

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Esquema de Escalonamento Baseado na Regularidade Local de Fluxos de Dados Internet

Autor: Christian Jorge
Orientador: Prof. Dr. Lee Luan Ling

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: **Telecomunicações e Telemática.**

Banca Examinadora

Prof. Dr. Lee Luan Ling DECOM/FEEC/UNICAMP
Prof. Dr. Dalton Soares Arantes DECOM/FEEC/UNICAMP
Prof. Dr. Paulo Cardieri DECOM/FEEC/UNICAMP

Campinas, 31 de janeiro de 2006

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

J768e Jorge, Christian
 Esquema de escalonamento baseado na regularidade local
de fluxos de dados Internet / Christian Jorge. --Campinas,
SP: [s.n.], 2006.

 Orientador: Lee Luan Ling
 Dissertação (Mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

 1. Telecomunicações – Tráfego. 2. Internet (Redes de
computação). 3. Wavelets (Matemática). 4. Kalman,
Filtragem de. I. Lee, Luan Ling. II. Universidade Estadual
de Campinas. Faculdade de Engenharia Elétrica e de
Computação. III. Título.

Título em Inglês: A stream scheduling scheme based on local regularity of Internet traffic

Palavras-chave em Inglês: Network traffic, Multifractals, Wavelets, Hölder exponent,
Kalman filter, Scheduling

Área de concentração: Telecomunicações e Telemática

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: Dalton Soares Arantes e Paulo Cardieri

Data da defesa: 31/01/2006

RESUMO

Nas redes de comunicações, a atual integração de vários tipos de serviços, cada qual com características estatísticas e requisitos de qualidade de serviço distintos, traz consigo a necessidade de esquemas eficientes de gerenciamento e controle de congestionamento do tráfego presente. Em pequenas escalas de tempo, os esquemas atuais podem ter sua eficiência reduzida devido à alta irregularidade do tráfego. Desta forma, neste presente trabalho, tendo como base a disciplina de escalonamento *Generalized Processor Sharing* (GPS), propõe-se um esquema de escalonamento de fluxos de dados que utiliza o expoente de Hölder pontual para caracterização local de cada fluxo. Para isso, propõe-se conjuntamente um estimador dinâmico destes expoentes e um preditor. Os expoentes de Hölder pontuais são estimados dinamicamente por meio do decaimento dos coeficientes wavelets em janelas de tempo. O preditor proposto possui características adaptativas e baseia-se no filtro de Kalman e no filtro de Mínimos Médios Quadrados Normalizado (*Normalized Least-Mean-Square* - NLMS). As avaliações realizadas mostram que este esquema de escalonamento contribui para o controle dinâmico preventivo no sentido de se obter uma menor perda de dados e um melhor uso da taxa de transmissão do enlace, em comparação com o GPS convencional.

Palavras-chave: Tráfego, Redes, Multifractais, Wavelets, Expoente de Hölder, Filtro de Kalman, Escalonamento.

ABSTRACT

Today network traffic is composed of many services with different statistical characteristics and quality of service requirements. This integration needs efficient traffic congestion control and management schemes. Dynamic and preventive schemes usually anticipate traffic conditions by means of a prediction process. Nevertheless, at fine-grained time scales, traffic exhibits strong irregularities and more complex scaling law that make this prediction process a non-trivial task. In this work we model network traffic flows as multifractal processes and introduce the pointwise Hölder exponent as an indicator of the local regularity degree. Also we propose a new traffic flow scheduling scheme based on the Generalized Processor Sharing (GPS) discipline that incorporate the pointwise Hölder exponent to locally characterize each data flow. For this end we explicitly present both dynamic pointwise Hölder exponent estimation and prediction mechanisms. The pointwise Hölder estimation is carried out dynamically based on the decay of the wavelet coefficients in the selected time windows. The proposed predictor is adaptive and implemented with both Kalman and Normalized Least Mean Squares (NLMS) filters. Experimental evaluations have validated the proposed scheduling scheme, resulting in low data loss rate and a better sharing of the network resources in comparison with the usual GPS scheme.

Keywords: Network traffic, Multifractals, Wavelets, Hölder exponent, Kalman filter, Scheduling.

AGRADECIMENTOS

Gostaria de agradecer ao prof. Dr. Lee Luan Ling por ter me admitido como seu orientando, ter acreditado no meu potencial e pelo apoio moral e técnico dispensado durante a execução deste trabalho.

Agradeço aos colegas do laboratório pelo auxílio e companheirismo. Em especial, gostaria de expressar minha gratidão ao amigo Flávio Henrique Teles Vieira pelo suporte, orientação e contribuição, indispensáveis à conclusão deste trabalho.

Aos meus pais, irmã e minha namorada Camila por terem acreditado no meu trabalho, pelo apoio e pela paciência durante este processo e principalmente por todo o amor e carinho.

SUMÁRIO

LISTA DE FIGURAS	1
LISTA DE TABELAS	3
GLOSSÁRIO	5
LISTA DE SÍMBOLOS	6
1 INTRODUÇÃO	7
2 ESTIMAÇÃO DO EXPOENTE DE HÖLDER PONTUAL	10
2.1 Introdução do capítulo.....	10
2.2 Caracterização da regularidade local.....	10
2.2.1 <i>Expoente de Hölder</i>	11
2.2.2 <i>Singularidades não-oscilantes vs. singularidades oscilantes</i>	13
2.2.3 <i>O espectro multifractal</i>	14
2.3 Estimação da regularidade local.....	15
2.3.1 <i>Análise wavelet</i>	15
2.3.2 <i>Estimação do espectro multifractal</i>	16
2.3.3 <i>Estimação via WTMM</i>	17
2.3.4 <i>Estimação do expoente de Hölder pontual para singularidades oscilantes ou não-oscilantes</i>	18
2.4 Avaliação do estimador de expoente de Hölder pontual utilizado.....	20
2.4.1 <i>Singularidade não-oscilante</i>	21
2.4.2 <i>Singularidade oscilante</i>	21
2.4.3 <i>Sinal monofractal</i>	22
2.4.4 <i>Sinal multifractal</i>	24
2.5 Estimação da regularidade local do tráfego Internet.....	25
2.5.1 <i>Estimação do expoente de Hölder pontual de referência</i>	26
2.5.2 Proposta: <i>estimação do expoente de Hölder pontual por meio de janelas de tempo</i>	28
2.6 Conclusões do capítulo.....	31
3 PREDIÇÃO DO EXPOENTE DE HÖLDER PONTUAL	32
3.1 Introdução do capítulo.....	32
3.2 Filtros adaptativos preditivos.....	34
3.2.1 <i>Filtro NLMS</i>	34
3.2.2 <i>Filtro de Kalman</i>	38

3.2.3 <i>Proposta: estimaco dinmica dos rudos do sistema</i>	40
3.3 Avaliaco do preditor proposto.....	41
3.3.1 <i>Avaliaco dos rudos do sistema</i>	43
3.3.2 <i>Avaliaco comparativa com outros preditores</i>	44
3.3.3 <i>Predico das caractersticas do trfego Internet</i>	47
3.4 Avaliaco conjunta da estimaco e predico dos expoentes de Hlder pontuais.....	49
3.5 Concluses do captulo.....	52
4 ESCALONAMENTO UTILIZANDO O EXPOENTE DE HLDER PONTUAL	54
4.1 Introduco do captulo.....	54
4.2 Escalonamento de fluxos de dados.....	54
4.2.1 <i>Conceito de escalonamento</i>	54
4.2.2 <i>Generalized Processor Sharing</i>	56
4.3 Proposta: esquema de escalonamento com estimaco e predico dos expoentes de Hlder pontuais	57
4.3.1 <i>GPS com expoentes de Hlder pontuais</i>	57
4.3.2 <i>Composio do escalonador proposto</i>	59
4.3.3 <i>Simulaes com o escalonador proposto</i>	62
4.3.4 <i>Resultados das simulaes</i>	63
4.3.5 <i>Anlise dos resultados das simulaes</i>	66
4.4 Concluses do captulo.....	67
5 CONCLUSES	68
REFERNCIAS	70
ARTIGOS PUBLICADOS	75

LISTA DE FIGURAS

2.1	Esquerda: exemplo de sinal e suas respectivas linhas de máximos. Direita: cone referente a uma linha de máximos.....	18
2.2	Esquerda: função $ x ^{0,7}$; Direita: “nuvem de pontos” e reta associados ao ponto com singularidade não-oscilante na função.....	21
2.3	Esquerda: função $ x ^{0,7} \text{sen}(1/ x ^{1,2})$; Direita: “nuvem de pontos” e reta associada ao ponto com singularidade oscilante na função.....	22
2.4	Acima à esquerda: fBM com $H=0,85$. Abaixo à esquerda: expoentes de Hölder pontuais estimados para o mesmo fBM. Acima à direita: espectro multifractal do mesmo fBM. Abaixo à direita: diferentes espectros multifractais para um fBM com $H=0,7$, computados com 2^{10} até 2^{16} amostras.....	23
2.5	Acima à esquerda: cascata binomial com $p=0,3$ e $n=14$; Acima à direita: Expoentes de Hölder pontuais estimados para a cascata; Embaixo: Espectro multifractal teórico da cascata.....	25
2.6	Acima à esquerda: amostras de tráfego da série <i>dec-pkt-1</i> na escala de tempo de 100 ms; Acima à direita: Expoentes de Hölder pontuais referentes às amostras citadas. Abaixo: Espectro multifractal.....	26
2.7	Acima à esquerda: amostras de tráfego da série <i>lbl-pkt-5</i> na escala de tempo de 100 ms; Abaixo à direita: Expoentes de Hölder pontuais referentes às amostras citadas. Abaixo: Espectro multifractal.....	27
2.8	Acima à esquerda: amostras de tráfego da série <i>lbl-pkt-5</i> na escala de tempo de 200 ms; Acima à direita: Expoentes de Hölder pontuais referentes às amostras citadas. Abaixo: Espectro multifractal.....	27
2.9	Utilização de janelas de tempo na estimação dos expoentes de Hölder pontuais de amostras de um sinal de tráfego.....	28
2.10	Expoentes de Hölder pontuais de amostras de três séries de tráfego Internet: <i>dec-pkt-1</i> na escala de tempo de 100 ms (acima à esquerda), <i>lbl-pkt-5</i> na escala de tempo de 100 ms (acima à direita) e <i>lbl-pkt-5</i> na escala de tempo de 200 ms (abaixo).....	30
3.1	Função de autocorrelação das amostras de uma série de tráfego e de seus respectivos expoentes de Hölder pontuais. Acima à esquerda: série <i>dec-pkt-1</i> na escala de tempo de 100 ms. Acima à direita: série <i>lbl-pkt-5</i> na escala de tempo de 100 ms. Abaixo: série <i>lbl-pkt-5</i> na escala de tempo de 200 ms.....	33
3.2	Filtro transversal.....	35
3.3	Representação gráfica do fluxo de um sistema linear dinâmico e discreto.....	38

3.4	Esquerda: função de autocorrelação das amostras de um dos ruídos contidos no vetor de ruído de processo $\eta_1(k)$. Direita: histograma dos valores das amostras do mesmo ruído.....	43
3.5	Esquerda: função de autocorrelação das amostras do ruído de medida $\eta_2(k)$. Direita: histograma dos valores das amostras do mesmo ruído.....	44
3.6	Função de correlação cruzada entre amostras do ruído $\eta_2(k)$ e amostras de um dos ruídos do vetor $\eta_1(k)$	44
3.7	Decaimento do erro quadrático de predição dos expoentes de Hölder pontuais de amostras da série <i>dec-pkt-1</i> na escala de tempo de 100 ms (acima à esquerda) e da série <i>lbl-pkt-5</i> nas escalas de tempo de 100 ms (acima à direita) e 200 ms (abaixo)..	47
3.8	Expoentes de Hölder pontuais de referência, expoentes de Hölder pontuais estimados e predição destes últimos, todos a partir da série <i>dec-pkt-1</i> . Acima à esquerda: escala de tempo de 100 ms e janela 13. Acima à direita: escala de tempo de 200 ms e janela 13. Abaixo: escala de tempo de 200 ms e janela 12.....	50
3.9	Expoentes de Hölder pontuais de referência, expoentes de Hölder pontuais estimados e predição destes últimos, todos a partir da série <i>lbl-pkt-5</i> . Acima à esquerda: escala de tempo de 100 ms e janela 13. Acima à direita: escala de tempo de 200 ms e janela 13. Abaixo: escala de tempo de 200 ms e janela 12.....	51
4.1	Sistema GPS com n filas.....	56
4.2	Tarefas realizadas pela unidade de predição dos expoentes de Hölder pontuais.....	59
4.3	Tarefas realizadas pela unidade de predição dos expoentes de Hölder pontuais.....	60
4.4	Tarefas realizadas pela unidade de escalonamento dos fluxos.....	61
4.5	Esquema completo do escalonador proposto.....	62

