al.* [7] e Garrett [8] que investigaram a natureza estatística do tráfego de vídeo em redes de pacotes. Em redes WAN/LAN Paxson *et. al.* [9] apresentam resultados da característica estatística do tráfego nestas redes, mostrando que os modelos poissonianos aplicam-se apenas a tráfegos muito específicos como é o caso do tráfego gerado por TELNET e pelo tráfego gerado pelo controle de conexões FTP. Outros trabalhos como o apresentado por Crovella *et. al.* [10, 11] buscam caracterizar o tráfego produzido na WWW.

1.3 Tratamento Estatístico

Algumas distribuições, tais como Poisson, Exponencial, etc., foram e ainda são de grande utilidade no estudo do tráfego. No entanto, modelos de tráfego que usam essas distribuições como base são aplicados apenas em modelagem de tráfego de voz e em alguns casos específicos de redes de pacotes, como é o caso do tráfego TELNET e FTP [9]. Isso devido a nova natureza do tráfego presente nas moder-

nas redes de telecomunicações.

Desde os estudos empíricos apresentados por Leland et. al. [6] os modelos clássicos entraram em cheque. Isso acarretou um aumento significativo pela busca de novos modelos para representar a natureza estatística no tráfego nas redes de pacotes. Como um dos principais modelos podemos citar o FGn [3, 12] e o FBM [13, 2] que são amplamente referenciados na literatura como modelos capazes de representar a auto-similaridade do tráfego presente nas redes de pacotes.

No entanto, devido a complexidade desses modelos inúmeras outras propostas surgem para representar o tráfego com característica auto-similar [14, 15, 16, 17] e outras propostas podem ser vistas na vasta bibliografia apresentada por Willinger, Taqqu e Erramilli [18].

1.4 Proposta do Trabalho

A principal proposta deste trabalho é o desenvolvimento de um modelo matemático capaz de representar o comportamento auto-similar presente no tráfego das redes de pacotes, assim como investigar a aplicação desse modelo utilizando-se a teoria clássica de filas.

Para alcançar o objetivo proposto neste trabalho será apresentado no Capítulo-2, de forma sintetizada, algumas técnicas para analisar o comportamento estatístico do tráfego, como também é colocada a definição de auto-similaridade, que é uma característica estatística evidenciada no tráfego das redes de pacotes. Finalizando esse capítulo são mencionadas diversas propostas para modelagem de tráfego auto-similar considerando diversos tipos de tecnologia de rede. O objetivo desse capítulo é obter uma base teórica para o desenvolvimento dos capítulos posteriores.

O Capítulo-3, traz a proposição do modelo e o seu desenvolvimento matemático bem como sua respectiva validação; ao final desse capítulo são apresentados algumas conclusões a respeito do modelo proposto assim como a sua contribuição. Um importante resultado neste capítulo é a incorporação do parâmetro de Hurst na fdp de Pareto, o que caracteriza uma contribuição para essa distribuição quando usada para modelar tráfego auto-similar.

Como consequência do desenvolvimento feito no Capítulo-3, é apresentada uma análise da fdp de Pareto com parâmetro de Hurst em um sistema de filas do tipo $G/M/1$. Como resultado desse estudo identificou-se que o modelo proposto não admite solução analítica quando o mesmo é usado para modelar o tempo entre chegada das entidades neste tipo de fila. Esse estudo é o assunto do Capítulo-4.

Essa dissertação é concluída com o Capítulo-5, onde são abordados os aspectos de possíveis trabalhos futuros, dando assim continuidade a esse estudo.

Capítulo 2

Caracterização e Modelagem do Tráfego nas Redes de Pacotes

2.1 Introdução

A comunidade científica/acadêmica tem buscado continuamente caracterizar o tráfego das redes atuais, como já foi feito no passado nas redes de telefonia. Para isso dispõe-se atualmente de um grande número de ferramentas (*hardware e software*) para coleta de dados, sendo estas de fundamental importância para se investigar a natureza estatística do tráfego.

Estudos empíricos realizados por Leland *et. al.* [19, 20, 6], no início da última década, revelaram que o tráfego nas redes de pacotes apresenta característica de auto-similaridade. A partir destas pesquisas, inúmeros trabalhos surgiram visando investigar essa nova característica do tráfego, assim como apresentar modelos matemáticos capazes de descrever esse novo comportamento – veja [2, 3, 21, 10, 11, 9, 18, 22, 17, 23, 24, 25, 26, 27, 14, 13, 28, 12]. Desde então, o tráfego das redes vem sendo classificado de duas formas:

tráfego com dependência de longa duração, e

tráfego com dependência de curta duração.

Para compreender melhor as características estatísticas presentes no tráfego das redes de pacotes, este capítulo é dividido em duas seções: a primeira, responsável por apresentar o conceito qualitativo e quantitativo da auto-similaridade; a segunda, responsável por apresentar alguns modelos de tráfego que são capazes de representar essa nova característica presente no tráfego das redes de pacotes. Finalizando este capítulo são apresentadas algumas conclusões.

2.2 Auto-similaridade

A auto-similaridade refere-se à propriedade de um determinado fenômeno manter certas características quando observado em diferentes escalas (tempo, espaço, etc.). Este conceito foi inicialmente apresentado por Kolmogorov* em 1940. Em 1962, Lamperti [29] apresentou o tratamento matemático da auto-similaridade. Três anos depois, em 1965, Mandelbrot [30] introduziu este conceito em sistemas de comunicação, bem como o conceito de ruído gaussiano fracional [31], em 1968.

Visualmente, a auto-similaridade do tráfego manifesta-se como uma invariância em relação à escala de tempo. Por exemplo: se forem construídos gráficos “Pacotes/Unidade de Tempo” vs. “Unidade de Tempo”, para diversas escalas de tempo, pode-se verificar que o tráfego parece o mesmo. Pode-se visualizar esse conceito através da Figura-2.1.

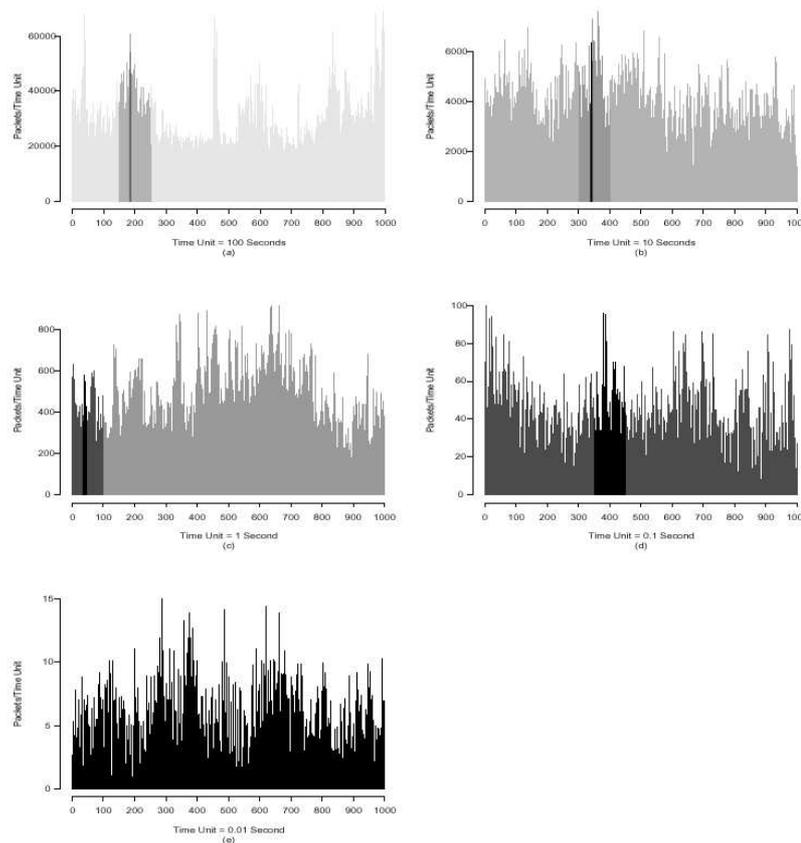


Figura 2.1: Tráfego Ethernet observado em cinco diferentes escalas de tempo. (Figura extraída do artigo “On the Self-Similar Nature of Ethernet Traffic”)

*Andrei Nikolaevich Kolmogorov, nasceu em 25 abril de 1903 em Tambov, Russia, morreu em outubro de 1987 em Moscou. Ele foi um dos mais famosos matemáticos soviéticos e um dos grandes matemáticos do século passado.

Porém, o tráfego nas redes Ethernet não são predominantemente auto-similares como pode ser visto na através da Figura-2.2. Em (a) tem-se o tráfego, no *gateway* da rede do INATEL, observado na escala de cinco minutos, em (b), o tráfego é medido na escala de 30 minutos, e em (c), o tráfego é visto na escala de duas horas.

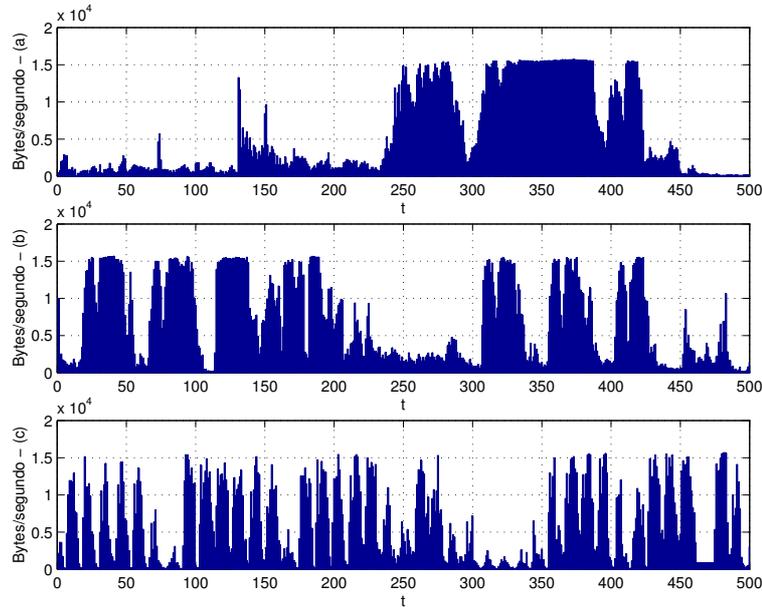


Figura 2.2: *Ilustração do conceito de auto-similaridade do tráfego nas redes de pacotes. Pode-se observar nas diferentes escalas de tempo, (a)–(c), que o comportamento do tráfego de entrada da rede de pacotes do INATEL, apresenta ligeira semelhança.*

Ao contrário do que se pode pensar, a auto-similaridade não se expressa, unicamente, de forma intuitiva. Como será visto mais adiante, a auto-similaridade possui uma definição matemática rigorosa. No entanto, a auto-similaridade pode ser descrita através de poucos parâmetros, onde um deles é o parâmetro de Hurst – H , sendo que $0.5 < H < 1$ e, $H = 0.5$ indica que não existe auto-similaridade e $H = 1$ indica uma perfeita auto-similaridade.

A seguir, é feita uma breve descrição das definições relacionadas a processos estocásticos auto-similares. Finalizando esta seção, são apresentados métodos para determinar se um processo estocástico apresenta características de auto-similaridade ou não. Maiores detalhes sobre este assunto podem ser encontrados em [29, 32, 33] e [34].

2.2.1 A matemática da auto-similaridade

Existem varias referências que trazem a representação matemática da auto-similaridade. Por este motivo não será feita uma exposição rigorosa a respeito desse assunto,

pois estaria fora do escopo deste trabalho.

Os processos auto-similares são definidos da seguinte forma: seja $\{X_t : t = 0, 1, 2, \dots\}$ um processo estocástico com variância $Var(X_t)$, e função de autocorrelação $r(k), k \geq 0$. Sejam $\{X_k^{(m)}\}$ as novas séries obtidas pela divisão proporcional da série original em blocos não sobrepostos de tamanho $m, m = 1, 2, 3, \dots$. Portanto, as novas séries serão descritas da seguinte forma:

$$X_k^{(m)} = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i, k = 1, 2, 3, \dots \quad (2.1)$$

As novas séries $X_k^{(m)}$ têm variância $Var(X_k^{(m)})$ e função de autocorrelação $r^{(m)}(k)$. Um processo X_t é auto-similar com parâmetro H de auto-similaridade se, para todo m positivo, as seguintes seqüências tiverem a mesma distribuição:

$$\{mX_k^{(m)}, k \geq 1\} \triangleq \{m^H X_k, k \geq 1\}, \quad (2.2)$$

ou seja, a distribuição e as características estatísticas são as mesmas, tanto para a soma agregada como para a série original. Lamperti [29], demonstrou que para a equação (2.2) ser um processo estocástico auto-similar deve existir um único valor de m e H que satisfaz essa equação. Logo, se um processo é auto-similar, a série original tem a mesma função de autocorrelação das séries obtidas pela divisão proporcional da mesma.