LISTA DE TABELAS

2.1	EQMN dos expoentes de Hölder pontuais das séries de tráfego na escala de tempo de 100 ms.....	30
2.2	EQMN dos expoentes de Hölder pontuais das séries de tráfego na escala de tempo de 200 ms.....	30
3.1	EQMNs para os expoentes de Hölder pontuais da série de tráfego <i>dec-pkt-1</i> , na escala de tempo de 100 ms, estimados usando janela 13.....	45
3.2	EQMNs para os expoentes de Hölder pontuais da série de tráfego <i>dec-pkt-1</i> , na escala de tempo de 200 ms, estimados usando janela 13.....	45
3.3	EQMNs para os expoentes de Hölder pontuais da série de tráfego <i>dec-pkt-1</i> , na escala de tempo de 200 ms, estimados usando janela 12.....	46
3.4	EQMNs para os expoentes de Hölder pontuais da série de tráfego <i>lbl-pkt-5</i> , na escala de tempo de 100 ms, estimados usando janela 13.....	46
3.5	EQMNs para os expoentes de Hölder pontuais da série de tráfego <i>lbl-pkt-5</i> , na escala de tempo de 200 ms, estimados usando janela 13.....	46
3.6	EQMNs para os expoentes de Hölder pontuais da série de tráfego <i>lbl-pkt-5</i> na escala de tempo de 200 ms, estimados usando janela 12.....	46
3.7	EQMNs para predição de amostras da série de tráfego <i>dec-pkt-1</i> , na escala de tempo de 100 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.....	48
3.8	EQMNs para predição de amostras da série de tráfego <i>dec-pkt-1</i> , na escala de tempo de 200 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.....	48
3.9	EQMNs para predição de amostras da série de tráfego <i>dec-pkt-2</i> , na escala de tempo de 100 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.....	48
3.10	EQMNs para predição de amostras da série de tráfego <i>dec-pkt-2</i> , na escala de tempo de 200 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.....	49
3.11	EQMNs para predição de amostras da série de tráfego <i>lbl-pkt-5</i> , na escala de tempo de 100 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.....	49
3.12	EQMNs para predição de amostras da série de tráfego <i>lbl-pkt-5</i> , na escala de tempo de 200 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.....	49

3.13	EQMNs dos expoentes de Hölder pontuais das séries de tráfego na escala de tempo de 100 ms.....	52
3.14	EQMNs dos expoentes de Hölder pontuais das séries de tráfego na escala de tempo de 200 ms.....	52
4.1	Características dos fluxos.....	64
4.2	Performance do Escalonador proposto para taxa de 2,25 Mbps e <i>buffer</i> de 16 Kbytes.....	65
4.3	Performance do Escalonador proposto para taxa de 2,25 Mbps e <i>buffer</i> de 25 Kbytes.....	65
4.4	Performance do Escalonador proposto para taxa de 2,75 Mbps e <i>buffer</i> de 25 Kbytes.....	65
4.5	Características dos Fluxos.....	65
4.6	Performance do Escalonador proposto para taxa de 2,25 Mbps e <i>buffer</i> de 16 Kbytes.....	65
4.7	Performance do Escalonador proposto para taxa de 2,25 Mbps e <i>buffer</i> de 25 Kbytes.....	66
4.8	Performance do Escalonador proposto para taxa de 2,75 Mbps e <i>buffer</i> de 25 Kbytes.....	66

GLOSSÁRIO

DEC	Digital Equipment Corporation
EQM	Erro Quadrático Médio
EQMN	Erro Quadrático Médio Normalizado
fBM	fractional Brownian Motion
FCFS	First-Come First-Served
GPS	Generalized Processor Sharing
IP	Internet Protocol
Kbps	Kilobits por segundo
LAN	Local Area Network
LBL	Lawrence Berkeley Laboratory
LMS	Least Mean Squares
Mbps	Megabits por segundo
ms	milissegundos
NLMS	Normalized Least Mean Squares
QoS	Quality of Service
TCP	Transmission Control Protocol
WAN	Wide Area Network
WFQ	Weighted Fair Queueing
WF ² Q	Worst-case Fair Weighted Fair Queueing
WTMM	Wavelet Transform Modulus Maxima

LISTA DE SÍMBOLOS

$\ \cdot \ _1$	norma um.
$\ \cdot \ _2$	norma dois ou norma euclidiana.
∂^m	operador diferencial de ordem m .
$B(x_0, \rho)$	bola aberta com centro em x_0 e raio ρ .
$C^\alpha(x_0)$	classe de funções com α derivadas contínuas no ponto x_0 .
$d_H(\cdot)$	dimensão de Hausdorff
$E[\cdot]$	operador esperança matemática
$Tr[\cdot]$	operador traço

1 INTRODUÇÃO

A atual integração de serviços nas redes de comunicações traz consigo a necessidade de gerenciamento de vários tipos de tráfego, tais como tráfego de dados, de voz e de vídeo. Os tráfegos gerados por esses serviços podem ser substancialmente diferentes quanto às suas características estatísticas e requisitos para a manutenção da qualidade de serviço necessária (*QoS*). Por exemplo, o tráfego de voz pode necessitar de uma taxa de transmissão de alguns kilobits por segundo (Kbps) e ser sensível ao atraso, enquanto o tráfego de dados de alta velocidade, utilizado para transferência de arquivos e interconexão de LANs e WANs, pode necessitar de centenas de megabits por segundo (Mbps) para a transmissão e ser sensível à perdas.

Com o intuito de suportar os requisitos de cada tipo de serviço, comumente utilizam-se mecanismos de gerenciamento e controle de congestionamento do tráfego presente. Recomenda-se que estes mecanismos busquem um adequado equilíbrio entre a utilização de recursos da rede e a provisão da qualidade de serviço. Políticas de gerenciamento e controle muito conservativas podem resultar em desperdício de recursos e baixa utilização total da rede. Por outro lado, uma elevada utilização da rede, resultante do aumento do número de usuários atendidos, pode deteriorar a qualidade de serviço prestado, causando insatisfação aos próprios usuários [OST01].

Mecanismos estáticos de controle normalmente dependem de um modelo de tráfego para a determinação do comportamento estatístico real do tráfego. Tais mecanismos podem se tornar altamente ineficientes em situações nas quais este comportamento não possa ser determinado ou quando há discordância entre os parâmetros de tráfego declarados pelo usuário e os parâmetros de tráfego reais. [SIR03].

Mecanismos dinâmicos (ou adaptativos) de controle são uma alternativa aos mecanismos estáticos [SIR03]. Como obtêm de maneira *on-line* as informações necessárias para o controle de tráfego, os mecanismos dinâmicos não precisam considerar um modelo de tráfego, nem se preocupar com o impacto causado por parâmetros declarados de maneira imprecisa pelo usuário. Tais mecanismos podem ser divididos em duas categorias: reativa e preventiva [HAL00] [SIR03]. O controle reativo, ou de malha fechada, depende de sinais de

realimentação (*feedback*) da rede para o ajuste da taxa de envio dos pacotes pela fonte, conforme a indicação do nível de congestionamento. O nível de congestionamento é determinado comparando-se o comprimento da fila no *buffer* do multiplexador com um determinado limiar (*threshold*). Uma desvantagem deste tipo de mecanismo de controle é a dificuldade para determinar-se o valor ótimo deste limiar. Outra desvantagem é a possível perda de dados por parte das fontes, antes que cada uma receba seu respectivo sinal de realimentação. Esta desvantagem pode ocorrer principalmente em redes de alta velocidade.

Em contraste, mecanismos de controle preventivo, ou de malha aberta, alocam recursos com antecedência, a fim de se evitar um possível congestionamento. Nestes mecanismos, o preditor das características do tráfego desempenha um papel fundamental [OST01]. Por exemplo, um mecanismo de escalonamento aliado a um esquema de predição pode prevenir que algum fluxo mal-comportado e irregular receba um valor da taxa de transmissão compartilhada acima do valor considerado justo a ele.

Para aplicações em tempo real, o processo de predição deve ser feito rapidamente, a partir de uma mínima quantidade de informações armazenada. O intervalo de predição deve considerar o tempo de processamento das informações e outras latências. Além disso, para possibilitar uma alocação correta dos recursos, a estimativa deve ser precisa o suficiente [OST01]. Infelizmente, a precisão das estimativas diminui conforme se aumenta o intervalo de predição [SAN00]. Claramente há um compromisso entre um grande intervalo de predição e um pequeno erro na estimativa. Isto reflete outro compromisso, entre a escala de tempo de controle e a eficiência da rede [SAN00] [SHA03].

Outro fator que pode influenciar o desempenho da predição é a própria natureza do sinal analisado [OST01] [QIA04]. Em se tratando de sinais de tráfego de redes, ocorreram mudanças significativas no modo de compreendê-los na última década. Inicialmente, houve a descoberta da propriedade de invariância à escala (*scaling*), do tráfego de pacotes, em grandes escalas de tempo [LEL94] [PAX95]. Sabe-se que este fenômeno é resultante da agregação de fluxos com “cauda pesada” (*heavy-tailed*) [PAR00], que por sua vez são oriundos de arquivos cujos tamanhos possuem distribuições também com cauda pesada [CRO97]. Nesta situação, o tráfego costuma ser modelado como um processo monofractal, cuja lei de escalas é determinada pelo parâmetro de Hurst. Este parâmetro não apenas caracteriza a regularidade global de um sinal, como também o grau das rajadas de dados (*burstiness*) do tráfego.

A descoberta do tráfego de redes como um processo multifractal, uma forma mais significativa (e complexa) de processo invariante à escala, tornou a modelagem das WANs mais robusta e completa [RIE97]. Nesta situação, devido ao fato de os processos multifractais apresentarem regularidade local variante no tempo, costuma-se substituir o parâmetro de Hurst pelo expoente de Hölder. Em [FEL98] tornou-se claro pela primeira vez a evidência do comportamento invariante à escala do tráfego TCP/IP em dois regimes separados. Empiricamente, determinou-se que o limite entre estes dois regimes estava na faixa de algumas centenas de milissegundos (ms). Desta forma, concluiu-se que o tráfego possuía comportamento multifractal em pequenas escalas (regendo frequências mais altas) e com dependência de longo prazo (auto-similaridade assintótica de segunda ordem) em escalas maiores (regendo frequências mais baixas). A causa sugerida para este fenômeno é a redistribuição hierárquica dos dados pelo protocolo TCP/IP.

Em termos gerais, a análise da regularidade local é importante tanto para o gerenciamento e controle de congestionamento da rede, quanto para políticas de alocação de recursos de maneira justa. Em análise de redes, sabe-se que um tráfego de dados com alto grau de rajadas proporciona um menor aproveitamento dos recursos [ABR02]. Desta forma, como contribuição ao gerenciamento e controle preventivo do tráfego de redes, propomos por meio deste trabalho, um mecanismo de escalonamento de fluxos de dados que utiliza a regularidade local de cada fluxo presente, como um meio de se obter uma melhor distribuição da taxa de transmissão de um enlace e uma menor perda de dados. Para atingirmos este objetivo, inicialmente, no Capítulo 2, propomos um esquema dinâmico para estimação da regularidade local de um dado fluxo. Em seguida, no Capítulo 3, um preditor adaptativo baseado no filtro de Kalman e no filtro de Mínimos Médios Quadrados Normalizado (*Normalized Least-Mean-Square* - NLMS) é proposto, com o intuito de estimarmos a regularidade local no próximo instante de tempo para cada fluxo. No Capítulo 4, tanto o estimador dinâmico como o preditor citados são incorporados num esquema de escalonamento, modelado com base na disciplina de escalonamento *Generalized Processor Sharing* (GPS). Por último, no Capítulo 5, são apresentadas as conclusões e possíveis extensões a este trabalho.

2 ESTIMAÇÃO DO EXPOENTE DE HÖLDER PONTUAL

2.1 Introdução do capítulo

Este capítulo tem como proposta a estimação dos expoentes de Hölder pontuais de amostras de sinais de tráfego de redes. Nosso objetivo é a utilização destes expoentes como um indicador das características locais de um dado sinal de tráfego.

Para atingirmos este objetivo, inicialmente na Seção 2.2, caracterizamos a regularidade local de um sinal por meio dos expoentes de Hölder. Em seqüência, na Seção 2.3, apresentamos alguns métodos de estimação dos expoentes de Hölder pontuais para sinais. Na Seção 2.4 estimamos os expoentes de Hölder pontuais de amostras de alguns tipos de sinais. Na Seção 2.5, aplicamos o método de estimação destes expoentes para amostras de sinais de tráfego Internet. Nesta mesma seção, propomos um estimador com caráter dinâmico. Por último, na Seção 2.6, apresentamos as conclusões obtidas.

2.2 Caracterização da regularidade local

Em análise de sinais, algumas vezes nos deparamos com situações nas quais não nos importa o comportamento regular e suave de um determinado sinal e sim as singularidades presentes, tais como rápidas mudanças, picos, descontinuidades e transientes em frequências.

Em uma imagem, as descontinuidades em sua intensidade podem nos indicar a localização do contorno do objeto, o que é útil para a área de reconhecimento de padrões [VOO03]. Em outros tipos de sinais, desde sinais de eletrocardiograma até sinais de fala, a informação de interesse está contida nas singularidades presentes [DAO95] [WES04]. Particularmente para sinais de tráfego de redes de computadores, o grau da sua regularidade está intimamente ligado ao grau das rajadas de dados (*burstiness*) [KRI02] [SEU00].

Existem várias maneiras de se medir a regularidade de um sinal. Um método muito utilizado é de natureza geométrica e trata de encontrar a dimensão fractal do gráfico de uma função. De uma maneira grosseira, precisamos determinar como esse gráfico preenche o espaço em pequenas escalas. A precisão deste método é muito sensível à dimensão fractal

utilizada, sendo as mais freqüentes a de Hausdorff, a da caixa e a de Tricot [FAL90]. Neste trabalho, consideramos os métodos baseados nos espaços de Hölder. Estes são os métodos mais utilizados para a caracterização da regularidade local de um sinal e consistem na utilização do expoente de Hölder local ou, em especial, o expoente de Hölder pontual. A definição formal dos expoentes de Hölder é apresentada a seguir.

2.2.1 Expoente de Hölder

Conforme citado anteriormente, tanto o expoente de Hölder local como o expoente de Hölder pontual podem servir como quantificadores do grau da regularidade local de um sinal. O tipo de expoente escolhido depende da aplicação a ser feita. Antes de definirmos estes dois tipos de expoentes, inicialmente precisamos definir alguns conceitos oriundos da Análise Funcional, para uma melhor compreensão:

Definição 2.2.1.1: Seja $\Omega \subset \mathbb{R}^n$, $k > 0$ e a função $f : \Omega \rightarrow \mathbb{R}$. $C^k(\Omega)$ define o conjunto de funções f que são k vezes diferenciáveis, com derivadas contínuas.

Definição 2.2.1.2 (Espaço métrico): Um espaço métrico é um par ordenado (X, ρ) em que X é um conjunto não-vazio com uma função $\rho : X \times X \rightarrow [0, \infty)$, satisfazendo:

$$\text{i) (não-degeneração) } \rho(x, y) = 0 \Leftrightarrow x = y \quad (2.1)$$

$$\text{ii) (simetria) } \rho(x, y) = \rho(y, x) \text{ para todo } x, y \in X \quad (2.2)$$

$$\text{iii) (desigualdade triangular) } \rho(x, z) \leq \rho(x, y) + \rho(y, z), \text{ para todo } x, y, z \in X \quad (2.3)$$

A função ρ é chamada uma métrica em X e o número real não-negativo $\rho(x, y)$ é chamado de distância de x a y .

Definição 2.2.1.3: Seja (X, ρ) um espaço métrico, r um número real estritamente positivo e $x, x_0 \in X$. O conjunto

$$B(x_0, r) = \{x : \rho(x_0, x) < r\} \quad (2.4)$$

é chamado de bola aberta com centro em x_0 e raio r .

Definição 2.2.1.4 (Expoente de Hölder local): Seja α um número real estritamente positivo, $x_0 \in \mathbb{R}$, e uma função $f: \mathbb{R} \rightarrow \mathbb{R}$. Caso $m < \alpha < m+1$ ($m \in \mathbb{N}$), podemos dizer que $f \in C^\alpha(B(x_0, r))$, se existir uma constante K tal que, para todo x, y em $B(x_0, r)$,

$$|\partial^m f(x) - \partial^m f(y)| \leq K |x - y|^{\alpha - m} \quad (2.5)$$

na qual ∂^m é um operador diferencial de ordem m . Particularmente, caso $0 < \alpha < 1$, podemos dizer que $f \in C^\alpha(B(x_0, r))$, se existir uma constante K tal que, para todo x, y em $B(x_0, r)$,

$$|f(x) - f(y)| \leq K |x - y|^\alpha \quad (2.6)$$

Então seja

$$\alpha_l(x_0, B(x_0, r)) = \sup\{\alpha : f \in C^\alpha(B(x_0, r))\} \quad (2.7)$$

Note que $\alpha_l(x_0, B(x_0, r))$ é não crescente em função de r . O expoente de Hölder local α_l de f em x_0 é definido como

$$\alpha_l(x_0) = \lim_{r \rightarrow 0} \alpha_l(x_0, B(x_0, r)) \quad (2.8)$$

Este tipo de expoente de Hölder é estável sob ação de operadores (pseudo)diferenciais ou integradores. No entanto, sua principal desvantagem é que a função de Hölder local, $x \rightarrow \alpha_l(f, x)$, é uma função semi-contínua inferior, ou seja,

$$\forall x_0 \in \mathbb{R}, \forall \varepsilon, \exists \eta :$$

$$y \in B(x_0, \eta) \Rightarrow \alpha_l(f, y) > \alpha_l(f, x_0) - \varepsilon \quad (2.9)$$

Segundo [HOB58], duas funções semi-contínuas inferiores, coincidentes em um conjunto totalmente denso, são iguais. Assim, por exemplo, não é possível observar um ponto isolado “regular” em um ambiente “irregular” por meio do expoente de Hölder local. [DAO98] [GUI98]. Uma solução alternativa e mais versátil é considerarmos o expoente de Hölder pontual.

Definição 2.2.1.5 (Expoente de Hölder pontual): Seja α um número real estritamente positivo, K uma constante e $x_0 \in \mathbb{R}$. A função $f: \mathbb{R} \rightarrow \mathbb{R}$ é $C^\alpha(x_0)$ se existe um polinômio P_n , de grau $n < \alpha$, tal que

$$|f(x) - P_n(x - x_0)| \leq K |x - x_0|^\alpha \quad (2.10)$$

O expoente de Hölder pontual α_p da função f em x_0 é definido como

$$\alpha_p(x_0) = \sup\{\alpha > 0 \mid f \in C^\alpha(x_0)\} \quad (2.11)$$

Esta caracterização da regularidade de uma função é amplamente utilizada em Análise Fractal, tanto na teoria como na prática. Por exemplo, em [SEU00], mostra-se que o expoente de Hölder pontual pode quantificar o grau da variação instantânea de um sinal de tráfego de redes. Mais precisamente, este expoente pode indicar o grau das rajadas de dados presentes neste sinal.

De um ponto de vista teórico, funções de Hölder pontuais (ou seja, funções do tipo $x \rightarrow \alpha_p(f, x)$), são funções contínuas inferiores [DAO98]. Isto permite a aplicação destas funções em uma grande variedade de situações.

2.2.2 Singularidades não-oscilantes vs. singularidades oscilantes

Para uma correta descrição e estimação da regularidade local de um sinal, precisamos considerar a possível presença de oscilações em torno das singularidades. Desta forma, apresentamos nesta subsecção, duas categorias de singularidades relevantes para a caracterização da regularidade local de um sinal: a singularidade não-oscilante (*cusps*) e a singularidade oscilante (*chirp*).

Definição 2.2.2.1 (Singularidade não-oscilante): Seja $f_\alpha^{(-n)}$ a primitiva de n -ésima ordem da função f_α . Dizemos que f_α possui uma singularidade não-oscilante, com expoente de Hölder pontual α em x_0 , se

$$\forall n \in \mathbb{N} : f_\alpha^{(-n)} \in C^{\alpha+n}(x_0) \quad (2.12)$$

Desta forma, a função f_α deve ser regular o suficiente, para que dado um ponto x_0 com expoente de Hölder α , a derivada de f_α apresente expoente de Hölder $\alpha-1$ em x_0 e a integral de f_α apresente expoente de Hölder $\alpha+1$ no mesmo ponto. De uma maneira grosseira, podemos dizer que este ponto ganha um grau de suavidade na integração e perde um grau de suavidade na diferenciação [SEU00]. A função $f_\alpha(x) = |x - x_0|^\alpha$ representa o exemplo mais simples de singularidade não-oscilante em x_0 .

Definição 2.2.2.2 (Singularidade oscilante): seja $g_{\alpha,\beta}^{(-n)}$ a primitiva de n-ésima ordem da função $g_{\alpha,\beta}$. Dizemos que $g_{\alpha,\beta}$ possui uma singularidade oscilante em x_0 , com expoente de Hölder pontual α e expoente de oscilação $\beta > 0$, se

$$\forall n \in \mathbb{N} : g_{\alpha,\beta}^{(-n)} \in C^{\alpha+n(1+\beta)}(x_0) \quad (2.13)$$

Este tipo de singularidade ocorre em funções cujo expoente de Hölder pontual em x_0 aumenta mais que a unidade quando integradas. Isto se deve a efeitos de suavização das oscilações presentes, causados pela operação de integração da função [MAL92]. A função $g_{\alpha,\beta}(x) = |x - x_0|^\alpha \operatorname{sen}\left(\frac{1}{|x - x_0|^\beta}\right)$ representa o exemplo mais simples de singularidade oscilante em x_0 .

2.2.3 O Espectro multifractal

O espectro multifractal (ou espectro de singularidades) provê informações sobre quais singularidades ocorrem em um dado sinal e quais singularidades predominam. O espectro é uma curva unidimensional, normalmente com perfil côncavo, cuja abcissa representa os expoentes de Hölder pontuais que efetivamente existem no sinal, e cuja ordenada está relacionada com a quantidade de pontos em que uma dada singularidade é encontrada.

Para definirmos o espectro multifractal, primeiro precisamos definir a **dimensão de Hausdorff** [FAL90]. Assim, seja $\{U_i\}$ uma coleção finita de conjuntos com diâmetro de no máximo δ , que cobre um subconjunto F em \mathbb{R}^n . Neste caso, chamamos $\{U_i\}$ de coleção de conjuntos de cobertura δ de F .

Definição 2.2.3.1 : Seja $|U_i|$ o diâmetro do conjunto U_i . Define-se $H_\delta^s(F)$ para algum $\delta > 0$ como

$$H_\delta^s(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ seja uma coleção de conjuntos de cobertura } \delta \text{ de } F \right\}. \quad (2.14)$$

A medida de Hausdorff s -dimensional de F é definida como o limite de $H_\delta^s(F)$ quando δ tende a zero. Ou seja,

$$H^s(F) = \lim_{\delta \rightarrow 0} H_\delta^s(F). \quad (2.15)$$

O limite que define a medida de Hausdorff existe para qualquer subconjunto F em \mathbb{R}^n , sendo seu valor usualmente igual a 0 ou ∞ .

Definição 2.2.3.2: O valor crítico de s para o qual a medida de Hausdorff s -dimensional $H^s(F)$ muda instantaneamente de ∞ para 0 é definido como dimensão de Hausdorff de F .

Definição 2.2.3.3: Seja f uma função: $[a, b] \rightarrow \mathbb{R}$, $a < b$ e seja $\alpha(x)$ o expoente de Hölder pontual de f em cada ponto $x \in [a, b]$. O espectro multifractal $D(\alpha)$ de f é definido como

$$D(\alpha) = d_H(\{x \mid \alpha(x) = \alpha\}), \quad (2.16)$$

em que d_H denota a dimensão de Hausdorff.

A dimensão de Hausdorff descreve precisamente a geometria fractal de um conjunto fractal composto por pontos com um mesmo grau de singularidade α . Assim, o espectro multifractal provê a informação geométrica pertinente à dimensão fractal dos conjuntos de pontos que possuem um dado expoente de Hölder pontual α em um sinal.

2.3 Estimação da regularidade local

Nesta seção apresentamos alguns estimadores da regularidade local de um sinal. Nossa intenção é caracterizar quantitativamente a regularidade de um determinado sinal em cada amostra presente. Para isto, consideramos o expoente de Hölder pontual como quantificador numérico do grau da regularidade local. A estimação deste expoente é feita por meio do decaimento do valor absoluto dos coeficientes wavelet do sinal analisado. Para isto, faremos uma breve apresentação da transformada wavelet a seguir.

2.3.1 Análise wavelet

A transformada wavelet é uma ferramenta poderosa para caracterização da regularidade local de um sinal. Esta é a maior motivação para estudá-la na matemática [MEY90] e em suas aplicações [GRO84]. Conceitualmente, a transformada wavelet é um produto-convolução do sinal analisado com a wavelet mãe ψ . Neste processo, a wavelet-mãe deve ser ajustada a uma determinada escala j e transladada até um ponto $2^j k$ do sinal, com $j, k \in \mathbb{Z}$. Desta forma, o coeficiente wavelet $d_{j,k}$ é dado por [DAU92]

$$d_{j,k} = 2^{-j} \int_{-\infty}^{\infty} f(x) \psi(2^{-j}x - k) dx \quad (2.17)$$

O uso do fator 2^j deve-se ao fato de obtermos $\|\psi_{j,k}\|_1 = \|\psi\|_1$, em que $\|\cdot\|_1$ é a norma um. Com o aumento do valor de $d_{j,k}$ temos um aumento da variação do sinal a uma distância 2^j , a partir do ponto $2^j k$.

Uma importante propriedade de uma wavelet $\psi(x)$ é a sua quantidade de **momentos de desvanecimento** (*vanishing moments*). Pode-se dizer que $\psi(x)$ possui N momentos de desvanecimento se e somente se, para $i=0, \dots, N-1$, satisfaz-se:

$$\int_{-\infty}^{\infty} x^i \psi(x) dx = 0 \quad (2.18)$$

Por meio desta propriedade, nota-se que as regiões de um sinal que podem ser aproximadas por um polinômio de grau máximo igual a N , possuem os coeficientes wavelets nulos. Esta propriedade é importante para a caracterização das singularidades de um sinal, por meio do decaimento do valor absoluto dos coeficientes wavelet [MAL92].

2.3.2 Estimação do espectro multifractal

Uma maneira muito utilizada para a estimação do espectro multifractal de um sinal com suporte compacto consiste na aplicação da transformada de Legendre. Para isso, inicialmente define-se a **função de partição** do sinal analisado. A função de partição $S(q, j)$ é definida utilizando-se os coeficientes wavelets do sinal, pela seguinte equação

$$S(q, j) = \sum_k |d_{j,k}|^q \quad (2.19)$$

Definindo a **função estrutura** $\tau(q)$ como

$$\tau(q) = \lim_{j \rightarrow \infty} \frac{\log S(q, j)}{j \log 2}, \quad (2.20)$$

o espectro multifractal $D(\alpha)$ do sinal analisado é definido como

$$D(\alpha) = \tau^*(\alpha), \quad (2.21)$$

em que $\tau^*(\alpha)$ é a transformada inversa de Legendre da função estrutura dada por $\tau^*(\alpha) = \min_q (\alpha q - \tau(q))$.

Neste trabalho, o espectro multifractal é estimado utilizando-se o *FracLab*¹, para verificação das estimativas dos expoentes de Hölder pontuais. O *FracLab* é uma *toolbox* do *MatLab* e é largamente utilizada para análise de sinais (multi)fractais.