Mais formalmente, processos auto-similares são caracterizados por uma função de autocorrelação ($\rho(k) \approx k^{-\eta}$ com $0 < \eta < 1$ e $k \rightarrow \infty$) que decai hiperbolicamente[†]. Outro indício que um processo apresenta características auto-similares, pode ser visto através da função de autocorrelação apresentando soma infinita, através da densidade espectral do processo que diverge na origem ($\mathcal{S}(f) \approx f^{-\theta}$ com $0 < \theta < 1$ e $f \rightarrow 0$), através do índice de dispersão (IDC) que cresce monotonicamente por várias escalas de tempo ou ainda através da distribuição de probabilidade que apresenta cauda pesada [6, 3, 11, 17, 35, 36].

2.2.2 Identificando a característica de auto-similaridade em processos estocásticos

Existem atualmente diferentes técnicas que podem ser empregadas para se medir o grau de auto-similaridade de um processo estocástico. Dentre elas, temos: representação gráfica Variância-Tempo [2, 3], análise R/S (Estatística de Amplitude

[†]Nome dado ao efeito do decaimento da função de autocorrelação ser mais lento do que uma função exponencial, ou mais precisamente seguindo uma lei de potência.

Reescalada Ajustada) [6, 37], método do Periodograma [3, 11], estimador de máxima verossimilhança de Whittle [6, 35], etc. dentre estes, pode-se destacar o método da representação gráfica da Variância-Tempo, que apresenta baixa complexidade, e o método da análise R/S, também conhecido como diagrama de *por* [6], por ser um método preciso [37] para aferir o grau de auto-similaridade de um processo estocástico.

Nesta subseção é apresentado conceitualmente o método da análise R/S e também é feita uma ilustração do método aplicando-o para medir o grau de auto-similaridade do tráfego de entrada da rede de pacotes do INATEL.

Análise R/S – Estatística de Amplitude Reescalada Ajustada

O método da análise R/S, considera uma seqüência finita de observações $\{X_k, k = 1, 2, \dots, n\}$ com média $Y(n)$ e variância das amostras dada por $S^2(n)$, então a amplitude reescalada ajustada é dada por:

$$\frac{R(n)}{S(n)} = \frac{1}{S(n)} \left[\max_{0 \leq t \leq n} (Y(t) - tY(n)) - \min_{0 \leq t \leq n} (Y(t) - tY(n)) \right], \quad (2.3)$$

sendo que $S(n)$ é definido como

$$S(n) = \left[\frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2 \right]^{\frac{1}{2}}, \quad (2.4)$$

e $Y(n)$ como

$$Y(n) = \frac{1}{n} \sum_{i=1}^n X_i. \quad (2.5)$$

Hurst [38] mostrou que muitas séries temporais satisfazem a relação

$$E \left[\frac{R(n)}{S(n)} \right] \approx cn^H, n \rightarrow \infty \quad (2.6)$$

onde c é uma constante que não depende de n e $S(n)$ é o desvio padrão.

A declividade da reta ajustada para o gráfico $(\log(n), \log(R(n)/S(n)))$ é uma estimativa do parâmetro de Hurst – H . Como já foi mencionado anteriormente, este parâmetro traz a informação do grau de auto-similaridade de um processo estocástico.

A Figura-2.3 traz o gráfico da análise R/S aplicado ao tráfego de saída e de entrada do *gateway* da rede do INATEL.

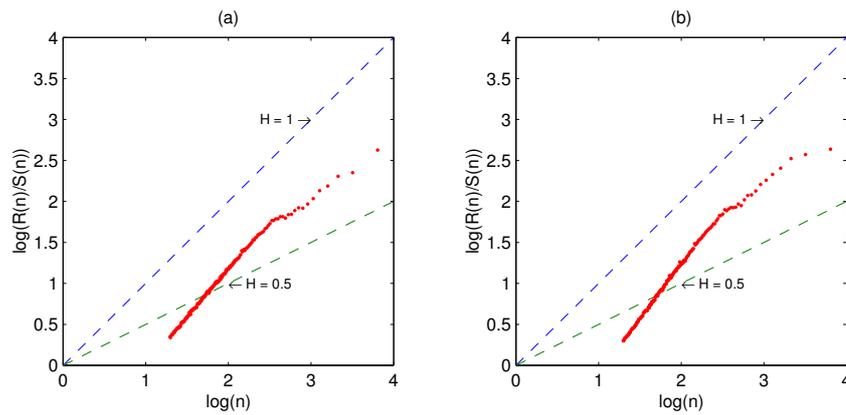


Figura 2.3: Em (a) pode-se ver a análise R/S do tráfego de saída da rede de pacotes do INATEL. Em (b) é apresentada a análise R/S do tráfego de entrada da rede de pacotes do INATEL.

2.3 Modelos de tráfego

Como já foi mencionado anteriormente, os dados coletados nas redes de pacotes demonstram, convincentemente, que o tráfego pode apresentar característica de auto-similaridade [19, 20, 6, 21, 9, 38, 39]. O conceito de auto-similaridade, ou mais precisamente auto-proximidade, também é conhecido como *fractality* ou dependência de longa duração.

Desde que foi introduzido pela primeira vez, esse conceito tem apresentado um destaque singular nas técnicas de compressão, processamento de sinais e mais recentemente na análise do tráfego de redes de pacotes [6, 39, 40, 15].

Como pode ser visto em [2, 3, 21, 18] existe uma extensa bibliografia que aborda este assunto. Devido a essa abrangência, serão apresentados, nesta seção, alguns modelos estocásticos que podem ser empregados para efetuar a modelagem do comportamento auto-similar do tráfego nas redes atuais, destacando-se apenas as tecnologias com maior evidência.

2.3.1 Modelo para redes Ethernet

Uma das maiores evidências da auto-similaridade do tráfego em redes Ethernet foi apresentado por Leland *et.al.* [19, 20, 6]. Uma explicação para esse fenômeno pode ser dada através da teoria de processos On/Off. Nesta, cada estação de trabalho é considerada como On (enviando pacotes para a rede) ou como Off (não utilizando a rede), então a superposição de tais fontes On/Off pode ser assintoticamente auto-similar [10, 41]. Outro modelo que descreve a dependência de longa duração do tráfego Ethernet é o modelo baseado em Wavelets, como apresentado em [42].

Modelos markovianos utilizados no passado, se forem devidamente adaptados, também podem ser empregados para descrever o tráfego auto-similar como se pode ver em [16]. Este tipo de modelagem traz um ganho no que se refere ao tratamento matemático, devido à sua simplicidade comparada com outros modelos. De acordo com Ryo *et. al.* [43] o tráfego em redes Ethernet pode ser modelado através de Processos de Contagem, que é capaz de tratar a auto-similaridade apresentada por esse tipo de tráfego.

Como visto, existem vários modelos matemáticos capazes de tratar auto-similaridade apresentada pelo tráfego em redes Ethernet. Pode-se observar que os modelos se diferenciam uns dos outros através da complexidade inerente a cada um deles. Em [2] e [21] são descritos outros modelos capazes de tratar o tráfego em redes Ethernet.

2.3.2 Modelos para redes WAN

Análise de dados coletados em redes WAN, como apresentada por Paxson e Floyd [9], demonstram que modelos poissonianos falham em capturar a dependência de longa duração presentes no tráfego deste tipo de rede.

Modelos que utilizam fontes On/Off, mencionado anteriormente, também podem ser usados para fazer a modelagem de tráfego em rede WAN. Outro método que pode ser empregado para representar o tráfego em redes WAN é o modelo $M/G/\infty$, onde os clientes (pacotes) chegam de conformidade com uma distribuição de Poisson e tem um tempo de serviço, modelado de acordo com uma distribuição de cauda pesada [6, 44, 45, 9].

Pode-se encontrar, atualmente, uma série de modelos que são capazes de representar o tráfego em redes WAN, como por exemplo os modelos baseados em Wavelets [45], MMPP [23] e Pareto [9, 23]. É possível, ainda, encontrar outros modelos que podem ser aplicados ao tráfego das redes WAN, na extensa bibliografia apresentada por Hlavacs *et. al.* [2] e Ruela *et. al.* [21] e Willinger *et. al.* [18].

2.3.3 Modelos para redes ATM

A tecnologia ATM foi amplamente difundida na última década, onde seu diferencial era suportar uma ampla variedade de serviços. Alguns tipos de serviços suportados por essa tecnologia, podem ser vistos na Tabela-2.1.

Devido à sua grande flexibilidade em suportar diferentes tipos de serviços, surgiram inúmeros modelos de tráfego disponíveis para tecnologia ATM.

Um modelo que pode ser utilizado para representar o tráfego ATM é o proposto por Fan e Georganas [24], que modela o tráfego auto-similar através de uma

Tabela 2.1: Exemplos de aplicações suportadas pela tecnologia ATM, classificados de acordo com os tipos de classes de serviços

Classe	A	B	C	D
Exemplo de Serviços	Emulação de circuitos e video a taxa constante	Vídeo e áudio a taxas variáveis	Transf. de dados orientada a conexão	Transf. de dados não orientada a conexão
Protocolo AAL	AAL1	AAL2	AAL3,4	AAL5

distribuição Poisson-Zeta, com parâmetros α e ρ , sendo que o tráfego agregado é gerado através de processos On/Off. Este modelo foi empregado para descrever o tráfego de entrada de comutadores ATM, com o objetivo de verificar seu respectivo desempenho mediante um tráfego com característica auto-similar.

Modelos clássicos, como o modelo $M/G/1$ também podem ser utilizados, se forem devidamente ajustados. Um exemplo disso são os trabalhos apresentados por Tsybakov *et. al.* [25, 17, 26], onde são apresentados modelos para dimensionamento de *buffer* para redes ATM considerando um tráfego auto-similar.

Existem, ainda, outros modelos capazes de descrever o tráfego auto-similar nas redes ATM,. Dentre eles tem-se os modelos propostos por Diamond *et. al.* [27] e Daniels *et. al.* [14], que são baseados em distribuições de cauda pesada e processos On/Off .

2.3.4 Modelos para WWW

Muitos estudos têm sido feitos não apenas com o intuito de se determinar modelos de tráfego que sejam apropriados para representar o comportamento da *Web*. Trabalhos como o apresentado por Crovella e Bestavros [10, 11] buscam identificar, também, a natureza da auto-similaridade deste tipo de tráfego.

Crovella e Bestavros apresentam um modelo para o tráfego Web usando a superposição de alguns processos de renovação, onde a renovação está restrita aos valores 0 e 1 e, o intervalo de tempo entre as renovações segue uma distribuição de cauda pesada. Outros modelos como os apresentados por Feldmann *et. al.* [45], que utilizam como base Wavelets e $M/G/\infty$, também podem ser utilizados para representar o comportamento desse tipo de tráfego.

2.4 Conclusões

Neste capítulo buscou-se sumariar as principais técnicas empregadas para analisar o comportamento do tráfego nas redes de pacotes, com o objetivo de efetuar

uma caracterização precisa do seu comportamento. Além disso, pode-se observar através da análise R/S aplicada no tráfego de entrada e saída do *gateway* da rede do INATEL, que o tráfego pode não apresentar simetria. Isso dificulta, ainda mais, o dimensionamento, a análise de desempenho e a otimização das redes de pacotes.

Foram apresentados, também, diversos modelos probabilísticos que podem ser empregados para descrever matematicamente a característica de auto-similaridade presente no tráfego das redes de pacotes. No entanto, os modelos apresentados possuem um tratamento matemático “difícil” bem como a fato de não existir um modelo definitivo e concreto que descreva o comportamento do tráfego nas redes de pacotes. Daí surge a principal motivação para se buscar modelos, de representação de tráfego de redes de pacotes, que sejam factíveis para serem utilizados no dimensionamento e na análise de desempenho de redes de pacotes.

Como resultado de nossa investigação, no próximo capítulo estaremos apresentando um modelo, matemático, que seja capaz de descrever o comportamento auto-similar do tráfego das redes de pacotes e que também seja de baixa complexidade matemática.