2.3.3 Estimação via WTMM

O método dos Máximos em Módulos da Transformada Wavelet (*Wavelet Transform Modulus Maxima* - WTMM) é largamente utilizado para detecção de singularidades em sinais [STR00]. O WTMM é definido da seguinte forma [MAL92]:

Definição 2.3.3.1 : Sejam $d_{j,k}$'s os coeficientes wavelet :

- Chamamos de **máximos em módulo**, qualquer ponto (j_0, k_0) tal que $|d_{j_0,k}| < |d_{j_0,k_0}|$, na qual k pertence tanto à vizinhança à esquerda ou à direita de k_0 .
- Chamamos de **linha de máximos**, qualquer curva conectada no espaço-escala (j,k) formada por pontos que são máximos em módulo.

Na realidade, a linha de máximos contém os pontos mais significativos dentre aqueles que estão imersos em um cone invertido, definido por

$$|2^j k - x_0| \leq K 2^j \quad (2.22)$$

na qual K é uma constante. Dentro deste cone, a linha de máximos converge para a singularidade presente no ponto $x_0 = 2^j k_0$ de um sinal, obedecendo a seguinte relação:

$$|d_{j,k}| \leq A 2^{j\alpha} \quad (2.23)$$

ou, equivalentemente:

$$\log |d_{j,k}| \leq \log A + \alpha \log(2^j) \quad (2.24)$$

¹ <http://www.irccyn.ec-nantes.fr/hebergement/FracLab/>

nas quais A é uma constante.

A desigualdade (2.24) mostra que o expoente de Hölder pontual α em x_0 é o maior coeficiente angular das retas que estão acima de $|d_{j,k}|$, na escala logarítmica. A Figura 2.1, à esquerda, mostra várias linhas de máximos referentes ao sinal exibido. Cada uma destas linhas converge para uma singularidade presente. Na mesma figura, à direita, podemos também observar uma representação do cone, referente a uma das linhas de máximos, e de alguns pontos que são máximos em módulo.

Apesar de possuir uma abordagem relativamente simples, a estimação via WTMM possui duas principais limitações: somente pode ser aplicada eficientemente em sinais cujas singularidades sejam isoladas e de natureza não-oscilante. Uma alternativa a essas duas limitações é apresentada a seguir.

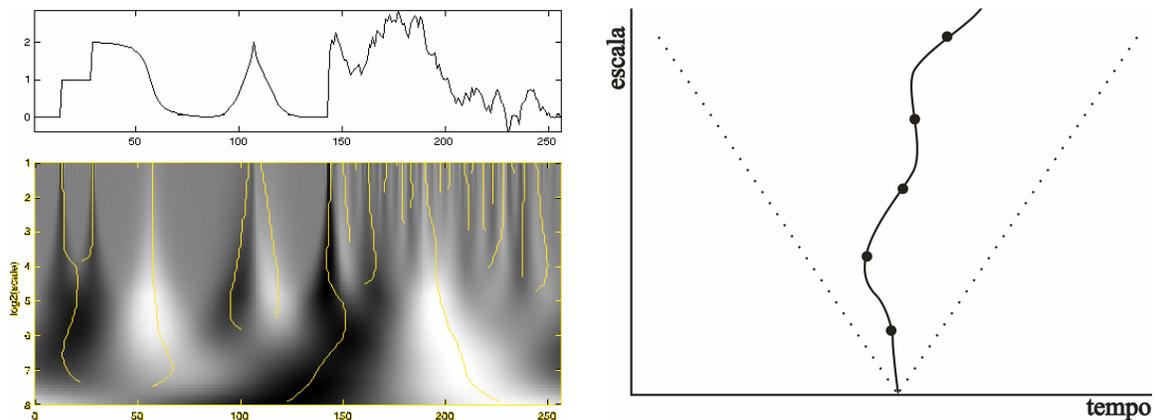


Figura 2.1 - Esquerda: exemplo de sinal e suas respectivas linhas de máximos. Direita: cone referente a uma linha de máximos

2.3.4 Estimação do expoente de Hölder pontual para singularidades oscilantes ou não-oscilantes

Como citado anteriormente, o método WTMM pode falhar na estimação do expoente de Hölder pontual para singularidades oscilantes. Isso se deve ao fato de não termos mais um cone limitando os coeficientes wavelets máximos em módulo, para um ponto x_0 com uma singularidade oscilante (α, β) . Nesta situação temos um cone “parabólico”, abrangendo um maior número de pontos do que o tipo de cone definido na seção anterior. O novo tipo de cone obedece a seguinte equação [ARN98]:

$$|2^j k - x_0|^{1+\beta} \leq K 2^j \quad (2.25)$$

Dentro deste cone, os coeficientes wavelet obedecem a seguinte lei de decaimento dos coeficientes wavelets:

$$|d_{j,k}| \leq A 2^{j\alpha(1+\beta)} \quad (2.26)$$

Podemos perceber que, para uma singularidade não-oscilante ($\beta = 0$), retornamos ao caso da seção anterior.

Diante destas observações, pode-se concluir que para a descrição eficiente do comportamento de uma função num determinado ponto, devemos levar em consideração não somente a amplitude dos coeficientes wavelet, mas também sua localização no tempo (representada por k). O seguinte teorema fundamental confirma nossa discussão [JAF89]:

Teorema 2.3.4.1 : Seja f uma função $R \rightarrow R$. Se $f \in C^\alpha(x_0)$, então

$$|d_{j,k}| \leq K 2^{j\alpha} (1 + |k - k_0|)^\alpha \quad (2.27)$$

Conforme pode-se constatar por meio da desigualdade (2.27), para uma singularidade em um ponto x_0 , o coeficiente wavelet máximo absoluto que a representa em uma escala j , não é necessariamente $d_{j,k}$, mas possivelmente $d_{j+m,k}$, com m real estritamente positivo [SEU00]. Tomando-se isto como base, Seuret et. al. propuseram um algoritmo de estimação dos expoentes de Hölder pontuais, considerando tanto as singularidades oscilantes quanto as singularidades não-oscilantes presentes em sinais [SEU00]. Este algoritmo, apresentado a seguir, permite o cálculo do valor da regularidade pontual nos instantes de tempo desejados.

Assim, seja um sinal amostrado, contendo 2^n amostras. Seja também $d_{j,k}$ a transformada wavelet não-dizimada (ou seja, redundante). A estimação do expoente de Hölder pontual para cada amostra k_0 é feita da seguinte maneira:

- Construa, em uma mesma figura, para cada $0 < j \leq n$, a seguinte curva paramétrica (com parâmetro $k \leq 2^n$):

$$x_j(k) = \log_2(2^j + |2^j k - x_0|) \quad (2.28)$$

$$y_j(k) = \log_2(|d_{j,k}|) \quad (2.29)$$

- Encontre todas as retas $D: y = \alpha x + C$ que satisfaçam as duas restrições a seguir:

1. D está acima de todos os pontos $(x_j(k), y_j(k))$, ou seja:

$$\forall j, \forall k, \quad y_j(k) \geq \alpha x_j(k) + C \quad (2.30)$$

2. D toca uma das curvas paramétricas; ou seja, existe uma seqüência de pares (j_m, k_m) tal que:

$$\lim_{m \rightarrow \infty} y_{j_m}(k_m) - (\alpha x_{j_m}(k_m) + C) = 0 \quad (2.31)$$

- Considere α_{max} o maior coeficiente angular encontrado entre todas as retas D que satisfaçam as restrições (1) e (2). O coeficiente α_{max} é o **expoente de Hölder pontual** do sinal para a amostra k_0 .

Este é o método de estimação do expoente de Hölder pontual utilizado neste trabalho daqui em diante.

2.4 Avaliação do estimador de expoente de Hölder pontual utilizado

Nesta seção, apresentamos uma avaliação do estimador utilizado para vários tipos de sinais de teste: sinal com singularidade não-oscilante, sinal com singularidade oscilante, sinal monofractal e sinal multifractal.

Devido ao fato de trabalharmos com sinais amostrados de comprimento finito, utilizamos uma função wavelet com suporte compacto. Conforme comentado anteriormente, a transformada wavelet deve ser do tipo não-dizimada. Este tipo de transformada produz redundância nas informações, necessária para uma estimativa mais precisa. Como na prática não se consegue um número infinito de escalas para tornar a transformada wavelet não-dizimada, a solução é o uso do maior número de escalas possível. Também devemos considerar a escala máxima utilizada, devido ao fato de estarmos lidando com um número muito grande de amostras. Nesta situação, a possível ocorrência de um fenômeno de periodização poderia afetar os resultados em escalas maiores.

Outra consideração feita é a quantidade de pontos $(x_j(k), y_j(k))$ utilizados para a estimação de cada expoente de Hölder pontual do sinal. Em relação a isto, usamos a restrição $3 \leq x(j, k) \leq \log_2(2^n) - 2$, sendo 2^n , o número de amostras utilizadas. Esta restrição, obtida de maneira empírica, nos permite considerar os coeficientes wavelet máximos absolutos em uma extensão suficiente de expoentes de oscilação β e gerar uma quantidade de pontos $(x_j(k), y_j(k))$ suficiente para construção correta da reta que toca o topo desta “nuvem de pontos”.

Desta forma, utilizamos 2^{14} amostras para cada sinal de teste. A quantidade de amostras utilizada é escolhida de forma a tender uma boa precisão dos resultados e uma alta velocidade de simulação. Utilizamos a *Morlet* como wavelet-mãe, cujo suporte efetivo é limitado pelo intervalo $[-4, 4]$. A escala máxima utilizada foi $j=12$.

2.4.1 Singularidade não-oscilante

Seja a função $f(x)=|x|^{0,7}$. Esta função é infinitamente diferenciável em \mathbb{R}^* e possui uma singularidade não-oscilante isolada em 0, onde o expoente de Hölder pontual possui valor 0,7. A Figura 2.2 nos mostra a função f , a nuvem de pontos $(x_j(k), y_j(k))$ e a respectiva reta que toca o seu topo, relativas ao local desta singularidade.

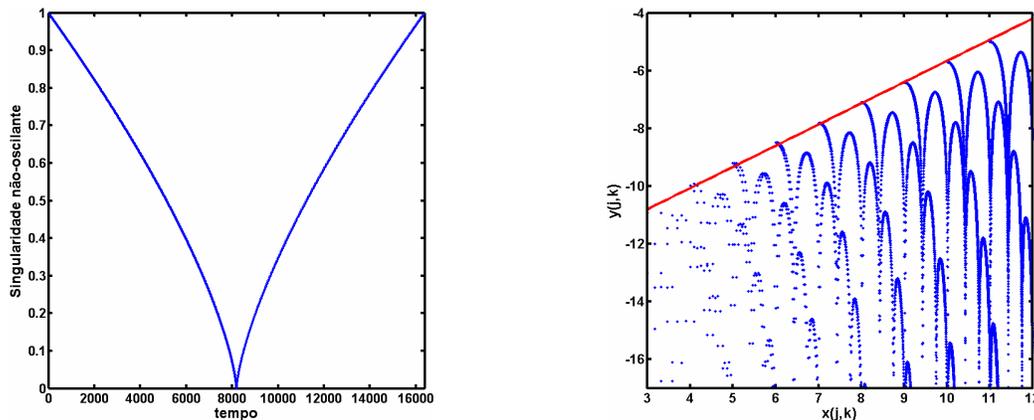


Figura 2.2 - Esquerda: função $|x|^{0,7}$. Direita: “nuvem de pontos” e reta associados ao ponto com singularidade não-oscilante da função.

2.4.2 Singularidade oscilante

Seja a função $g(x) = x^{0,7} \sin \frac{1}{x^{1,2}}$. Esta função é infinitamente diferenciável em todo ponto, exceto em zero, onde há uma singularidade oscilante com expoente de Hölder pontual cujo valor é 0,7. A Figura 2.3 nos mostra a função g , a nuvem de pontos e a respectiva reta que toca o seu topo, relativas ao local da singularidade.

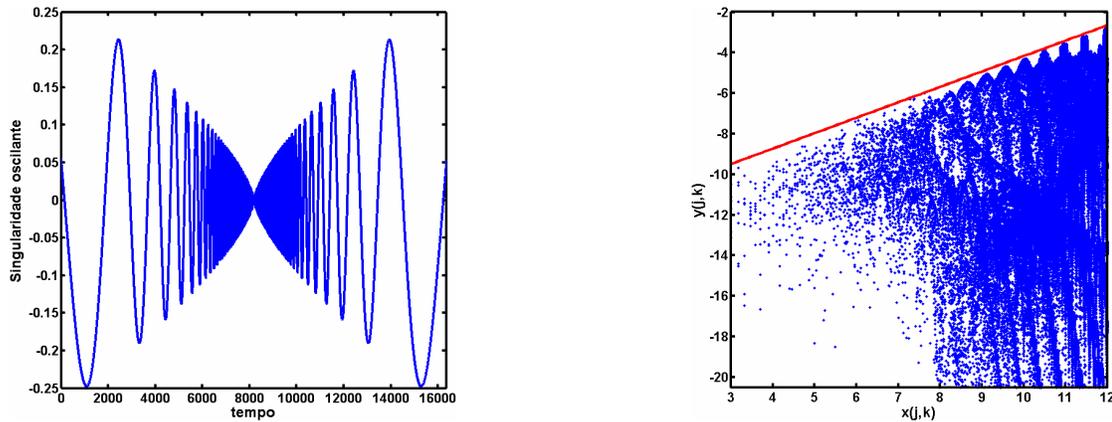


Figura 2.3 - Esquerda: função $|x|^{0.7}\text{sen}(1/|x|^{1.2})$. Direita: “nuvem de pontos” e reta associada ao ponto com singularidade oscilante da função.

Agora, a diferença entre uma singularidade oscilante e uma não-oscilante pode tornar-se mais clara. Embora as duas funções possuam expoente de Hölder com o mesmo valor 0,7 no ponto zero, suas nuvens de pontos possuem características diferentes. No caso de uma singularidade não-oscilante, pode-se observar que o topo da sua respectiva nuvem é formado pelos coeficientes wavelet máximos absolutos em cada escala. Estes coeficientes são exatamente aqueles que formam a linha de máximos (seção 2.3.3) que tendem à singularidade. Não há necessidade de interação entre coeficientes de escalas diferentes. No caso de uma singularidade oscilante, podemos notar que devemos considerar uma quantidade maior de pontos $(x_j(k), y_j(k))$ para a determinação correta da reta que se encaixa no topo da nuvem. Desta forma, um coeficiente wavelet oriundo uma escala $j+m$ pode caracterizar uma escala j no topo da nuvem de pontos, conforme discutido no Teorema 2.3.4.1.

2.4.3 Sinal monofractal

Como exemplo de um sinal monofractal, consideramos o *Movimento Browniano Fracionário* (fBM). O fBM é uma função aleatória, não-estacionária, gaussiana, com média zero e função de autocorrelação definida como

$$R_{B_H}(s, t) = C_H \left(|s|^{2H} + |t|^{2H} - |t-s|^{2H} \right) \quad (2.32)$$

em que $0 < H < 1$ é o parâmetro de Hurst e C_H é uma constante que depende somente de H . Pode-se verificar que tal processo é estatisticamente auto-similar com parâmetro H [SAM94]. Assim, em cada ponto desta função, temos uma singularidade não-oscilante com expoente de Hölder com valor H . Isto é muito útil para testarmos nosso estimador.

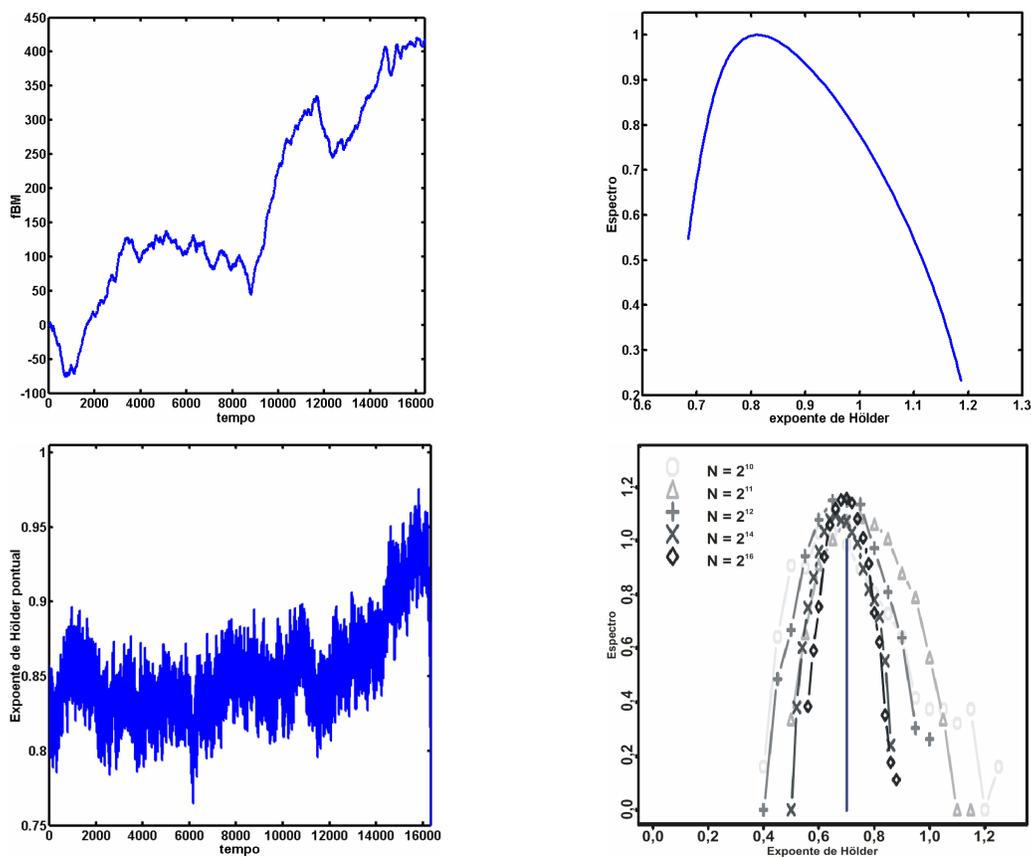


Figura 2.4 - Acima à esquerda: fBM com $H=0,85$. Abaixo à esquerda: expoentes de Hölder pontuais estimados para o mesmo fBM. Acima à direita: espectro multifractal do mesmo fBM. Abaixo à direita: diferentes espectros multifractais para um fBM com $H=0,7$, computados com 2^{10} até 2^{16} amostras.

A Figura 2.4 (acima à esquerda) apresenta um sinal fBM com $H = 0,85$, gerado por meio do *FracLab*. Na mesma figura, encontram-se o resultado da estimação dos expoentes de Hölder pontuais (abaixo à esquerda da figura) e o espectro multifractal do sinal (acima à direita), estimado por meio do *FracLab*. Podemos observar que os expoentes estimados não possuem valor único constante igual a 0,85, embora demonstrem valores muito próximos àqueles encontrados sob a curva do espectro estimado. A ocorrência deste fenômeno deve-se à utilização de uma quantidade ainda insuficiente de amostras para a estimação dos expoentes do espectro [SEU00]. Segundo demonstrado por Seuret et. al., à medida que utilizamos uma quantidade maior de amostras, naturalmente consideramos uma maior quantidade de escalas. Desta forma, as estimativas tornam-se mais precisas e convergem ao valor esperado. Abaixo à direita da Figura 2.4, podemos observar esta tendência. Conforme aumenta-se o número de amostras de $N = 2^{10}$ até $N = 2^{16}$, para um fBM gerado com $H = 0,7$, o espectro multifractal estimado converge ao espectro teórico, ou seja, um impulso localizado em $H = 0,7$.

2.4.4 Sinal multifractal

Cascatas são processos obtidos de maneira recursiva e servem como um interessante teste ao nosso estimador. Isto se deve ao fato de as mesmas serem multifractais e permitirem a obtenção do seu espectro multifractal teórico.

Seja o número real $p \in (0,1)$. A partir do intervalo inicial $I_0=[0,1]$ com massa $\mu(I_0) = 1$, dividimos com comprimento igual cada intervalo restante. O subintervalo da esquerda possui peso igual a p e o subintervalo da direita possui peso igual a $1-p$. Este processo se repete em cada subintervalo. No estágio n , a massa inicial de valor unitário é distribuída entre os 2^n intervalos de acordo com a seguinte regra:

$$\mu(I_k) = p^{n_0} (1-p)^{n_1} \quad (2.33)$$

em que n_0 e n_1 é o número de vezes que p e $1-p$ são multiplicados, respectivamente. Note que $n_0+n_1=n$.

Esta é a **cascata binomial determinística** utilizada como teste para o nosso estimador. Em cada ponto, a mesma possui uma singularidade não-oscilante. Neste caso, diferente do fBM, cada ponto possui expoente de Hölder pontual cujo valor não necessariamente é igual ao valor do mesmo expoente em outro ponto. Este tipo de cascata multifractal permite a determinação do seu espectro multifractal teórico. Os limites deste espectro são dados por

$$\alpha_1 = \frac{\log_2[p^n]}{-n} \quad (2.34)$$

$$\alpha_2 = \frac{\log_2[(1-p)^n]}{-n} \quad (2.35)$$

Devido ao fato deste espectro apresentar simetria em torno do eixo vertical, seu cume (onde $D(\alpha) = 1$) localiza-se em α_{mean} , dado por:

$$\alpha_{mean} = \frac{\alpha_1 + \alpha_2}{2} \quad (2.36)$$

A Figura 2.5 exhibe uma cascata binomial determinística com parâmetros $p=0,3$ e $n=14$. Na mesma figura, são mostrados os expoentes de Hölder pontuais para este processo e seu respectivo espectro multifractal teórico estimado por meio do *FracLab*.

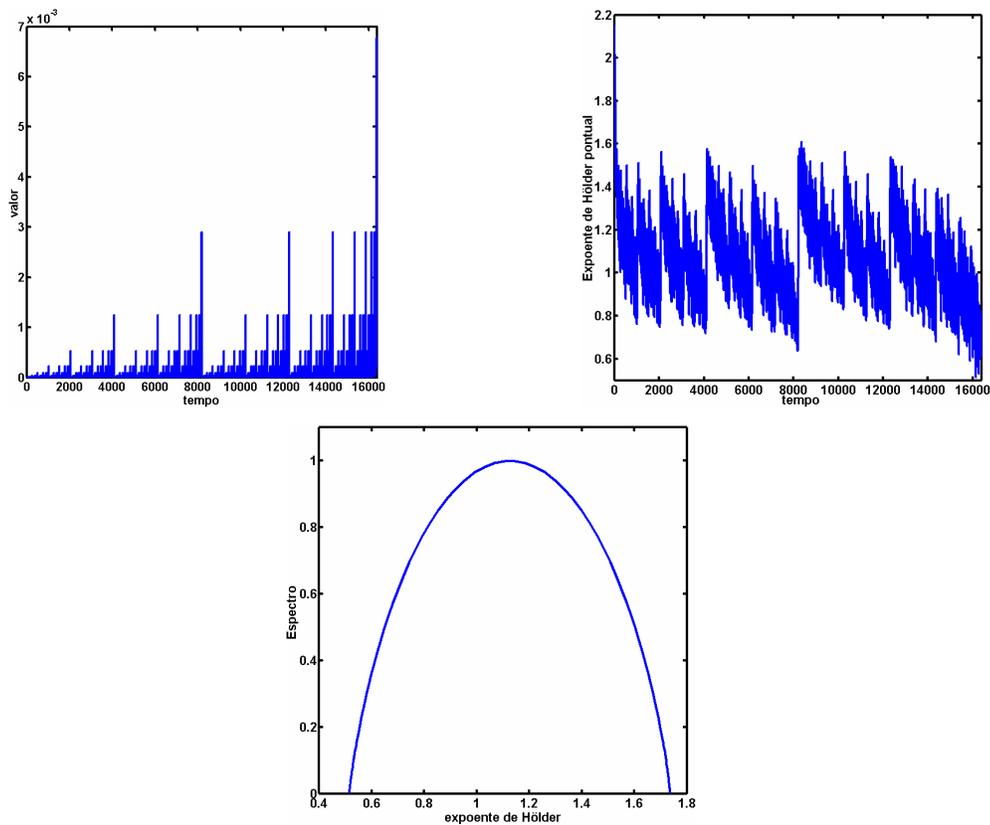


Figura 2.5 - Acima à esquerda: cascata binomial com $p=0,3$ e $n=14$. Acima à direita: expoentes de Hölder pontuais estimados para a cascata. Abaixo: espectro multifractal teórico da cascata.

2.5 Estimação da regularidade local do tráfego Internet

Nesta subseção, prosseguimos a avaliação do estimador de expoente de Hölder pontual com sinais de tráfego Internet. Para isto, utilizamos os traços de tráfego WAN TCP/IP *dec-pkt-1*, *dec-pkt-2* e *lbl-pkt-5*². Estes traços correspondem ao registro do tráfego de pacotes IP transmitidos em períodos de uma hora e foram coletados pela *Digital Equipment Corporation* (DEC) e pelo *Lawrence Berkeley Laboratory* (LBL) em seus respectivos pontos de acesso da Internet. Cada sinal de tráfego utilizado neste trabalho é constituído por 2^{14} amostras de cada um dos traços citados, agregados em intervalos (escalas) de tempo diferentes (100 e 200 ms). As amostras dos sinais resultantes contêm a quantidade de bytes transmitida em cada um dos intervalos de tempo citados.

² <http://ita.ee.lbl.gov/html/traces.html>

Inicialmente definiremos os expoentes de Hölder pontuais de referência para cada série de tráfego. Esta definição é importante para tornar-se possível a avaliação de uma nova proposta feita por nós. Esta proposta tem como objetivo a estimação dinâmica dos expoentes de Hölder pontuais de um sinal de tráfego por meio de janelas de tempo.

2.5.1 Estimação do expoente de Hölder pontual de referência

Neste trabalho, definimos como expoente de Hölder pontual de referência cada expoente de Hölder pontual estimado com uso de todas as 2^{14} amostras disponíveis de um sinal de tráfego.

As Figuras 2.6, 2.7 e 2.8 referem-se respectivamente às amostras dos sinais de tráfego *dec-pkt-1* (na escala de tempo de 100 ms), *lbl-pkt-5* (na escala de tempo de 100 ms) e *lbl-pkt-5* (na escala de tempo de 200 ms). Cada figura apresenta um destes sinais de tráfego, seus respectivos expoentes de Hölder pontuais, assim como seu espectro multifractal, estimado por meio do *FracLab*.