Capítulo 3

Modelagem do Tráfego com Característica Auto-Similar

3.1 Introdução

Quando se objetiva construir um modelo de tráfego para representar o comportamento de uma determinada rede ou mesmo de um elemento desta rede, um dos pontos cruciais é desenvolver um modelo matemático que seja capaz de representar a fonte geradora do tráfego. Outro aspecto de extrema relevância é conhecer as características estatísticas do mesmo.

Como pode ser visto no Capítulo-2, o tráfego nas redes de pacotes apresenta uma característica diferente em relação ao tráfego das redes de telefonia convencionais. Em virtude dessa nova característica os modelos baseados em *Processos de Poisson*, *Cadeias de Markov*, *Batch-Poisson*, *Markov Modulated Poisson Process (MMPP)*, *Modelos de Fluido*, etc., não são adequados para representar esse novo comportamento [3].

No entanto, como pode-se observar em [20, 6, 2, 10, 11, 18, 17, 24, 25, 26, 27, 14], a característica de auto-similaridade do tráfego nas redes de pacotes, pode ser representada através da sobreposição de processos de renovação [22, 46], através de modelos FBm (*Fractional Brownian Motion*) [13, 23, 28, 12], FGn (*Fractional Gaussian Noise*) [6, 3, 2], Wavelets [47, 2], etc.. Porém, estes modelos são muito sofisticados e complexos de serem manipulados matematicamente.

Portanto, a finalidade deste capítulo é desenvolver um modelo matemático que seja capaz de representar esse novo comportamento do tráfego nas redes de pacotes e que apresente complexidade matemática inferior aos modelos já mencionados. Para isso, este capítulo é dividido da seguinte forma: a primeira parte traz a proposição do modelo matemático e a inclusão do parâmetro de

Hurst no modelo; a segunda parte trata da validação teórica do modelo proposto; e finalizando esse capítulo, são apresentadas algumas conclusões a respeito do modelo proposto.

3.2 Proposição do Modelo de Tráfego

O comportamento do tráfego nas redes de pacotes apresenta característica diferenciada quando comparado com o tráfego nas redes de telefonia convencional. Essa nova característica, conhecida como dependência de longa duração ou auto-similaridade, começou a ganhar destaque após os trabalhos apresentados por Leland *et. al.* [19, 20], na primeira metade da década de 90. Desde então, inúmeros estudos foram realizados com a finalidade de investigar essa nova característica [6, 13, 7, 9, 41, 18, 33], bem como, apresentar técnicas para identificação desse novo comportamento [6, 3, 2, 17].

Como já pode ser visto no Capítulo–2, existem inúmeras maneiras para identificar, matematicamente, a auto-similaridade de um processo estocástico.

Nesta seção, far-se-á uso do conceito apresentado inicialmente por Mandelbrot [48] e explorando por Leland *et. al.* [6], Fernandez *et. al.* [23], Cappe *et. al.* [36], etc. para justificar a utilização de distribuições de probabilidade que são classificadas de cauda pesada como uma ferramenta de modelagem de tráfego auto-similar. A abordagem aqui apresentada está direcionada à distribuição de Pareto. Finalizando essa seção, será mostrado como a distribuição de Pareto pode incorporar o parâmetro de Hurst, que indica o grau de auto-similaridade de um processo estocástico.

3.2.1 Modelo Matemático do Tráfego

Pode ser visto na seção–2.2.1, que a auto-similaridade se manifesta matematicamente de diversas formas. De acordo com o que foi exposto, uma das maneiras de se identificar se um modelo matemático é capaz de modelar a auto-similaridade é se este apresenta uma função distribuição de probabilidade com cauda pesada.

Mandelbrot [48], apresenta a seguinte definição para modelos probabilísticos com cauda pesada.

Definição 3.1. *Uma variável aleatória T é dita ser de cauda pesada, se a função distribuição de probabilidade complementar (f_{dc}^{-1}) segue a seguinte forma:*

$$P[|T| \geq t] \approx \frac{L(t)}{t^\alpha}, \quad (3.1)$$

com variância infinita, $0 < \alpha < 2$ e $L(t)$ é uma função que varia lentamente quando $t \rightarrow \infty$.

A distribuição de Pareto, apresentada no Anexo-A, pode apresentar as características equivalentes a Definição-3.1 se o parâmetro α for devidamente escolhido.

De acordo com o desenvolvimento matemático apresentado no Anexo-A, a distribuição de Pareto atende a estes requisitos para os valores de α no intervalo $0 < \alpha < 2$. Portanto, pode-se afirmar que esta distribuição apresentada pela equação (A.1),

$$f_T(t) = \begin{cases} \frac{\alpha}{\beta} \left(1 + \frac{t}{\beta}\right)^{-\alpha-1} & 0 < t < \infty, 0 < \alpha < 2 \text{ e } \beta > 0 \\ 0 & t < 0 \end{cases}, \quad (3.2)$$

é adequada para modelar tráfego com característica auto-similar - veja Figura-ref.

Distribuição de Pareto para alguns valores de α e β . Em (a) tem-se $\alpha = 1.5$ e $\beta = 0.1$, em (b) tem-se $\alpha = 1.5$ e $\beta = 1$, em (c) tem-se $\alpha = 1.5$ e $\beta = 10$ e em (d) tem-se $\alpha = 1.5$ e $\beta = 100$

No entanto, se considerarmos todo o intervalo da variável α para a distribuição de Pareto, dada pela equação (3.2), pode-se notar que a média da distribuição será infinita para os valores de α que estiverem dentro do intervalo $(0, 1]$, veja Anexo-A. Por esse motivo iremos considerar apenas os valores de α que estiverem dentro do intervalo $1 < \alpha < 2$. Portanto, a equação (3.2) terá a seguinte forma,

$$f_T(t) = \begin{cases} \frac{\alpha}{\beta} \left(1 + \frac{t}{\beta}\right)^{-\alpha-1} & 0 < t < \infty, 1 < \alpha < 2 \text{ e } \beta > 0 \\ 0 & t < 0 \end{cases}. \quad (3.3)$$

3.2.2 Parâmetro de Hurst e a Distribuição de Pareto

De acordo com o que foi exposto, os valores de α que serão considerados estarão dentro do intervalo $1 < \alpha < 2$. No entanto, fica ainda a seguinte questão: como a distribuição de Pareto, apresentada pela equação (3.3), e mencionada em várias referências [6, 10, 23, 36, 49], poderá incorporar o parâmetro de Hurst, responsável por indicar o grau de auto-similaridade de um processo estocástico?

Para responder esta questão, será considerado, primeiramente, que a variável α apresenta alguma lei de relacionamento com o parâmetro H, ou seja,

$$\alpha \propto H. \quad (3.4)$$

Se for considerado que a variável α , na equação (3.4), é uma função de H, pode-se

reescrever a equação (3.4) da seguinte maneira,

$$\alpha = f(H). \quad (3.5)$$

Com o intuito de investigar a forma da função $f(H)$, apresentada na equação (3.5), considerar-se-á o plano formado pela interseção dos valores de α e H - veja Figura-3.1a. Para facilitar a identificação da função $f(H)$, será admitida uma variação linear entre a variável α e o parâmetro de Hurst - veja Figura-3.1b. Partindo, dessa suposição [23, 28], podemos reescrever a equação (3.5) da

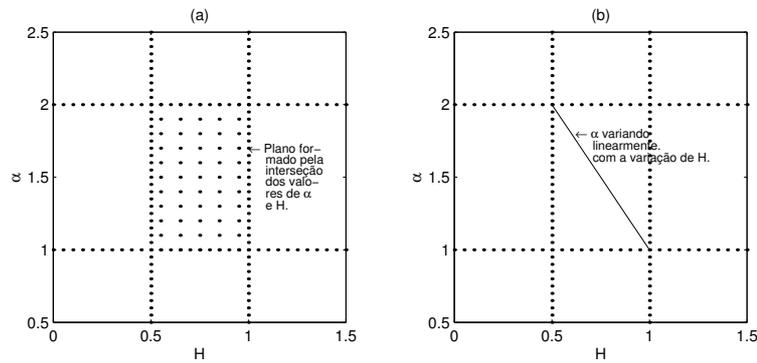


Figura 3.1: Gráfico da variação do parâmetro α , da distribuição de Pareto, em relação aos valores do parâmetro de Hurst - H . Em (a) observa-se o plano formado pela interseção das faixas de valores dos parâmetros α e H . Em (b) é apresentado o relacionamento entre os valores de α e H .

seguinte forma:

$$\alpha(H) = 3 - 2H. \quad (3.6)$$

Portanto, podemos considerar que a equação (3.3) pode ser reescrita, da seguinte forma,

$$f_T(t) = \begin{cases} \frac{\alpha(H)}{\beta} \left(1 + \frac{t}{\beta}\right)^{-\alpha(H)-1} & t > 0 \\ 0 & t < 0 \end{cases}, \quad (3.7)$$

onde $\alpha(H)$ é calculado através da equação (3.6), $\beta > 0$ e $0.5 < H < 1$.

Como o novo valor da variável α apresenta restrições impostas pelo parâmetro de Hurst, será verificado se a variável β sofrerá algum tipo de restrição. Para efetuar essa verificação será usada a definição da função densidade de probabilidade, apresentada na equação (3.8).

$$\int_{-\infty}^{\infty} f_X(x) = 1. \quad (3.8)$$

Aplicando a equação (3.7) na definição apresentada em (3.8), tem-se:

$$\int_0^{\infty} \frac{\alpha(H)}{\beta} \left(1 + \frac{t}{\beta}\right)^{-\alpha(H)-1} dt = 1. \quad (3.9)$$

Resolvendo a equação (3.9), com o objetivo de determinar o valor da variável β , pode-se demonstrar que qualquer valor de $\beta : \beta \in \mathbb{R}$ satisfaz a equação (3.9). Portanto, a única restrição imposta ao valor de β é o da própria definição da distribuição de Pareto. Ou seja,

$$\beta > 0.$$

No entanto, apesar de β poder assumir qualquer valor no conjunto \mathbb{R}_+^* , não é possível afirmar, com exatidão, que a distribuição de Pareto, apresentada pela equação (3.7), levando-se em conta apenas a variável α como sendo uma função de H , pode ser considerada como um modelo capaz de representar o comportamento auto-similar do tráfego, com um grau de auto-similaridade dada pelo parâmetro de Hurst.

Com o objetivo de investigar a importância do valor da variável β , na equação (3.7), será considerada a estatística de segunda ordem de um processo estocástico estacionário com fdp dada pela equação (3.7). Ou seja,

$$R_T(\tau) = \frac{\beta^{\alpha(H)+1}}{(1 - \alpha(H))} \left[\frac{\alpha(H)(\tau + \beta) - (1 - \alpha(H))\beta}{(1 - \alpha(H))(\tau + \beta)^{\alpha(H)}} \right], \quad (3.10)$$

onde $0,5 < H < 1$ e $\beta > 0$, que é a função de autocorrelação do referido processo estocástico. O desenvolvimento da equação (3.10) é mostrado no Anexo-B.

Seja também um processo estocástico com função de autocorrelação dada por,

$$\rho(\tau) = \frac{1}{2} (|\tau + 1|^{2H} - 2|\tau|^{2H} + |\tau - 1|^{2H}), \quad \tau = 1, 2, 3, \dots \quad (3.11)$$

O processo estocástico estacionário descrito pela função de autocorrelação, apresentada pela equação (3.11) é conhecido como Ruído Gaussiano Fracional ou FGn [6, 2, 3, 12, 50]. Este processo estocástico é capaz de descrever, de forma adequada, o tráfego presente nas redes de pacotes .

Se for feita uma comparação gráfica entre as equações (3.10) e (3.11) para um mesmo valor de H e diferentes valores de β , pode-se ver, através da Figura-3.2, que o valor de β também deve ser considerado para que a distribuição de probabilidade apresentada pela equação (3.9) seja capaz de modelar tráfego auto-similar, onde o grau de auto-similaridade é dado pelo parâmetro de Hurst.

Portanto, observa-se, através da Figura-3.2, que a variável β também deve ser considerada para que a distribuição de Pareto, apresentada na equação (3.7),

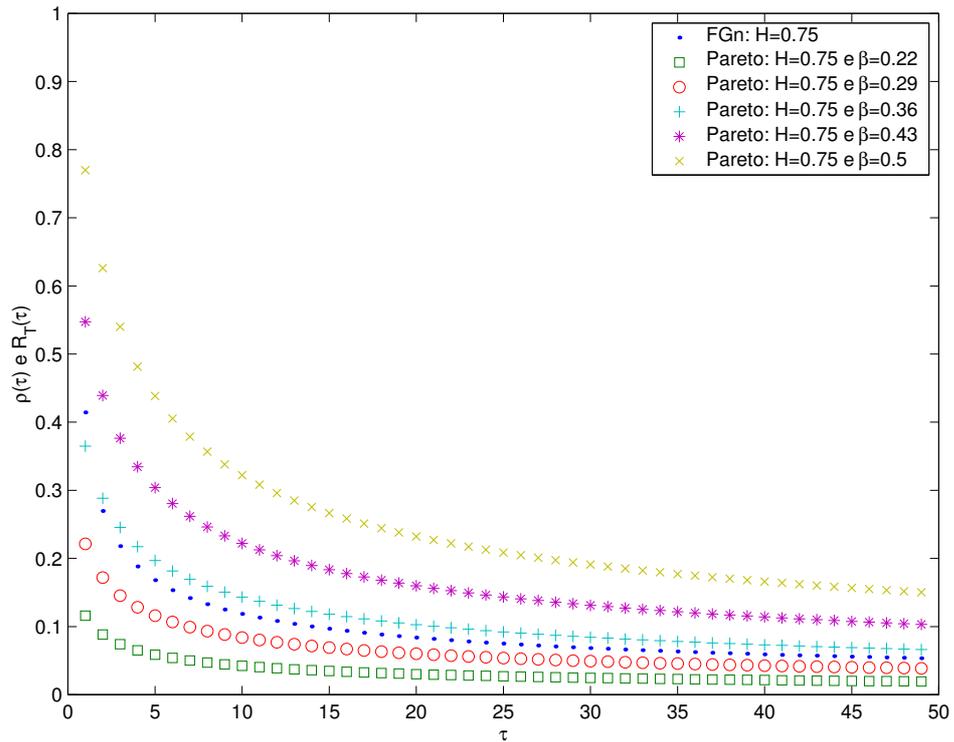


Figura 3.2: Comparação entre a função de autocorrelação do modelo FGn , com $H=0.75$, e o função de autocorrelação do Modelo de Pareto considerando apenas a variável α como sendo função do parâmetro H e $\beta > 0$.

possa ser capaz de capturar a característica auto-similar presente no tráfego das redes de pacotes.

Para determinar o valor de β na distribuição de Pareto, na equação (3.7), é suposto que a variável β também apresenta algum tipo de relacionamento com o parâmetro de Hurst. Ou seja, considera-se que,

$$\beta \propto H. \quad (3.12)$$

Se for considerado que a variável β , na equação (3.12), é uma função de H , pode-se reescrever a equação (3.12) da seguinte maneira,

$$\beta = g(H). \quad (3.13)$$

No entanto, não se pode afirmar que o relacionamento entre a variável β e H apresente alguma característica de linearidade, como foi feito com a variável α . Portanto, para determinar a forma da função $g(H)$, será tomado o erro médio quadrático entre a função de autocorrelação de Pareto (3.10) e do Ruído Gaus-

siano Fracional (3.11):

$$\varepsilon(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{\tau=1}^T [R_T(\tau) - \rho(\tau)]^2. \quad (3.14)$$

Substituindo-se as equações (3.10) e (3.11) na equação (3.14), tem-se:

$$\varepsilon(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{\tau=1}^T \left\{ \frac{\beta^{\alpha(H)+1}}{(1-\alpha(H))} \left[\frac{\alpha(H)(\tau + \beta) - (1-\alpha(H))\beta}{(1-\alpha(H))(\tau + \beta)^{\alpha(H)}} \right] - \frac{1}{2} (|\tau + 1|^{2H} - 2|\tau|^{2H} + |\tau - 1|^{2H}) \right\}^2. \quad (3.15)$$

Para determinar o relacionamento entre a variável β e o parâmetro de Hurst – H, deve-se resolver a equação (3.15) em função de β , para um dado valor de H, que minimize o erro médio quadrático.

Devido à complexidade apresentada pela equação (3.15), tomar-se-á uma solução numérica para diferentes valores de H, com o objetivo de determinar o valor da variável β . A Figura-3.4a apresenta alguns valores de β em relação ao parâmetro H. Na Figura-3.4b, tem-se a curva obtida através da interpolação dos valores produzidos a partir da solução da equação (3.15) com o objetivo de determinar o relacionamento entre a variável β e o parâmetro H.

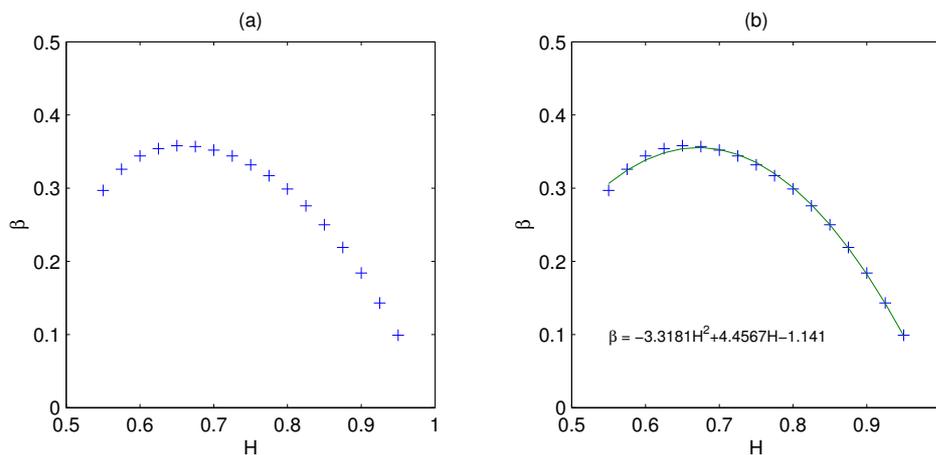


Figura 3.3: Relação entre a variável β e o parâmetro de Hurst. Em (a) vemos alguns valores de β em função de H. Em (b) é apresentada a curva obtida através de interpolação dos pontos apresentados em (a).

A partir do resultado da interpolação feita com os resultados obtidos da

solução da equação (3.15), pode-se expressar a função $g(H)$, da seguinte forma:

$$g(H) \approx -3.3181H^2 + 4.4567H - 1.141. \quad (3.16)$$

Substituindo a equação (3.16) na equação (3.13), temos:

$$\beta(H) = -3.3181H^2 + 4.4567H - 1.141. \quad (3.17)$$

Finalmente, pode-se reescrever a equação (3.7) da seguinte forma:

$$f_T(t) = \begin{cases} \frac{\alpha(H)}{\beta(H)} \left(1 + \frac{t}{\beta(H)}\right)^{-\alpha(H)-1} & t > 0 \\ 0 & t < 0 \end{cases}, \quad (3.18)$$

onde $0.5 < H < 1$, $\alpha(H)$ determinado a partir da equação (3.6) e $\beta(H)$ determinado a partir da equação (3.17).

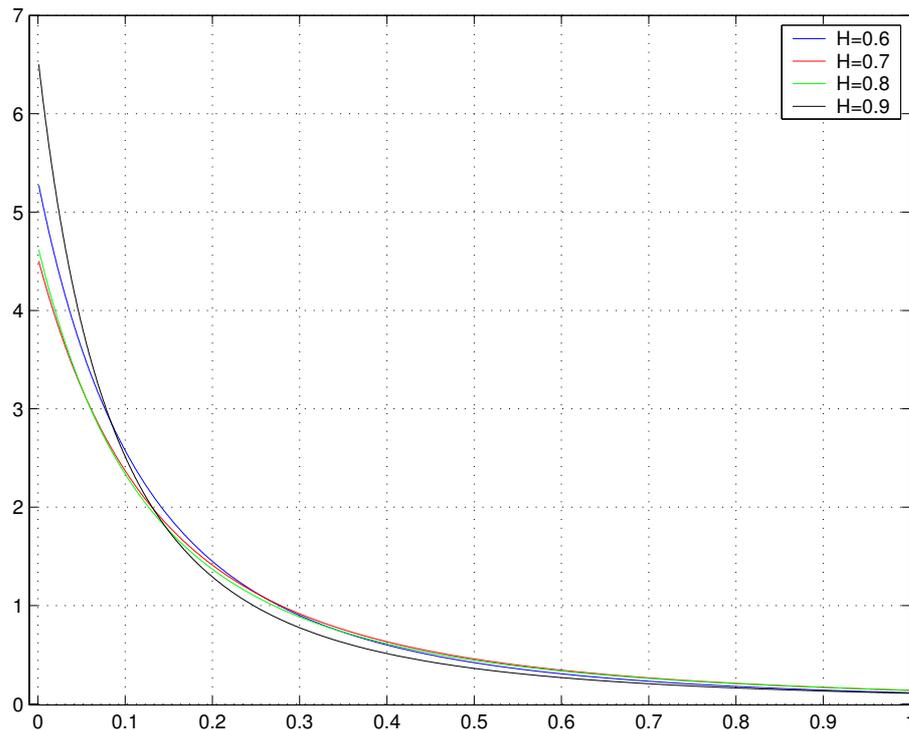


Figura 3.4: Função densidade de probabilidade de Pareto com parâmetro de Hurst.

Portanto, a função densidade de probabilidade de Pareto, apresentada na equação (3.18), considerando o parâmetro de Hurst, pode ser aplicada para descrever o comportamento do tráfego com característica auto-similar.

3.2.3 Validação Teórica do Modelo de Tráfego

Como já foi apresentado, no Capítulo-2, a auto-similaridade se manifesta de várias maneiras. Com o objetivo de validar o modelo apresentado, na subseção anterior, vamos considerar as definições apresentadas em [6, 3], ou seja,

Definição 3.2. *Processos estocásticos com características auto-similares apresentam função de autocorrelação, da seguinte forma:*

$$r(k) \approx a_1 k^{-\eta}, \quad (3.19)$$

com $k \rightarrow \infty$, $a_1 > 0$ e $0 < \eta < 1$.

Esta definição caracteriza o decaimento hiperbólico da função de autocorrelação.

Para validar o modelo matemático, será considerado um processo estocástico estacionário com fdp de Pareto, veja Anexo-B. Analisando o comportamento da função de autocorrelação desse processo quando $\tau \rightarrow \infty$, tem-se:

$$R_T(\tau) \rightarrow \frac{\alpha\beta^{\alpha+1}}{(1-\alpha)^2}(\tau + \beta)^{1-\alpha} \text{ para } \tau \rightarrow \infty \text{ e } 1 < \alpha < 2. \quad (3.20)$$

Seja ainda,

$$a_1 = \frac{\alpha\beta^{\alpha+1}}{(1-\alpha)^2}, \quad (3.21)$$

e

$$(\tau + \beta)^{1-\alpha} \rightarrow \tau^{1-\alpha} \text{ para } \tau \gg \beta. \quad (3.22)$$

Substituindo a expressão (3.21) e (3.22) na expressão (3.20), tem-se:

$$R_T(\tau) \approx a_1 \tau^{1-\alpha} \text{ para } \tau \rightarrow \infty \text{ e } 1 < \alpha < 2. \quad (3.23)$$

Comparando o resultado de (3.23) com a Definição-3.2 o processo estocástico apresenta característica auto-similar, se

$$-1 < 1 - \alpha < 0. \quad (3.24)$$

Substituindo o resultado da equação (3.6) na inequação (3.24), tem-se,

$$0.5 < H < 1, \quad (3.25)$$

que valida o modelo apresentado.

3.3 Conclusões

Nas inúmeras referências bibliográficas utilizadas nesta dissertação, observa-se que existem muitas propostas de modelos matemáticos para representar o comportamento auto-similar do tráfego nas redes de pacotes. Exemplos disso, são os modelos baseados em fontes ON/OFF [20, 6, 2, 10, 11, 18, 17, 24, 25, 26, 27, 14] assim como a utilização da distribuição de Pareto [6, 10, 28, 23, 36, 49] são mencionadas como modelos capazes de representar a auto-similaridade do tráfego. No entanto, não é feita nenhuma menção a respeito do grau de auto-similaridade apresentado por esses modelos.

Neste capítulo buscou-se investigar qual deve ser o relacionamento entre as variáveis α e β com o parâmetro de Hurst, para que a distribuição de Pareto 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 que é necessário 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 [51].

Finalmente, neste capítulo não se buscou apresentar uma nova descoberta a respeito da distribuição de Pareto como modelo de tráfego auto-similar, mas o objetivo foi apresentar uma nova contribuição ao modelo existente, através da incorporação do parâmetro de Hurst a essa distribuição, mostrando também, que a variável de localização dessa distribuição deve ser considerada. Portanto, pode-se concluir que o modelo apresentado pela equação (3.18) é capaz de modelar processos estocásticos com grau de auto-similaridade dado pelo parâmetro de Hurst.