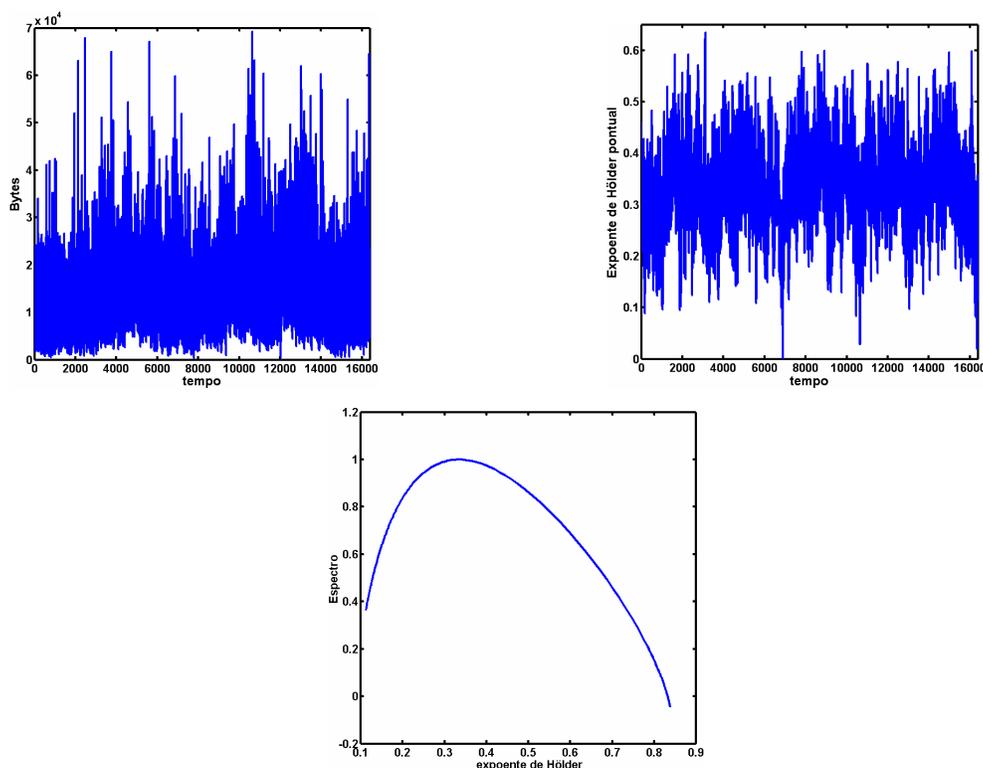


Figura 2.6 - Acima à esquerda: amostras de tráfego da série *dec-pkt-1* na escala de tempo de 100 ms. Acima à direita: expoentes de Hölder pontuais referentes às amostras citadas. Abaixo: espectro multifractal.

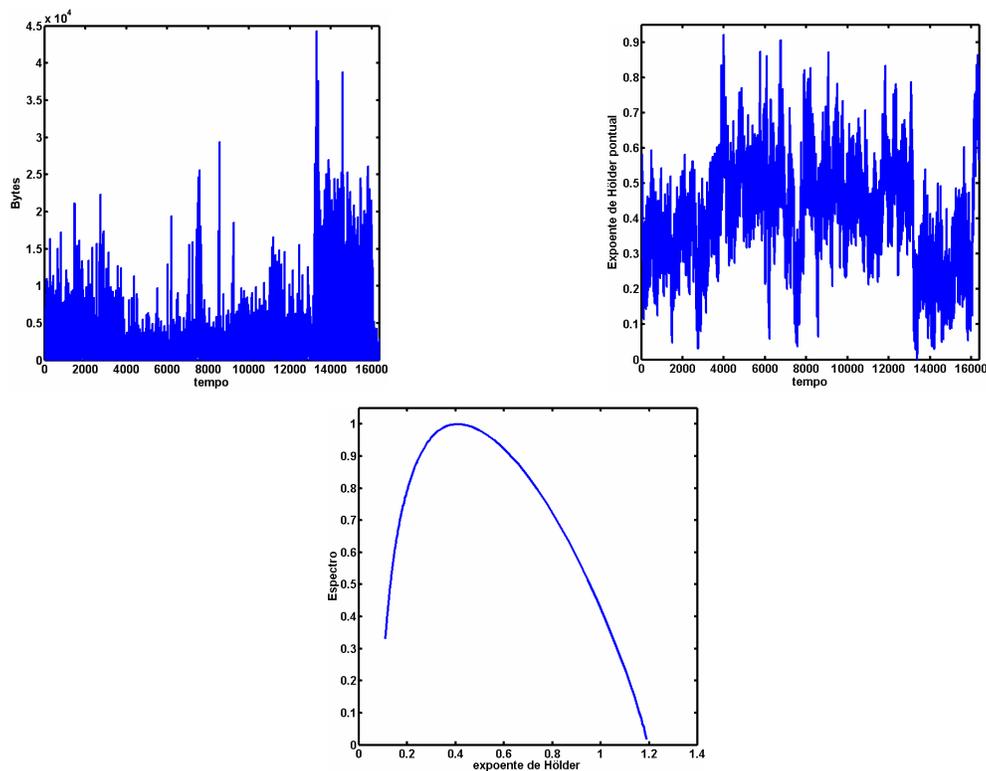


Figura 2.7 - Acima à esquerda: amostras de tráfego da série *lbl-pkt-5* na escala de tempo de 100 ms. Acima à direita: expoentes de Hölder pontuais referentes às amostras citadas. Abaixo: espectro multifractal.

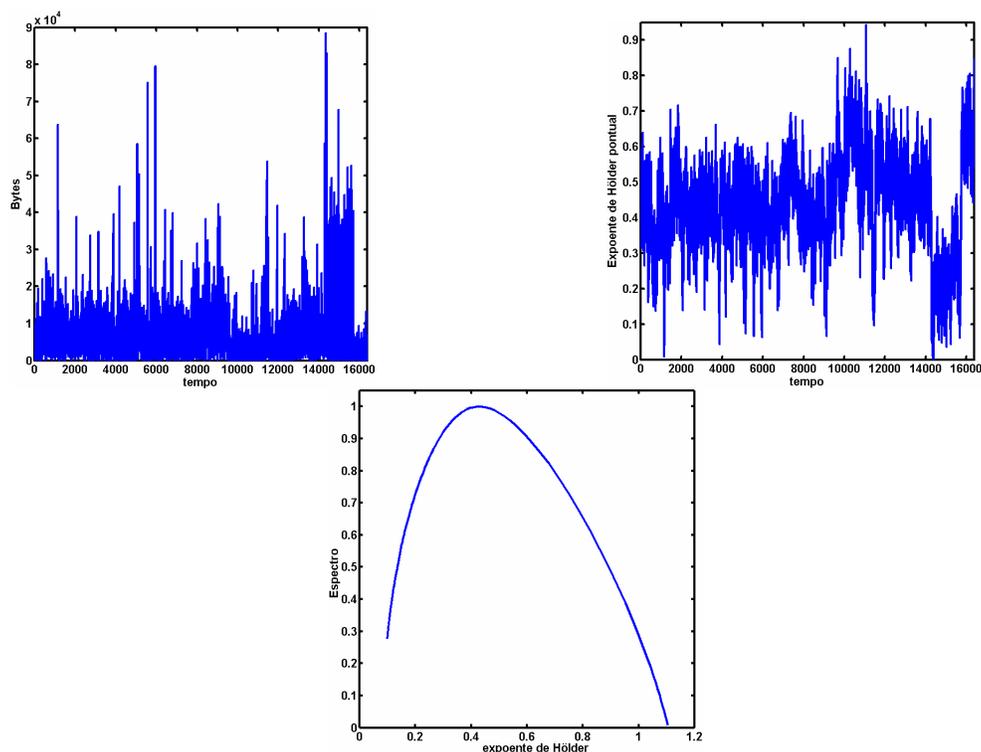


Figura 2.8 - Acima à esquerda: amostras de tráfego da série *lbl-pkt-5* na escala de tempo de 200 ms. Acima à direita: expoentes de Hölder pontuais referentes às amostras citadas. Abaixo: espectro multifractal.

À primeira vista, podemos observar a proximidade dos valores dos expoentes estimados com os valores dos expoentes presentes no espectro. Pode-se notar, inclusive, a queda dos valores destes expoentes em instantes de tempo em que há um aumento acentuado da amplitude do sinal. Por meio destas duas observações, podemos deduzir que os valores dos expoentes estimados são próximos dos valores reais dos expoentes. Estes mesmos valores, em sua grande maioria menores do que 0,5, demonstram a intensa irregularidade e o caráter de rajadas (*burstiness*) do tráfego Internet, em pequenas escalas de tempo.

2.5.2 Proposta: estimação do expoente de Hölder pontual por meio de janelas de tempo

Nesta subseção, propomos uma estratégia dinâmica para a estimação do expoente de Hölder pontual para um dado sinal. Esta estratégia se baseia na utilização de uma quantidade fixa de amostras consecutivas (enquadrada em uma janela de tempo) para a estimação do expoente de Hölder pontual de cada amostra do sinal. Como análise, nesta subseção consideramos as janelas de tempo seqüenciais e sem amostras em comum, conforme a Figura 2.9.

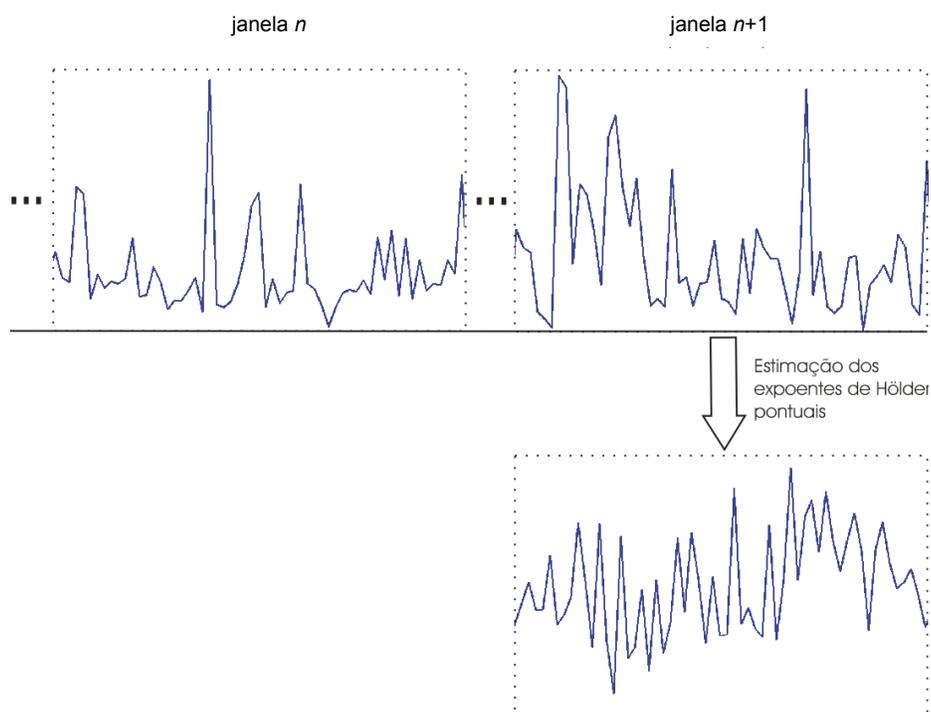


Figura 2.9 - Utilização de janelas de tempo na estimação dos expoentes de Hölder pontuais de amostras de um sinal de tráfego

Cada amostra tem seu respectivo expoente estimado utilizando-se somente as amostras da janela à qual pertence. Desta forma, conseguimos utilizar uma menor quantidade de amostras (assim como uma menor quantidade de coeficientes wavelet) para a estimação do expoente em cada instante de tempo. A diminuição da quantidade de coeficientes wavelet utilizada em cada estimação nos possibilita a vantagem de um processamento mais rápido das informações. Como consequência negativa, ao restringirmos a quantidade de amostras, diminuimos a precisão da estimativa. São exatamente estas imprecisões que analisamos a seguir, comparando os expoentes de Hölder pontuais estimados por meio de janelas de tempo com os expoentes de Hölder pontuais de referência.

Como medida de desempenho da estimação, utilizamos o erro quadrático médio normalizado (EQMN), definido a seguir:

$$EQMN = \frac{E[(h_{ref} - h_{jan})^2]}{\sigma^2_{h_{ref}}} \quad (2.37)$$

em que h_{jan} é o expoente de Hölder pontual com uso de janelas de tempo, h_{ref} é o expoente de Hölder pontual usado como referência (sem uso de janelas de tempo), $\sigma^2_{h_{ref}}$ é a variância dos expoentes de Hölder pontuais de referência e $E [.]$ denota o operador esperança matemática. O EQMN compara a média do erro quadrático (EQM), resultante do uso das janelas no processo de estimação, com a variância dos expoentes de Hölder pontuais de referência. Idealmente, uma boa estimação resulta em valores de EQMNs pequenos; EQMNs iguais ou superiores à 1, resultam em estimativas ruins, pois neste caso o EQM possui valor igual ou superior à variância dos expoentes de referência.

Para as simulações, adotamos três tamanhos de janelas: 2^{11} amostras, 2^{12} amostras e 2^{13} amostras. Como conveniência, passaremos a referenciá-las respectivamente por janela 11, janela 12 e janela 13. Conforme explicado no início da seção 2.4, condicionamos a escala máxima de cada transformada wavelet ao valor 2^{j-2} , sendo 2^j o tamanho da janela utilizada.

As tabelas 2.1 e 2.2 nos apresentam os EQMNs obtidos na estimação dos expoentes de Hölder pontuais para os sinais de tráfego considerados, por meio do uso de três tamanhos distintos de janelas de tempo. Conforme esperava-se, os valores dos EQMNs aumentam conforme diminui-se o tamanho da janela de tempo usada na estimação. Isto confirma o aumento da imprecisão da estimativa com a diminuição do tamanho da janela. A Figura 2.10

nos dá uma noção visual das imprecisões causadas pelo processo de estimação com o uso de dois tamanhos de janelas de tempo. Estas imprecisões podem ser notadas principalmente em pontos localizados nos picos presentes nos gráficos.

Tabela 2.1 - EQMN dos expoentes de Hölder pontuais das séries de tráfego na escala de tempo de 100 ms.

	Janela 11	Janela 12	Janela 13
série <i>dec1</i>	0,4899	0,2483	0,0426
série <i>dec2</i>	0,7099	0,4879	0,2674
série <i>lbl5</i>	0,5423	0,4332	0,2116

Tabela 2.2 - EQMN dos expoentes de Hölder pontuais das séries de tráfego na escala de tempo de 200 ms.

	Janela 11	Janela 12	Janela 13
série <i>dec1</i>	0,5471	0,3536	0,2147
série <i>dec2</i>	0,4594	0,31	0,2030
série <i>lbl5</i>	0,5743	0,3626	0,1420

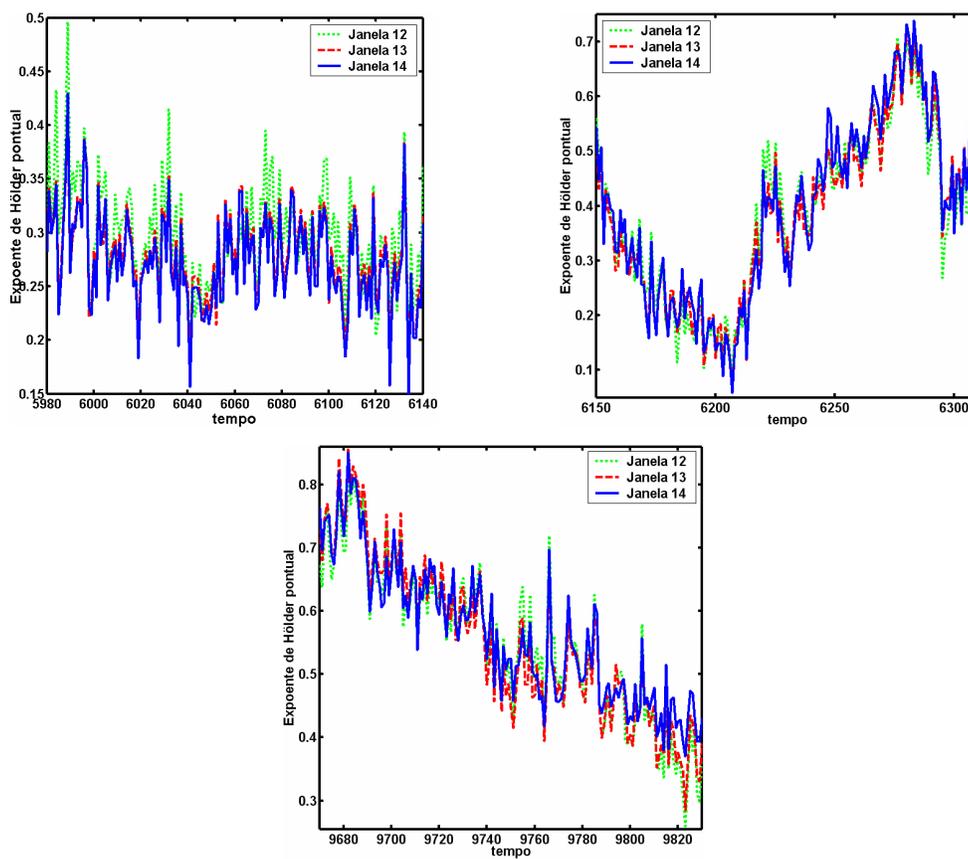


Figura 2.10 - Expoentes de Hölder pontuais de amostras de três séries de tráfego Internet: *dec-pkt-1* na escala de tempo de 100 ms (acima à esquerda), *lbl-pkt-5* na escala de tempo de 100 ms (acima à direita) e *lbl-pkt-5* na escala de tempo de 200 ms (abaixo).

2.6 Conclusões do capítulo

Neste capítulo, apresentamos o expoente de Hölder pontual como indicador da regularidade local de um sinal. Em especial para sinais de tráfego de redes, este expoente pode caracterizar o grau das rajadas presentes no tráfego.

A regularidade local de um sinal de tráfego pode ser avaliada, de modo global, por meio do espectro multifractal ou, de modo local, pela estimativa do expoente de Hölder pontual em instantes de tempo desejados. Para uma estimativa correta destes expoentes, devemos considerar o caráter oscilatório do sinal em torno de cada instante de tempo analisado. Para isto, devemos considerar não somente a amplitude dos coeficientes wavelet, mas também suas localizações no tempo.

Mostramos que as estimativas dos expoentes de Hölder pontuais de um sinal podem ser obtidas de maneira dinâmica, por meio do uso de janelas de tempo durante o processo de estimação. Esta técnica, embora simples, gera estimativas cujos tempos de obtenção e precisão são dependentes da quantidade de amostras contidas na janela utilizada. Quanto menor o tamanho da janela, mais rápido é o processo de estimação dos expoentes e mais imprecisas serão as estimativas. Quanto a isto, mostramos que mesmo para janelas contendo 2^{11} amostras, os resultados obtidos possuíam imprecisões aceitáveis.

Pretendemos, a seguir, utilizar a série de expoentes de Hölder pontuais estimados a partir de cada sinal de tráfego e aplicá-la em um esquema de previsão. Nossa meta é verificar a previsibilidade do tráfego de rede por meio do expoente de Hölder pontual e adiantar o comportamento do tráfego a um passo na escala de tempo.

3 PREDIÇÃO DO EXPOENTE DE HÖLDER PONTUAL

3.1 Introdução do capítulo

No capítulo anterior pudemos apresentar o expoente de Hölder pontual como um quantificador do grau das rajadas de um sinal de tráfego de redes. No mesmo capítulo propusemos um método dinâmico de estimação deste expoente para um determinado sinal. O objetivo do atual capítulo é a predição a um passo do expoente de Hölder pontual de amostras do sinal de tráfego analisado. Isto nos serve como um meio de antecipação das condições do próprio tráfego. Para isto, utilizamos estimativas dos expoentes deste sinal e um preditor proposto neste mesmo capítulo. A predição destes expoentes é importante para a realização do escalonamento de fluxos de dados no capítulo a seguir.

Para a caracterização do tráfego em um instante de tempo, podemos considerar o expoente de Hölder pontual ou um quantificador mais direto do tráfego, tal como a quantidade de bytes recebida naquele instante. Entre estes dois tipos de caracterizadores locais do tráfego, a escolha do expoente de Hölder pontual tem sua justificativa. Esta justificativa se baseia na comparação do decaimento da função de autocorrelação das séries de amostras consideradas. A Figura 3.1 nos apresenta a função de autocorrelação das amostras de três sinais de tráfego e de suas respectivas séries de expoentes de Hölder pontuais estimados. Nesta figura, podemos notar um decaimento assintótico mais lento da função de autocorrelação do sinal de tráfego, confirmando sua propriedade de dependência de longo prazo [QIA04]. Isto pode justificar o uso de técnicas mais complexas (e computacionalmente custosas) de predição do tráfego, tais como Redes Neurais Artificiais e/ou Lógica Nebulosa [ALA05][BOR00]. Na mesma figura, o decaimento assintótico mais rápido das funções de autocorrelação das séries de expoentes de Hölder pontuais nos indica que tais séries não apresentam dependência de longo prazo. Isto nos sugere a possibilidade do uso de técnicas clássicas e mais simples para predição de séries temporais, tais como o filtro de Mínimos Médios Quadrados (*Least-Mean Squares* - LMS) e o filtro de Kalman [HAY91].

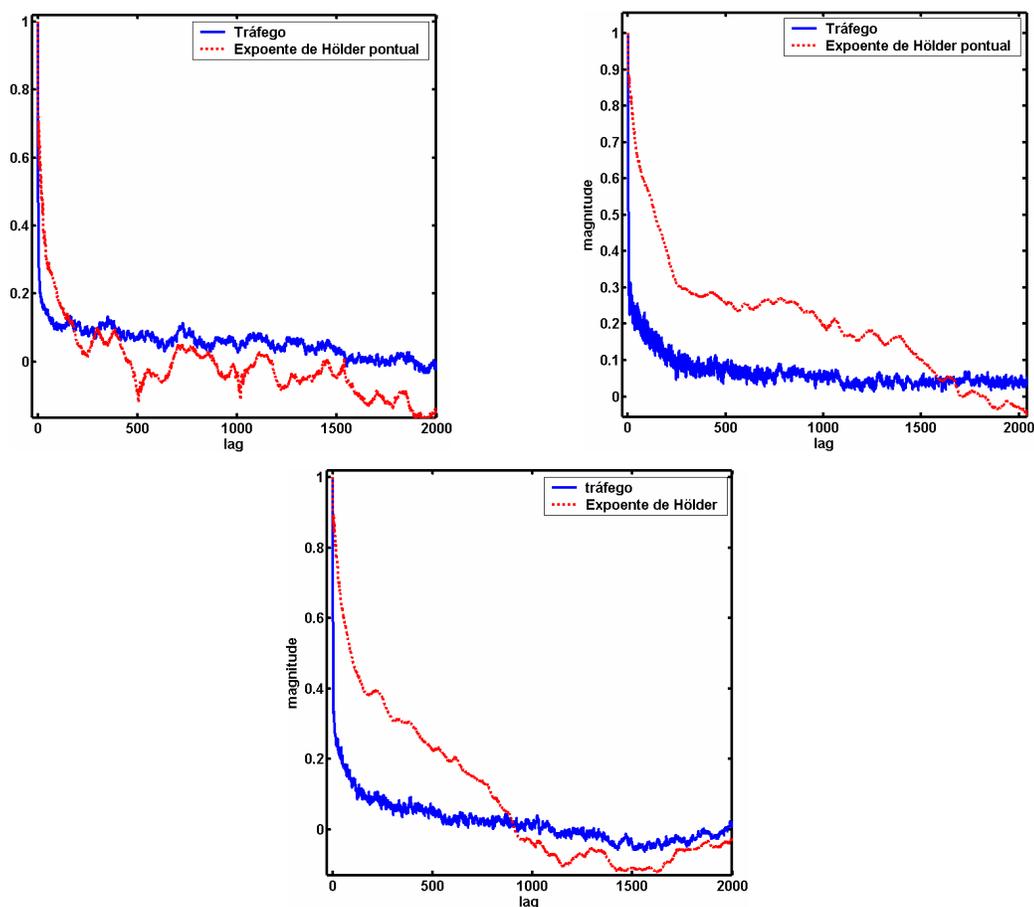


Figura 3.1 - Função de autocorrelação das amostras de uma série de tráfego e de seus respectivos expoentes de Hölder pontuais. Acima à esquerda: série *dec-pkt-1* na escala de tempo de 100 ms. Acima à direita: série *lbl-pkt-5* na escala de tempo de 100 ms. Abaixo: série *lbl-pkt-5* na escala de tempo de 200 ms.

Para atingirmos nosso objetivo, inicialmente, na Seção 3.2, apresentamos dois preditores clássicos e tomando-os como base, propomos um novo esquema de predição. A seguir, na Seção 3.3, avaliamos o desempenho do preditor proposto. Esta avaliação é feita utilizando-se os expoentes de Hölder pontuais estimados a partir dos sinais analisados. O desempenho do preditor proposto também é comparado com o desempenho de preditores considerados clássicos. Na seqüência, a Seção 3.4 avalia conjuntamente os erros de estimação e de predição destes expoentes para os sinais de tráfego considerados. Por último, na Seção 3.5, concluímos este capítulo.

3.2 Filtros adaptativos preditivos

Filtro é um dispositivo utilizado para a extração de uma informação de interesse, a partir de um conjunto de informações ruidosas. Este dispositivo normalmente é útil para três tipos de operações básicas em processamento de sinais: filtragem, suavização e predição. Neste trabalho, nossa atenção está especialmente voltada para este último tipo de operação.

Em processamento de sinais, predição refere-se à antecipação do valor de um determinado sinal, relativo ao instante de tempo futuro $t + \tau$, $\tau > 0$, usando o conjunto de valores medidos até (e inclusive) o instante t . Em situações nas quais o sinal analisado possui características estatísticas completamente conhecidas à priori, o processo de predição pode ser feito por meio de técnicas mais simples, como o filtro de Wiener [HAY91].

Nas situações em que não há conhecimento completo acerca das características estatísticas do sinal analisado, ou estas variam com o tempo, um filtro adaptativo pode ser usado como estratégia para a predição do sinal. Este tipo de filtro normalmente utiliza um algoritmo recursivo e sua operação envolve dois processos básicos: predição e adaptação. O processo de predição produz o valor predito do sinal analisado, baseado em valores já disponíveis e nos parâmetros atuais do filtro. Já o processo de adaptação nos fornece um mecanismo de atualização de um conjunto de parâmetros utilizados no processo de predição. Em geral, o processo de adaptação se baseia no erro de predição.

Nesta seção, apresentamos dois filtros preditores clássicos: o filtro LMS Normalizado (NLMS) e o filtro de Kalman. Tomando-se ambos como base, apresentaremos uma nova proposta para predição adaptativa. Sua inovação situa-se justamente na estimação adaptativa dos ruídos do sistema que descreve o sinal analisado. Esta proposta demonstra sua maior utilidade nas situações em que não há possibilidade de conhecimento completo das características dos ruídos citados.

3.2.1 Filtro NLMS

O algoritmo dos Mínimos Médios Quadrados Normalizado (Normalized Least-Mean Squares - NLMS), utilizado pelo filtro NLMS para o processo de adaptação, como o próprio nome nos informa, é uma extensão do algoritmo dos Mínimos Médios Quadrados (Least-Mean Squares – LMS), característico do filtro LMS. Desta forma, é interessante inicialmente

apresentarmos o algoritmo LMS para a formulação do filtro NLMS. O filtro LMS é popularmente utilizado para predição adaptativa, devido à sua simplicidade, robustez e baixo custo computacional [HAY91]. Por ser um filtro preditor adaptativo, sua operação envolve um processo de adaptação e um processo de predição. O processo de adaptação consiste no ajuste automático de um conjunto de pesos, conforme veremos a seguir. Já o processo de predição consiste no uso de um filtro transversal para a obtenção do valor a ser predito e para a obtenção do erro de predição.