Capítulo 4

Modelo de fila $G/M/1$ e a Distribuição de Pareto com Parâmetro de Hurst

4.1 Introdução

A natureza estatística do tráfego nas redes vem sofrendo mudanças drásticas. Por isso novas ferramentas para dimensionamento, análise de desempenho e otimização se fazem necessárias.

Muitas das ferramentas empregadas atualmente para análise de redes de pacotes fazem uso de modelos de filas. Essas ferramentas também utilizam o recurso de geração de números aleatórios para emular o tráfego das redes.

Neste capítulo não há nenhuma intensão em se construir uma ferramenta para analisar o desempenho de redes, onde o tráfego apresenta característica auto-similar. No entanto, o objetivo principal desse capítulo é apresentar uma análise matemática, considerando um modelo de fila $G/M/1$, onde a distribuição do tempo entre chegadas é dada pela distribuição de Pareto com parâmetro de Hurst, apresentada no Capítulo-3. Além disso, será feita uma comparação entre o modelo $G/M/1$ e o modelo $M/M/1$, através de simulação. Finalizando esse capítulo, são apresentadas as conclusões sobre os resultados obtidos.

4.2 Análise Matemática do Modelo $G/M/1$

Existem diversas publicações [52, 53] que trazem o desenvolvimento matemático de modelos de filas considerando uma distribuição genérica para o tempo entre

chegadas das entidades*. Este equacionamento é apresentado na Tabela-4.2.

Tabela 4.1: Equações válidas para um sistema de filas do tipo G/M/1 para σ determinado a partir da solução da transformada Laplace–Stieltjes para a distribuição do tempo entre chegadas.

Descrição	Equação
Número médio de entidades no sistema	$\bar{K} = \frac{\rho}{1 - \sigma}$
Tempo médio de permanência	$\bar{T} = \frac{1}{\mu(1 - \sigma)}$
Var. do n. médio de entidades no sistema	$\sigma_K^2 = \frac{\rho(1 + \sigma - \rho)}{(1 - \sigma)^2}$
Comprimento médio da fila	$\bar{Q} = \frac{\rho\sigma}{1 - \sigma}$
Variância do comprimento da fila	$\sigma_Q^2 = \frac{\rho\sigma(1 + \sigma(1 - \rho))}{(1 - \sigma)^2}$
Tempo médio de espera das entidades na fila	$\bar{W} = \frac{\sigma}{\mu(1 - \sigma)}$
Prob. n. de entidades no sistema	$\pi_k = \rho(1 - \sigma)\sigma^{k-1}, k > 0$ $\pi_0 = 1 - \rho$
Distribuição do tempo de espera	$F_W(t) = \begin{cases} 1 - \sigma & t = 0 \\ 1 - \sigma e^{-\mu(1-\sigma)t} & t > 0 \end{cases}$

Pode-se observar que as expressões apresentadas na Tabela-4.2 somente se aplicam se o valor de σ for determinado. Isso poderá ser feito, solucionando-se a equação (4.1), em função de σ .

$$\sigma = A^{\sim}(\mu - \mu\sigma). \quad (4.1)$$

Esta equação é conhecida como transformada de Laplace–Stieltjes [52, 53], onde $A^{\sim}(s)$ é a transformada de Laplace para a distribuição que é empregada para modelar o tempo entre chegadas das entidades.

De acordo com Kleinrock [52], a equação (4.1) admite um único valor, real, para σ , tal que $0 < \sigma < 1$ e $0 < \rho < 1$, para que o sistema se mantenha em equilíbrio.

Com o objetivo de determinar a solução da equação (4.1) para identificar o

*Entidade pode se considerado como pacotes, células, etc..

valor de σ , seja, a equação (4.2) que representa a transformada de Laplace da distribuição de Pareto (3.18).

$$\mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} \left[\frac{\alpha(H)}{\beta(H)} \left(1 + \frac{t}{\beta(H)} \right)^{-\alpha(H)-1} \right] dt. \quad (4.2)$$

De acordo com o Anexo-C, a solução para a equação (4.2), é dada por:

$$\mathcal{L}\{f(t)\} = \alpha\Gamma(-\alpha) \left[(\beta s)^{\alpha} e^{s\beta} - \sum_{n=0}^{\infty} \frac{(\beta s)^n}{\Gamma(n - \alpha + 1)} \right], \quad (4.3)$$

onde α é determinado a partir da equação (3.6) e β a partir da equação (3.17). Substituindo a equação (4.3) na equação (4.1), tem-se:

$$\sigma = \alpha\Gamma(-\alpha) \left[[\beta(\mu - \mu\sigma)]^{\alpha} e^{(\mu - \mu\sigma)\beta} - \sum_{n=0}^{\infty} \frac{[\beta(\mu - \mu\sigma)]^n}{\Gamma(n - \alpha + 1)} \right], \quad (4.4)$$

onde μ , é a média da distribuição que modela o tempo entre chegadas das entidades. Para a distribuição de Pareto com parâmetro de Hurst, tem-se:

$$\mu = \frac{-3.3181H^2 + 4.4567H - 1.141}{(2 - 2H)}. \quad (4.5)$$

Para determinar o valor de σ , na equação (4.4), será usado o método gráfico tendo em vista que a referida equação é transcendental, o que inviabiliza a solução analítica. Portanto, seja,

$$g(\sigma) = \sigma \quad (4.6)$$

e

$$f(\sigma) = \alpha\Gamma(-\alpha) \left[[\beta(\mu - \mu\sigma)]^{\alpha} e^{(\mu - \mu\sigma)\beta} - \sum_{n=0}^{\infty} \frac{[\beta(\mu - \mu\sigma)]^n}{\Gamma(n - \alpha + 1)} \right]. \quad (4.7)$$

Se for tomado o ponto de interseção entre as funções $f(\sigma)$ e $g(\sigma)$ é possível determinar o valor de σ que satisfaz a equação (4.4), para diferentes valores de H .

A Figura-4.1 apresenta o ponto de intersecção entre as funções $f(\sigma)$ e $g(\sigma)$, considerando-se diferentes valores para o parâmetro de Hurst. Observa-se, que o único valor real positivo que satisfaz a equação (4.1) é $\sigma = 1$, o que resulta em uma solução trivial.

Portanto, conclui-se que a distribuição de Pareto com parâmetro de Hurst não pode ser empregada para determinar analiticamente o desempenho de filas $G/M/1$. Esse fato, não impede que o modelo proposto não possa ser empregado em um ambiente de simulação.

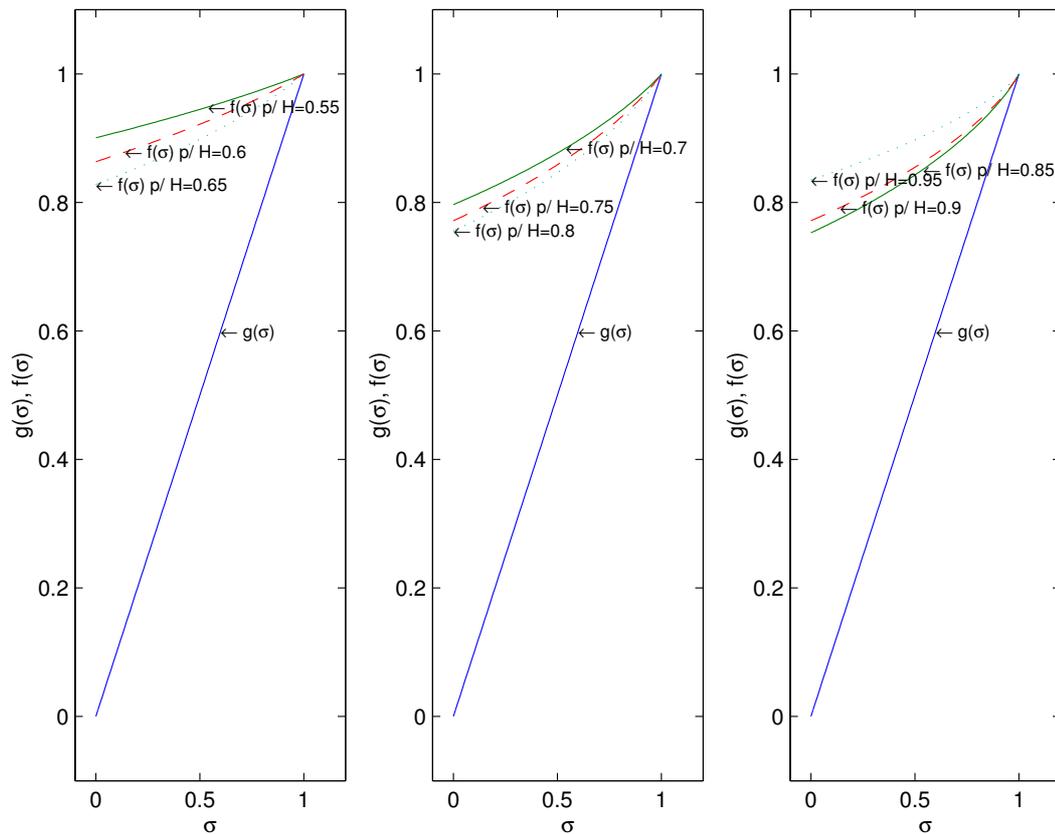


Figura 4.1: Ilustração gráfica da solução da equação (4.4) para diferentes valores de H .

4.3 Simulação: Modelo $G/M/1$ vs. $M/M/1$

De acordo com o que foi exposto na seção anterior, o modelo $G/M/1$ com a distribuição de Pareto modelando o tempo entre chegada de entidades não admite uma solução para σ , o que impede a comparação analítica entre os modelos de fila $M/M/1$ e $G/M/1$, onde G , que representa a distribuição do tempo entre chegada, é modelado de acordo com a equação (3.18).

Nesta seção, serão apresentados os resultados da comparação feita, através de simulação, utilizando-se a ferramenta *OPNET[®] Moduler 9.1 – Education Version*, entre o modelo $M/M/1$ e o modelo $G/M/1$. Primeiramente, será apresentado o modelo de simulação; e em seguida, serão apresentados os resultados obtidos.

4.3.1 Modelo de Simulação

O modelo empregado na simulação foi construído a partir do tutorial disponível na própria ferramenta *OPNET[®]*. Foi considerada uma fonte geradora de tráfego,

chamada *src*, um sistema de atendimento, chamado *queue*, e um sorvedouro, chamado *sink* – veja Figura-4.2

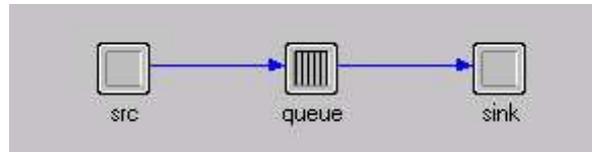


Figura 4.2: Modelo de simulação empregado para avaliar o comportamento de uma fila $M/M/1$ e $G/M/1$ considerando a mesma média para o tempo entre chegadas das entidades.

Com o objetivo de ter um referencial de comparação, o tempo médio entre chegada das entidades foi considerado o mesmo, tanto para a distribuição de Pareto com parâmetro de Hurst quanto para a distribuição exponencial, utilizada no modelo $M/M/1$.

O procedimento utilizado na simulação teve como objetivo extrair o tempo médio gasto na fila pelas entidades. Os passos empregados para comparar os dois modelos foram:

1. o bloco *src*, responsável pela geração das entidades, foi configurado inicialmente com a distribuição de Pareto, modelando o tempo entre chegada das entidades, onde as variáveis α e β são determinados a partir das equações (3.6) e (3.17), respectivamente;
2. o bloco *queue*, representando o sistema de atendimento, foi configurado para que o coeficiente de utilização do servidor variasse de 0.1 a 0.9;
3. considerando-se os diferentes fatores de utilização, dados do passo 2, juntamente com a variação do parâmetro de Hurst, determinado pelo passo 1, extrai-se o comportamento do sistema $G/M/1$.
4. altera-se a configuração do bloco *src*, substituindo-se a distribuição de Pareto pela distribuição Exponencial, considerando-se a mesma média para os diferentes valores de H ;
5. repete-se o passo de 2;
6. considerando-se os diferentes fatores de utilização, dados do passo 2, juntamente com a variação do parâmetro de Hurst no cálculo da média do tempo entre chegadas das entidades, determinado pelo passo 1, considerando-se a distribuição exponencial modelando o tempo entre as chegadas, extrai-se o comportamento do sistema $M/M/1$.

4.3.2 Resultados da Simulação

Os resultados provenientes das simulações realizadas, considerando-se o tempo médio gasto na fila pelas entidades, são apresentados nas Tabelas (4.2) e (4.3). O comprimento médio da fila foi determinado aplicando-se as equações de Little [53] – veja Figura-4.3.

A Tabela-4.2 traz os resultados referente ao tempo médio, dado em segundos, gasto na fila $G/M/1$, considerando que a distribuição empregada para modelar o tempo entre chegadas é a distribuição de Pareto com parâmetro de Hurst.

Tabela 4.2: Tempo médio, dado em segundos, gasto na fila $G/M/1$ para diferentes fatores de utilização, considerando-se que a distribuição do tempo entre chegadas é a distribuição de Pareto com parâmetro H .

H	ρ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.55	0.034	0.068	0.102	0.136	0.170	0.204	0.238	0.272	0.306
0.60	0.042	0.084	0.126	0.169	0.211	0.253	0.296	0.338	0.386
0.65	0.050	0.101	0.151	0.202	0.252	0.303	0.353	0.412	0.492
0.70	0.058	0.117	0.176	0.235	0.293	0.352	0.421	0.518	0.670
0.75	0.067	0.134	0.201	0.268	0.335	0.414	0.532	0.724	1.058
0.80	0.075	0.150	0.225	0.300	0.392	0.540	0.801	1.292	2.352
0.85	0.083	0.166	0.249	0.357	0.552	0.939	1.806	3.941	9.177
0.90	0.091	0.182	0.314	0.615	1.383	3.559	9.681	29.31	119.0
0.95	0.098	0.305	1.253	5.536	25.67	159.2	815.2	1934.8	3215.1

A Tabela-4.3 apresenta o tempo médio, dado em segundos, gasto na fila para que uma entidade inicie o seu atendida no sistema $M/M/1$.

Tabela 4.3: Tempo médio gasto na fila $M/M/1$ para diferentes fatores de utilização, considerando-se que a distribuição do tempo entre chegadas é exponencial com média dada pela distribuição de Pareto com parâmetro H .

H	ρ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.55	0.035	0.076	0.123	0.181	0.255	0.356	0.512	0.803	1.622
0.60	0.044	0.095	0.154	0.226	0.318	0.445	0.642	1.016	2.085
0.65	0.053	0.113	0.184	0.270	0.381	0.533	0.768	1.219	2.518
0.70	0.062	0.132	0.214	0.314	0.443	0.620	0.895	1.424	2.961
0.75	0.070	0.150	0.244	0.358	0.505	0.707	1.021	1.630	3.403
0.80	0.079	0.169	0.274	0.403	0.567	0.795	1.149	1.838	3.871
0.85	0.088	0.187	0.304	0.446	0.629	0.881	1.274	2.043	4.308
0.90	0.096	0.205	0.333	0.489	0.689	0.966	1.397	2.245	4.739
0.95	0.103	0.221	0.358	0.527	0.743	1.041	1.507	2.420	5.126

A Figura-4.3 traz o comportamento do tamanho médio da fila para dife-

rentes fatores de utilização considerando-se também diferentes índices de auto-similaridade.

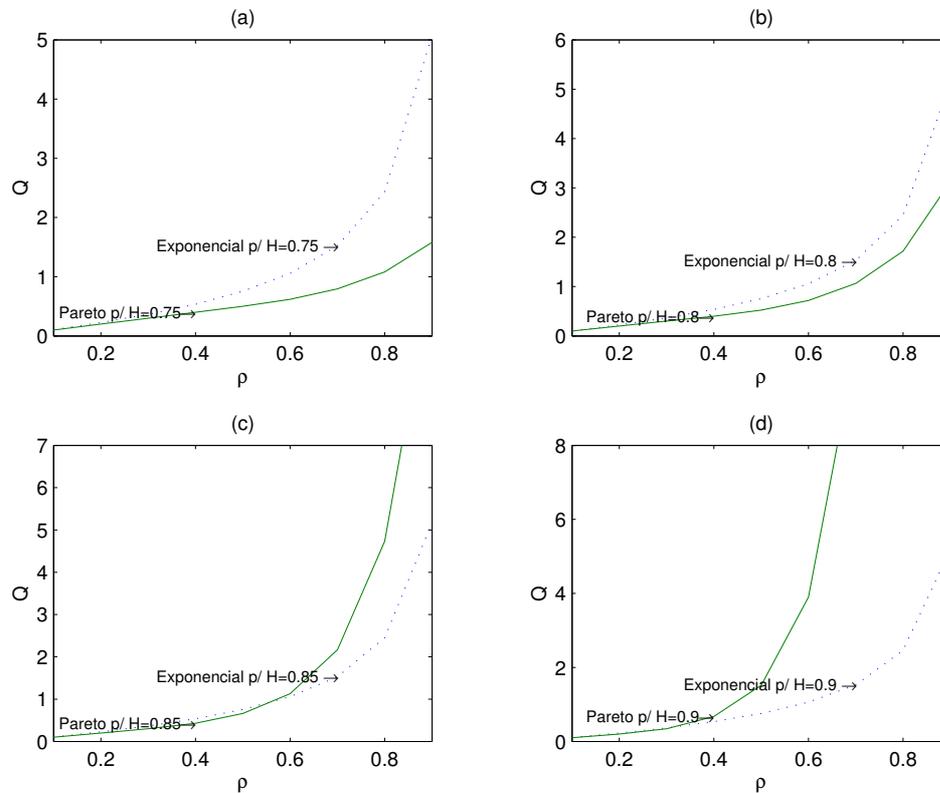


Figura 4.3: Comparação gráfica entre os modelos $G/M/1$ e $M/M/1$ levando-se em consideração o comprimento médio da fila.

Observa-se, através da Figura-4.3, que o comprimento médio da fila, usando a distribuição exponencial com média dada pela equação (4.5), modela de forma pessimista, o tráfego, para índices de auto-similaridade inferiores a 0.8 e passando a otimista para valores de H maior do que 0.85. É importante resaltar que o fator de utilização também deve ser considerado, como pode ser visto na figura referenciada.

O mesmo efeito estará presente se for considerado o tempo médio gasto no sistema por uma entidade e o número médio de entidades no sistema, de acordo com as equações de Little [53].

4.4 Conclusões

Neste capítulo foi demonstrado que a distribuição de Pareto com parâmetro de Hurst não admite solução analítica quando empregada para modelar filas $G/M/1$. O que caracteriza um ganho adicional, a essa distribuição, apresentada em varias

referências bibliográficas, como sendo capaz de modelar tráfego com característica auto-similar. No entanto, quando é considerado um em ambiente de simulação, o modelo proposto é factível.

A partir dos resultados apresentados nessa seção confirma-se que os modelos poissonianos não são mais capazes de tratar a natureza estatística do tráfego nas rede de pacotes, confirmando assim o que Paxson e Floyd apresentaram em [9]. Pode-se verificar, através dos resultados da simulação, que se considerarmos o modelo poissoniano, com média dada pela distribuição de Pareto, este pode ser empregado como uma aproximação na modelagem de tráfego auto-similar considerando-se fatores de utilização abaixo de 0.5.

Finalmente, conclui-se que o modelo apresentado traz uma melhoria nas várias propostas apresentadas [6, 11, 9, 36, 23, 49] quando afirmam que o modelo de Pareto é capaz de representar a auto-similaridade do tráfego, não fazendo menção a respeito de como o modelo descreve o grau de auto-similaridade dado pelo parâmetro de Hurst.

Capítulo 5

Conclusão

O principal objetivo a que se propôs esta dissertação foi investigar a natureza estatística do tráfego das redes de pacotes e apresentar uma modelagem matemática para representação desse tráfego. Adiciona-se, assim, mais um modelo aos inúmeros já existentes.

Com esse intuito, diversos estudos foram realizados compreendendo desde a natureza estatística do tráfego, passando pelos métodos para inferir a auto-similaridade do tráfego, até a análise de diversas propostas de modelos matemáticos. O Capítulo-2, é uma síntese desses estudos. No Capítulo-3, a proposição do modelo de tráfego é feita. Neste capítulo também é colocado o desenvolvimento matemático do modelo assim como a sua respectiva validação. Como uma consequência do desenvolvimento feito no Capítulo-3, o Capítulo-4 traz a investigação da aplicação do modelo desenvolvido, considerando a teoria clássica de filas, onde foi constatado que o mesmo não admite uma solução analítica podendo apenas ser empregado em ambiente de simulação.

Outras conclusões são encontradas no decorrer do texto dos próprios capítulos. A seguir, são apresentadas as contribuições desse trabalho e algumas proposições para futuros trabalhos, dando assim continuidade a essa dissertação.

5.1 Contribuições da Dissertação

Como primeira contribuição, destaca-se o desenvolvimento de uma expressão para a fdp de Pareto que incorpora o parâmetro de Hurst. Em segundo lugar foi apresentado uma análise matemática considerando o modelo de fila $G/M/1$ que mostra que a distribuição de Pareto com parâmetro de Hurst não admite uma análise matemática para determinar o desempenho desse tipo de fila. O desenvolvimento apresentado demonstra que os estudos apresentados por Leland [6], Paxson [9], Gordon [28] e Fernandez [23] consideram o modelo de Pareto com um

modelo de tráfego auto-similar, porém esse autores não fazem qualquer consideração a respeito do grau de auto-similaridade do modelo. Portanto, o modelo apresentado por essa dissertação caracteriza um ganho adicional a esse modelo, tão referenciado na bibliografia apresentada nessa dissertação.

5.2 Proposição para Trabalhos Futuros

Existem várias sugestões para futuros trabalhos com o objetivo de dar continuidade a esta dissertação. A primeira sugestão é o desenvolvimento de um simulador para redes de pacotes que seja capaz de considerar o parâmetro de Hurst para o tráfego emulado.

Um segundo trabalho seria efetuar uma comparação entre os modelos FGn e o modelo de Pareto com parâmetro de Hurst, considerando-se uma fila do tipo $G/M/1$, podendo assim acrescentar maiores detalhes ao trabalho de Fernandez *et. al.* [23] e para uma fila do tipo $G/M/1/K$, confrontando assim os resultados demonstrados por Koh *et. al.* [49].

Como terceiro trabalho a ser realizado pode-se sugerir a investigação de outras distribuições que apresente cauda pesada, para modelar tráfego com característica auto-similar, considerando o parâmetro de Hurst.

Anexo A

Distribuição de Pareto

Como pode ser visto através de várias referências bibliográficas, a distribuição de Pareto pode apresentar-se de diversas formas, veja [9, 36, 54, 49]. A forma utilizada nesta dissertação segue a mesma apresentada por Goovaert [54].

Função Densidade de Probabilidade - f_{dp}

$$f_T(t) = \begin{cases} \frac{\alpha}{\beta} \left(1 + \frac{t}{\beta}\right)^{-\alpha-1} & 0 < t < \infty \text{ e } \alpha, \beta > 0 \\ 0 & t < 0 \end{cases} \quad (\text{A.1})$$

Função de Distribuição Cumulativa - f_{dc}

$$\begin{aligned} F_T(t) &= P[T \leq t] \\ &= \int_{-\infty}^t f_T(x) dx \\ &= 1 - \int_t^{\infty} \frac{\alpha \beta^\alpha}{(x + \beta)^{\alpha+1}} dx \\ &= 1 - \alpha \beta^\alpha \left[\lim_{x \rightarrow \infty} \frac{(x + \beta)^{-(\alpha+1)+1}}{-\alpha + 1 + 1} - \frac{(t + \beta)^{-(\alpha+1)+1}}{-\alpha + 1 + 1} \right] \\ &= 1 - \alpha \beta^\alpha \left[\lim_{x \rightarrow \infty} \frac{1}{-\alpha(x + \beta)^\alpha} + \frac{1}{\alpha(t + \beta)^\alpha} \right] \\ &= 1 - \frac{\beta^\alpha}{(t + \beta)^\alpha} \end{aligned} \quad (\text{A.2})$$

Função de Distribuição Cumulativa Complementar - f_{dc}^{-1}

$$\begin{aligned}
 F_T^{-1}(t) &= P[T \geq t] \\
 &= 1 - P[T \leq t] \\
 &= \frac{\beta^\alpha}{(t + \beta)^\alpha}
 \end{aligned} \tag{A.3}$$

Momento de Primeira Ordem

O momento de primeira ordem da distribuição de Pareto é dado por:

$$\mu_1 = E[T] = \int_{-\infty}^{+\infty} t f_T(t) dt. \tag{A.4}$$

A solução da equação (A.4) é apresentada por Abramowitz [55]. Portanto, temos:

$$\begin{aligned}
 \mu_1 &= \int_0^{+\infty} \frac{t \alpha \beta^\alpha dt}{(t + \beta)^{\alpha+1}} \\
 &= \alpha \beta^\alpha \left[\lim_{b \rightarrow \infty} \left(\int_0^b \frac{t dt}{(t + \beta)^{\alpha+1}} \right) \right] \\
 &= \alpha \beta^\alpha \left[\lim_{b \rightarrow \infty} \left(\frac{(t + \beta)^{-\alpha+1}}{1 - \alpha} + \frac{\beta(t + \beta)^{-\alpha}}{\alpha} \right) \Big|_0^b \right] \\
 &= \alpha \beta^\alpha \left[\lim_{b \rightarrow \infty} \frac{(b + \beta)}{(1 - \alpha)(b + \beta)^\alpha} \right] + \frac{\beta}{\alpha - 1}
 \end{aligned} \tag{A.5}$$

O primeiro momento, ou o valor esperado, para a variável aleatória T , depende diretamente do valor de α . Portanto, a solução da equação (A.5) pode assumir dois valores, que são:

$$\mu_1 = \begin{cases} \frac{\beta}{\alpha - 1} & \alpha > 1 \\ +\infty & \alpha \leq 1 \end{cases} \tag{A.6}$$

Momento de Segunda Ordem

O momento de segunda ordem da distribuição de Pareto é dado por:

$$\mu_2 = E[T^2] = \int_{-\infty}^{+\infty} t^2 f_T(t) dt. \tag{A.7}$$

A solução da equação (A.7) é apresentada por Spiegel [56]. Portanto, temos:

$$\begin{aligned}
\mu_2 &= \int_0^{+\infty} \frac{t^2 \alpha \beta^\alpha dt}{(t + \beta)^{\alpha+1}} \\
&= \alpha \beta^\alpha \left[\lim_{b \rightarrow \infty} \left(\int_0^b \frac{t^2 dt}{(t + \beta)^{\alpha+1}} \right) \right] \\
&= \alpha \beta^\alpha \left[\lim_{b \rightarrow \infty} \left(\frac{(t + \beta)^2}{(2 - \alpha)(t + \beta)^\alpha} - \frac{2\beta(t + \beta)}{(1 - \alpha)(t + \beta)^\alpha} - \frac{\beta^2}{\alpha(t + \beta)^\alpha} \right) \Big|_0^b \right] \\
&= \alpha \beta^\alpha \left[\lim_{b \rightarrow \infty} \left(\frac{(b + \beta)^2}{(2 - \alpha)(b + \beta)^\alpha} - \frac{2\beta(b + \beta)}{(1 - \alpha)(b + \beta)^\alpha} \right) \right] + \frac{2\beta}{\alpha^2 - 3\alpha + 2}
\end{aligned} \tag{A.8}$$

O segundo momento para a variável aleatória T , depende diretamente do valor de α . Portanto, a solução da equação (A.8) pode assumir dois valores, que são:

$$\mu_2 = \begin{cases} \frac{2\beta^2}{\alpha^2 - 3\alpha + 2} & \alpha > 2 \\ +\infty & 0 < \alpha \leq 2 \end{cases} \tag{A.9}$$

Momento Central de Segunda Ordem – Variância

O segundo momento central da variável aleatória T , é dado por:

$$\sigma_T^2 = E[(T - \mu)^2] = \int_{-\infty}^{+\infty} (t - \mu)^2 f_T(t) dt \tag{A.10}$$

Para determinar o valor do segundo momento central da variável aleatória T , será desenvolvido o primeiro membro da equação (A.10). Portanto, temos:

$$\begin{aligned}
\sigma_T^2 &= E[(T - \mu)^2] \\
&= E[(T^2 - 2T\mu + \mu^2)] \\
&= E[T^2] - \mu^2
\end{aligned} \tag{A.11}$$

De acordo com a equação (A.11) o segundo momento central pode ser determinado a partir das equações (A.6) e (A.9), onde conclui-se que o segundo momento central assumirá duas soluções que são dependentes do valor de α , resultando em:

$$\sigma_T^2 = \begin{cases} \frac{2\beta^2}{\alpha^2 - 3\alpha + 2} - \frac{\beta^2}{(\alpha - 1)^2} & \alpha > 2 \\ +\infty & 0 < \alpha \leq 2 \end{cases} \tag{A.12}$$

Anexo B

Função de Autocorrelação de um Processo Estocástico Estacionário com fdp de Pareto

Este anexo apresenta o desenvolvimento da função de autocorrelação de um processo estocástico estacionário com função densidade de probabilidade de Pareto, apresentada no Anexo-A.

Função de Autocorrelação

$$\begin{aligned} R_T(t, t + \tau) &= E[T(t)T(t + \tau)] \\ &= E[T(t)T(t_1)] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} tt_1 f_{T(t)T(t_1)}(t, t_1) dt dt_1 \\ &= \lim_{b \rightarrow \infty} \int_0^b \int_0^b \frac{tt_1 \alpha^2 \beta^{2\alpha} dt dt_1}{(t + \beta)^{\alpha+1} (t_1 + \beta)^{\alpha+1}} \end{aligned} \tag{B.1}$$

$$\begin{aligned} R_T(t, t_1) &= \alpha^2 \beta^{2\alpha} \left\{ \lim_{b \rightarrow \infty} \left[\int_0^b \frac{t_1}{(t_1 + \beta)^{\alpha+1}} \times \right. \right. \\ &\quad \left. \left. \times \left(\int_0^b \frac{t dt}{(t + \beta)^{\alpha+1}} \right) dt_1 \right] \right\} \end{aligned} \tag{B.2}$$

$$\begin{aligned}
R_T(t, t_1) &= \alpha^2 \beta^{2\alpha} \left\{ \lim_{b \rightarrow \infty} \left[\int_0^b \frac{t_1}{(t_1 + \beta)^{\alpha+1}} \times \right. \right. \\
&\quad \left. \left. \times \left(\frac{(t + \beta)}{(1 - \alpha)(t + \beta)^\alpha} + \frac{\beta}{\alpha(t + \beta)^\alpha} \right) \Big|_0^b dt_1 \right] \right\} \\
&= \alpha^2 \beta^{2\alpha} \left(\lim_{b \rightarrow \infty} \left\{ \int_0^b \frac{t_1}{(t_1 + \beta)^{\alpha+1}} \times \right. \right. \\
&\quad \left. \left. \times \left[\frac{b + \beta}{(1 - \alpha)(b + \beta)^\alpha} - \left(\frac{\beta}{(1 - \alpha)\beta^\alpha} + \frac{\beta}{\alpha\beta^\alpha} \right) \right] dt_1 \right\} \right) \\
R_T(t, t_1) &= \alpha^2 \beta^{2\alpha} \left(\lim_{b \rightarrow \infty} \left\{ \int_0^b \frac{t_1}{(t_1 + \beta)^{\alpha+1}} \times \right. \right. \\
&\quad \left. \left. \times \left[\frac{b + \beta}{(1 - \alpha)(b + \beta)^\alpha} - \left(\frac{\beta}{(1 - \alpha)\beta^\alpha} + \frac{\beta}{\alpha\beta^\alpha} \right) \right] dt_1 \right\} \right) \\
&= \alpha^2 \beta^{2\alpha} \left(\lim_{b \rightarrow \infty} \left\{ \left[\frac{b + \beta}{(1 - \alpha)(b + \beta)^\alpha} - \right. \right. \right. \\
&\quad \left. \left. \left. - \left(\frac{\beta}{(1 - \alpha)\beta^\alpha} + \frac{\beta}{\alpha\beta^\alpha} \right) \right] \left(\frac{(t_1 + \beta)}{(1 - \alpha)(t_1 + \beta)^\alpha} + \frac{\beta}{\alpha(t_1 + \beta)^\alpha} \right) \Big|_0^b \right\} \right) \\
R_T(t, t + \tau) &= \alpha^2 \beta^{2\alpha} \left(\lim_{b \rightarrow \infty} \left\{ \left[\frac{b + \beta}{(1 - \alpha)(b + \beta)^\alpha} - \left(\frac{\beta}{(1 - \alpha)\beta^\alpha} + \frac{\beta}{\alpha\beta^\alpha} \right) \right] \times \right. \right. \\
&\quad \left. \left. \times \left(\frac{(t + \tau + \beta)}{(1 - \alpha)(t + \tau + \beta)^\alpha} + \frac{\beta}{\alpha(t + \tau + \beta)^\alpha} \right) \Big|_0^b \right\} \right) \\
R_T(\tau) &= \alpha^2 \beta^{2\alpha} \left(\lim_{b \rightarrow \infty} \left\{ \left[\frac{b + \beta}{(1 - \alpha)(b + \beta)^\alpha} - \left(\frac{\beta}{(1 - \alpha)\beta^\alpha} + \frac{\beta}{\alpha\beta^\alpha} \right) \right] \times \right. \right. \\
&\quad \left. \left. \times \left[\frac{(b + \tau + \beta)}{(1 - \alpha)(b + \tau + \beta)^\alpha} - \left(\frac{(\tau + \beta)}{(1 - \alpha)(\tau + \beta)^\alpha} + \frac{\beta}{\alpha(\tau + \beta)^\alpha} \right) \right] \right\} \right) \tag{B.3}
\end{aligned}$$

A soluo da equao (B.1)   dada por (B.3). Observa-se que o resultado de (B.3) possui dependencia direta do valor da vari vel α , portanto, temos duas solues para a equao (B.3) que so:

$$R_T(\tau) = \begin{cases} +\infty & \alpha \leq 1 \\ \frac{\beta^{\alpha+1}}{(1 - \alpha)} \left[\frac{\alpha(\tau + \beta) - (1 - \alpha)\beta}{(1 - \alpha)(\tau + \beta)^\alpha} \right] & \alpha > 1 \end{cases} . \tag{B.4}$$

Anexo C

Transformada de Laplace da fdp de Pareto

Neste anexo é feito o desenvolvimento do cálculo da transformada de Laplace para a distribuição de Pareto com parâmetro de Hurst.

Cálculo da Transformada de Laplace

Seja a fdp de Pareto da forma

$$f_T(t) = \begin{cases} \frac{\alpha}{\beta} \left(1 + \frac{t}{\beta}\right)^{-\alpha-1} & 0 < t < \infty \text{ e } \alpha, \beta > 0 \\ 0 & t < 0 \end{cases} \quad (\text{C.1})$$

então a transformada de Laplace é dada por,

$$\mathcal{L}\{f(t)\} = \int_0^{\infty} f_T(t)e^{-st} dt. \quad (\text{C.2})$$

Com o objetivo de identificar a solução para a equação (C.2), a equação (C.1) será reescrita da seguinte maneira,

$$\begin{aligned} f(t) &= \frac{\alpha}{\beta} \left(1 + \frac{t}{\beta}\right)^{-\alpha-1} \\ &= \frac{\alpha\beta^\alpha}{(\beta + t)^{-(\alpha+1)}} \\ &= \alpha\beta^\alpha(\beta + t)^{-(\alpha+1)}. \end{aligned} \quad (\text{C.3})$$

Substituindo a equação (C.3) na equação (C.2) e fazendo o seu desenvolvimento, tem-se,

$$\begin{aligned}\mathcal{L}\{f(t)\} &= \int_0^{\infty} \alpha\beta^{\alpha}(\beta+t)^{-(\alpha+1)}e^{-st}dt \\ &= \alpha\beta^{\alpha} \int_0^{\infty} (\beta+t)^{-(\alpha+1)}e^{-st}dt.\end{aligned}\tag{C.4}$$

Para resolver a equação (C.4), será feita a seguinte mudança de variável,

$$x = t + \beta \Rightarrow t = x - \beta.\tag{C.5}$$

Derivando a equação (C.5), tem-se

$$dx = dt.\tag{C.6}$$

Aplicando-se a mudança de variáveis apresentada pelas equações (C.5) e (C.6), e considerando também a mudança dos limites de integração imposto pela mudança de variáveis, a equação (C.5) pode ser reescrita da seguinte forma,

$$\begin{aligned}\mathcal{L}\{f(t)\} &= \alpha\beta^{\alpha} \int_{\beta}^{\infty} x^{-\alpha-1}e^{-s(x-\beta)}dx \\ &= \alpha\beta^{\alpha}e^{s\beta} \int_{\beta}^{\infty} x^{-\alpha-1}e^{-sx}dx.\end{aligned}\tag{C.7}$$

Empregado, novamente, o artifício de mudança de variáveis na equação (C.7), tem-se,

$$\mathcal{L}\{f(t)\} = \alpha(\beta s)^{\alpha}e^{s\beta} \int_{\beta s}^{\infty} z^{-\alpha-1}e^{-z}dz\tag{C.8}$$

para,

$$\begin{aligned}z = sx \quad \therefore \quad x &= \frac{z}{s} \\ dz = sdx \quad \therefore \quad dx &= \frac{dz}{s}\end{aligned}$$

De acordo com [55] a solução da integral na equação (C.8) pode ser expressa através de uma função gama incompleta, ou seja,

$$\mathcal{L}\{f(t)\} = \alpha(\beta s)^{\alpha}e^{s\beta}\Gamma(-\alpha, \beta s).\tag{C.9}$$

De acordo com Abramovitz [55] a função gama incompleta, pode ser representada através de serie de potências, para $|x| < \infty$, da seguinte forma,

$$\Gamma(a, x) = \Gamma(a) - e^{-x} x^a \Gamma(a) \sum_{n=0}^{\infty} \frac{x^n}{\Gamma(a + n + 1)}. \quad (\text{C.10})$$

Substituindo (C.10) na equação (C.9) a solução de (C.2) é dada por,

$$\mathcal{L}\{f(t)\} = \alpha \Gamma(-\alpha) \left[(\beta s)^\alpha e^{s\beta} - \sum_{n=0}^{\infty} \frac{(\beta s)^n}{\Gamma(n - \alpha + 1)} \right]. \quad (\text{C.11})$$

Referências Bibliográficas

- [1] Roberta R. Martine, *Basic Traffic Analysis*, Prentice Hall, Upper Saddle River, 1994.
- [2] G. Kotsis H. Hlavacs and C. Steinkellner, “Traffic source modeling”, Tech. Rep. TR-991001, Institute for Appl. Comp. Science and Inf. Systems, University of Vienna, 1999.
- [3] B. Melamed D. L. Jarferman and W. Willinger, *Stochastic Modeling of Traffic Processes*, CRC Press, in frontiers in queueing: models, methods and problems edition, 1996.
- [4] Douglas E. Comer, *Interligação Em Redes Com TCP/IP. Princípios, Protocolos e Arquitetura*, vol. I, Editora Campus, 1998.
- [5] L. Kleinrock and S. Lam, “Packet switching in a multiaccess broadcast channel: Performance evaluation”, *IEEE Transaction Communication*, vol. 23, no. 4, pp. 410–423, April 1975.
- [6] W. Leland M. Taqqu W. Willinger and D. Wilson, “On the self-similarity nature of ethernet traffic (extended version)”, *IEEE/ACM Trans. on Networking*, vol. 2, no. 1, pp. 1–15, February 1994.
- [7] M.S. Taqqu J. Beran, R. Sherman and W. Willinger, “Long-range dependence in variable-bit-rate video traffic”, *IEEE Transaction on Communications*, vol. 43, no. 2/3/4, pp. 1566–1579, Feb./Mar./Apr. 1995.
- [8] M. Garrett and W. Willinger, “Analysis, modeling and generation of self-similar VBR video traffic”, in *ACM Sigcom*, London, UK, 1994, pp. 269–280.
- [9] V. Paxson and S. Floyd, “Wide-area traffic: The failure of poisson modeling”, *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, 1995.
- [10] M. W. Crovella and A. Bestavros, “Explaining world wide web traffic self-similarity”, Tech. Rep. TR-95-015, Computer Science Department, Boston University, 1995.

-
- [11] M. Crovella and A. Bestavros, “Self-similarity in world wide web traffic: Evidence and possible causes”, *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835–846, 1997.
- [12] A. Adas, “Traffic models in broadband networks”, *IEEE Communications Magazine*, vol. 35, no. 7, pp. 82–89, 1997.
- [13] I. Norros, “A storage model with self-similar input”, *Queueing Systems*, vol. 16, pp. 387–396, 1994.
- [14] T. Daniels and C. Blondia, “A discrete-time ATM traffic model with long range dependence characteristics”, in *Proceedings of PMCCN’97*, Mar. 1997.
- [15] M. Zukerman R. Addie and T. Neame, “Performance of a single server queue with self-similar input”, in *Proc. IEEE Int. Conf. Commun.*, April 1995, vol. 43, pp. 1566–1579.
- [16] S. Robert and J. Y Le Boudec, “New models for pseudo self-similar traffic”, *Performance Evolution*, vol. 30, pp. 57–68, 1997.
- [17] Boris Tsybakov and Nicoals D. Georganas, “On self-similar traffic in atm queues: definitions, overflow probability bound, and cell delay distribution”, *IEEE/ACM Transactions on Networking (TON)*, vol. 5, no. 3, pp. 397–409, 1997.
- [18] M. Taqqu W. Willinger and A. Erramilli, “A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks”, in *Stochastic Networks: Theory and Applications*, Oxford Clarendon Press, Ed. F. P. Kelly, S. Zachary and I. Ziedins, 1996.
- [19] Henry J. Fowler and Will E. Leland, “Local area network traffic characteristics, with implications for broadband network congestion management”, *IEEE Journal of Selected Areas in Communications*, vol. 9, no. 7, pp. 1139–1149, 1991.
- [20] W. E. Leland M. S. Taqq W. Willinger and D. V. Wilson, “On the self-similar nature of Ethernet traffic”, in *ACM SIGCOMM*, Deepinder P. Sidhu, Ed., San Francisco, California, 1993, pp. 183–193.
- [21] A. Ruela and W. Kinsner, “A survey of traffic characterization techniques in telecommunicatios networks”, May 1996, IEEE Canadian Conference on Eletrical and Computer Engineering.
- [22] Walter Willinger, Murad S. Taqqu, Robert Sherman, and Daniel V. Wilson, “Self-similarity through high-variability: statistical analysis of Ethernet LAN

- traffic at the source level”, *IEEE/ACM Transactions on Networking*, vol. 5, no. 1, pp. 71–86, 1997.
- [23] F. Huebener J. M. Fernandez and D. Liu, “Queueing performance comparison for traffic models for internet traffic”, in *IEEE GlobeCom*, Sydney, November 1998.
- [24] Y. Fan and N. D. Georganas, “Performance analysis of ATM switches with self-similar input traffic”, *Systems and Computer Engineering J.*, vol. 1, 1997.
- [25] B. Tsybakov and N. Georganas, “On the self-similar traffic in ATM queues: Definitions, overflow probability bound and cell delay distributions”, *IEEE/ACM Transactions on Networking*, vol. 5, no. 3, pp. 397–409, June 1997.
- [26] B. Tsybakov N. Likhanov and N. Georganas, “Analysis of an ATM buffer with self-similar fractal input traffic”, in *Proceedings of IEEE INFOCOM’95*, Boston, MA, April 1995, pp. 985–992.
- [27] J. E. Diamond and A. S. Alfa, “Matrix analytical model of an ATM output with self-similar traffic”, *Performance Evaluation*, vol. 31, pp. 201–210, 1998.
- [28] J. Gordon, “Pareto process as a model of self-similar packet traffic”, in *Proc. IEEE Globecom’95*, 1995, pp. 2232–2236.
- [29] J. W. Lamperti, “Semi-stable processes”, *Trans. Amer. Math. Soc.*, vol. 104, pp. 62–78, 1962.
- [30] B. Mandelbrot, “Self-similar error clusters in communication systems and the concept of conditional stationarity”, *IEEE Transaction on Communication Technology*, pp. 71–90, 1965.
- [31] B. Mandelbrot and J. Van Ness, “Fractional brownian motions, fractional noises and applications”, *SIAM Review*, vol. 10, pp. 433–437, 1968.
- [32] P. Embrechts and M. Maejima, “An introduction to the theory of selfsimilar stochastic processes”, *International Journal of Modern Physics*, vol. B 14, pp. 1399–1420, 2000.
- [33] B. Tsybakov and N. D. Georganas, “Self-similar processes in communications networks”, *Transaction Information Theory*, vol. 44, no. 5, pp. 1713–1725, Sep 1998.
- [34] Paul Embrechts and Makoto Maejima, *Selfsimilar Processes*, Princeton Series in Applied Mathematics, 2 edition, 2002.

- [35] A. D. A. Miranda e A. Anzaloni, “Modelos de tráfego: Dependência de curta duração vs. dependência de longa duração”, in *XVIII Simpósio Brasileiro de Telecomunicações – SBrT’2000*, Setembro 2000.
- [36] E. M. Cappe J. C. Pesquet A. Petropolu and X. Yang, “Long-range dependence and heavy tail modeling for teletraffic data”, *IEEE Signal Processing Magazine*, vol. 19, no. 3, pp. 14–27, 2002.
- [37] B. Mandelbrot and M. S. Taqqu, “Robust R/S analysis of long run serial correlation”, in *Proceedings of 42nd Session of the International Statistical Institute*, Manila, Philippines, 1979, pp. 1–37.
- [38] H. E. Hurst, “Methods of long-term storage in reservoirs”, in *Proceedings of Institution Civil Engineers, Part*, 1955, pp. 519–577.
- [39] W. E. Leland and D. V. Wilson, “High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection”, in *Proc. Of IEEE INFOCOM’91*, FL, USA, April 1991, pp. 1360–1366.
- [40] M. Zukerman R. Addie and T. Neame, “Fractal traffic: Measurements, modelling and performance evaluation”, in *Proc. IEEE Infocom*, 1995, pp. 984–997.
- [41] R. Sherman W. Willinger, M. S. Taqqu and D. V. Wilson, “Self-similarity through high-variability: Statistical analysis of ethernet LAN traffic at the source level”, in *Proceedings SIGCOMM’95*, Boston, MA, 1995, pp. 100–113.
- [42] P. Abry and D. Veitch, “Wavelet analysis of long-range dependent traffic”, *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 2–15, 1998.
- [43] B. K. Ryo and S. B. Lowen, “Point process models for self-similar network traffic”, *Stochastic Models*, 1998.
- [44] Iowa State University Press, Ed., *Long-Range Dependence: A Review*. Iowa State Statistical Library, H. A. David and H. T. David, 1984.
- [45] W. Willinger A. Feldmann, A. C. Gilbert and T. G. Kurtz, “The changing nature of network traffic”, *Computer Communication Review*, vol. 28, no. 2, pp. 5–19, 1998.
- [46] Murad S. Taqqu, Walter Willinger, and Robert Sherman, “Proof of a fundamental result in self-similar traffic modeling”, *ACM SIGCOMM Computer Communication Review*, vol. 27, no. 2, pp. 5–23, 1997.

- [47] A. Feldmann A. Gilbert W. Willinger and T. Kurtz, “The changing nature of network traffic: Scaling phenomena”, *Computer Communication Review*, vol. 28, no. 2, April 1998.
- [48] B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco, 1982.
- [49] Kiseon Kim Younsuk Koh, “Loss probability behaviour of Pareto/M/1/K queue”, *IEEE Communications Letters*, vol. 7, no. 1, pp. 39–41, January 2003.
- [50] Tat-Keung Chan and V.O.K. Li., “Decomposition of network of queues with self-similar traffic.”, *GLOBECOM-98*, vol. 5, pp. 3001–3006, 1998.
- [51] Ernesto L. Andrade Neto Antônio M. Alberti e Leonardo S. Mendes, “Aplicação de modelos multifractais na simulação de redes ATM”, *XVIII Simpósio Brasileiro de Telecomunicações*, Setembro 2000.
- [52] Leonard Kleinrock, *Queueing Systems*, vol. I, John Wiley and Sons, New York, 1975.
- [53] Hermann de Meer Gunter Bolch, Stefan Greiner and Kishor S. Trivedi, *Queueing Networks and Markov Chains Modeling and Performance Evaluation with Computer Science Applications*, John Wiley and Sons, Inc., New York, 1998.
- [54] Marc J. Goovaerts and Nelson De Pril, “Survival probabilities based on pareto claim distributions”, *Astin Bulletin*, vol. 11, no. 2, pp. 154–157, 1980.
- [55] Milton Abramowitz and Irene A. Stegren, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., New York, 1970.
- [56] Ph. D Murray R. Spiegel, *Manual de Fórmulas e Tabelas Matemática*, McGraw-Hill do Brasil, São Paulo - Rio de Janeiro - Belo Horizonte, 1973.