Desta forma, para o processo de predição citado, seja o filtro transversal mostrado na Figura 3.2. O filtro transversal é um tipo de estrutura de filtro, composta por elementos de atraso, multiplicadores e somadores. Os elementos de atraso são identificados por meio do operador atraso unitário z^{-1} . O número de elementos de atraso, mostrado como N na figura citada, refere-se à **ordem do filtro**. Considere, para o mesmo filtro, os valores de entrada representados pela série temporal $u(0), u(1), u(2), \dots$, o vetor de entradas $\mathbf{u}(k) = [u(k), u(k-1), \dots, u(k-N)]^T$ e a resposta desejada $d(k)$, em algum instante de tempo discreto k , com $k \in \mathbb{Z}$.

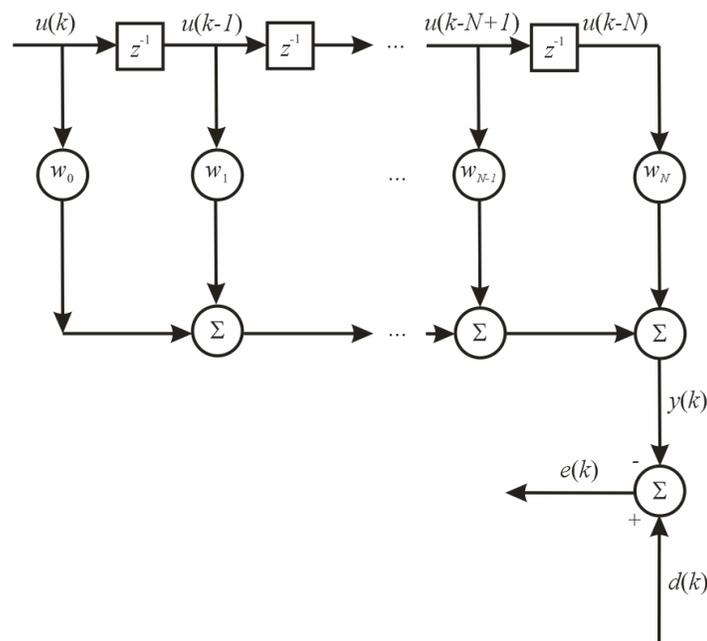


Figura 3.2 - Filtro transversal.

Desejamos encontrar o vetor $\mathbf{w} = [w_0, w_1, \dots, w_N]^T$, composto pelos coeficientes do filtro, tal que o valor predito $y(k)$, dado por

$$y(k) = \mathbf{w}^T \mathbf{u}(k) \quad (3.1)$$

seja tão próximo quanto possível de $d(k)$, no sentido de menor erro quadrático médio (EQM).

Assim, se $\mathbf{u}(k)$ e $d(k)$ forem conjuntamente estacionários e o erro de predição for definido como

$$e(k) = y(k) - d(k) \quad (3.2)$$

o vetor \mathbf{w} que minimiza a seguinte função de custo [HAY91]:

$$J(\mathbf{w}) = E[|e|^2] \quad (3.3)$$

é conhecido como solução de Wiener [HAY91]

$$\mathbf{w} = \mathbf{R}^{-1} \mathbf{p} \quad (3.4)$$

em que

$$\mathbf{R} = E[\mathbf{u}(k)\mathbf{u}^T(k)] \quad (3.5)$$

$$\mathbf{p} = E[d(k)\mathbf{u}(k)] \quad (3.6)$$

Nas equações acima, o operador $E[\cdot]$ denota a esperança matemática.

O vetor \mathbf{w} pode ser calculado recursivamente por meio do algoritmo de gradiente descendente (*steepest-descent*) [HAY91]

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu[\mathbf{p} - \mathbf{R}\mathbf{w}(k)] \quad (3.7)$$

Caso \mathbf{R} e \mathbf{p} sejam desconhecidos, podemos substituí-los por estimativas. Usando $\mathbf{u}(k)\mathbf{u}^T(k)$ e $d(k)\mathbf{u}(k)$ como estimativas instantâneas respectivamente de \mathbf{R} e \mathbf{p} , $\mathbf{w}(k+1)$ pode ser reescrito como

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu[d(k)\mathbf{u}(k) - \mathbf{u}(k)\mathbf{u}^T(k)\mathbf{w}(k)] \quad (3.8)$$

ou, rearranjando-se, obtemos o algoritmo LMS dado por

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu[\mathbf{u}(k)e(k)] \quad (3.9)$$

O parâmetro μ é conhecido como passo de adaptação (*step-size*) e seu valor afeta a estabilidade e a velocidade de convergência média quadrática do algoritmo. A região de estabilidade para esta convergência é dada por $0 < \mu < \frac{2}{\text{Tr}[\mathbf{R}]}$ [SOL95][HAY91]. Nesta desigualdade, $\text{Tr}[\cdot]$ indica o traço de \mathbf{R} . No entanto, nem sempre podemos ter conhecimento de \mathbf{R} para a definição correta de μ . Isto ocorre normalmente com sinais reais, os quais muitas vezes apresentam estatísticas variantes no tempo.

Para sinais reais, o algoritmo LMS Normalizado (NLMS) pode resolver o problema de indeterminação do passo de adaptação μ [HAY91][NAG67]. Este algoritmo determina o passo de adaptação $\mu(k)$, variante no tempo, que minimiza o erro quadrático médio (EQM) do próximo instante.

$$J_{k+1}(\mathbf{w}) = E[e^2(k+1)] \quad (3.10)$$

Depois de incorporarmos um passo de adaptação fixo auxiliar $\tilde{\mu}$, o algoritmo NLMS é escrito como

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\tilde{\mu}}{\|\mathbf{u}(k)\|^2} [\mathbf{u}(k)e(k)] \quad (3.11)$$

na qual $\|\cdot\|$ indica a norma euclidiana.

Os limites teóricos para a estabilidade deste algoritmo são dados por $0 < \tilde{\mu} < 2$ [HAY91]. Diferente do passo de adaptação μ do algoritmo LMS, o passo de adaptação $\tilde{\mu}$ é **adimensional**. Desta forma a região de estabilidade do algoritmo NLMS é **independente** das estatísticas do sinal de entrada. Isto nos permite determinar facilmente o passo de adaptação do algoritmo, além de garantir sua estabilidade. Além disso, o algoritmo NLMS é conhecido por convergir mais rápido que o algoritmo LMS [SLO93] [RUP93].

3.2.2 Filtro de Kalman

O filtro de Kalman é um conjunto de equações matemáticas eficiente para a estimação dos estados de um processo. Este filtro opera de modo a minimizar a média do erro quadrático dessa estimação e é muito poderoso em vários aspectos: suporta estimação de estados passados, presentes e futuros e inclusive o faz mesmo quando o sistema modelado não é totalmente conhecido. O filtro de Kalman é extremamente útil em várias aplicações do mundo real, tais como: controle de processos químicos [DIM91], controle da posição ou da velocidade de sistemas inerciais [GEO04], processamento de sinais [KOM02] e previsão. No campo da previsão constam desde previsões meteorológicas [HOU05], como financeiras [GRA05] e de características do tráfego de redes de computadores [KRI02][LIM00]. Este filtro possui modelagem relativamente simples e sua solução é computada recursivamente. Isto é interessante computacionalmente, pois não há a necessidade de armazenamento do conjunto completo de informações passadas; basta apenas a última estimação disponível e um novo dado de entrada.

Para o funcionamento apropriado do filtro, o processo a ser mensurado deve ser capaz de ser descrito como um sistema linear. Um sistema linear é um processo que pode ser descrito segundo as seguintes equações no espaço de estado [HAY91]:

$$\mathbf{w}(k+1) = \mathbf{A}\mathbf{w}(k) + \boldsymbol{\eta}_1(k) \quad (3.12)$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{w}(k) + \boldsymbol{\eta}_2(k) \quad (3.13)$$

Estas duas equações podem ser representadas como um gráfico de fluxo do processo, conforme a Figura 3.3.

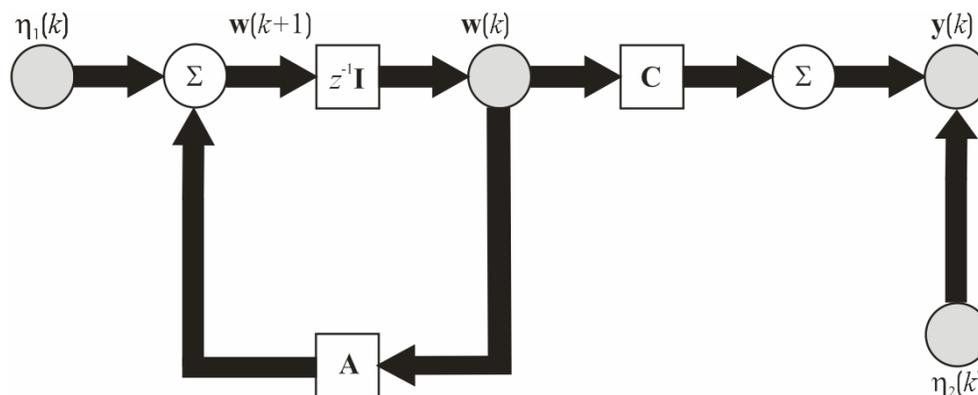


Figura 3.3 - Representação gráfica do fluxo de um sistema linear dinâmico e discreto.

Nestas equações de espaço de estado, \mathbf{A} e \mathbf{C} são matrizes previamente conhecidas, \mathbf{w} é o estado do sistema, \mathbf{y} é a saída medida e k é o instante de tempo discreto, com $k \in \mathbb{Z}$.

A variável $\boldsymbol{\eta}_1(k)$ é conhecida como ruído de processo, modelado como um processo ruído branco, gaussiano, com média zero e covariância \mathbf{Q} dada por

$$\mathbf{Q} = E[\boldsymbol{\eta}_1(k)\boldsymbol{\eta}_1^T(k)] \quad (3.14)$$

De modo semelhante, a variável $\boldsymbol{\eta}_2(k)$ é conhecida como ruído de medida e modelada como uma processo ruído branco, gaussiano, com média zero e covariância \mathbf{R} dada por

$$\mathbf{R} = E[\boldsymbol{\eta}_2(k)\boldsymbol{\eta}_2^T(k)] \quad (3.15)$$

Os ruídos $\boldsymbol{\eta}_1(k)$ e $\boldsymbol{\eta}_2(k)$ são considerados estatisticamente independentes, ou seja,

$$E[\boldsymbol{\eta}_1(n)\boldsymbol{\eta}_2^T(k)] = \mathbf{0} \text{ para qualquer } n \text{ e } k. \quad (3.16)$$

O filtro de Kalman é considerado um estimador ótimo caso satisfaça dois critérios:

- o valor médio da estimativa do estado é igual ao valor médio do verdadeiro estado
- a estimativa do estado possui erro com a menor variância possível

Desta maneira, podemos apresentar as seguintes equações que regem o filtro de Kalman para a operação de predição de $y(k+1)$ [HAY91]:

- Processo de adaptação:

$$\mathbf{K}(k) = \frac{\mathbf{A}\mathbf{P}(k)\mathbf{C}^T}{\mathbf{C}\mathbf{P}(k)\mathbf{C}^T + \mathbf{R}} \quad (3.17)$$

$$\mathbf{P}(k+1) = \mathbf{A}\mathbf{P}(k)\mathbf{A}^T + \mathbf{Q} - \mathbf{K}(k)\mathbf{C}\mathbf{P}(k)\mathbf{A}^T \quad (3.18)$$

- Processo de predição:

$$\hat{\mathbf{w}}(k+1) = \mathbf{A}\hat{\mathbf{w}}(k) + \mathbf{K}(k)[\mathbf{y}(k) - \mathbf{C}\hat{\mathbf{w}}(k)] \quad (3.19)$$

$$\hat{\mathbf{y}}(k+1) = \mathbf{C}\hat{\mathbf{w}}(k+1) \quad (3.20)$$

em que $\mathbf{K}(k)$ é chamado de ganho de Kalman, $\mathbf{P}(k+1)$ é a covariância dos erros de predição do estado e $\hat{\mathbf{w}}(k+1)$ e $\hat{\mathbf{y}}(k+1)$ são respectivamente a predição do próximo estado e a predição do próximo valor do processo.

Para que os resultados do filtro sejam ótimos, além dos dois critérios de otimalidade, necessita-se que as matrizes \mathbf{A} e \mathbf{C} e as covariâncias \mathbf{Q} e \mathbf{R} referentes aos ruídos do sistema analisado sejam conhecidas.

3.2.3 Proposta: estimação dinâmica dos ruídos do sistema

Conforme pudemos observar na seção anterior, o filtro de Kalman requer o conhecimento das covariâncias dos ruídos do sistema analisado para a obtenção de resultados ótimos. Na prática, inúmeras vezes nos deparamos com sistemas cujas matrizes \mathbf{Q} e \mathbf{R} , referentes às covariâncias citadas, ou são desconhecidas, ou são apenas conhecidas parcialmente. Uma modelagem incorreta tanto de \mathbf{Q} quanto de \mathbf{R} pode resultar em um desempenho sub-ótimo do filtro [HEF66] [NIS67].

Na literatura, alguns esquemas de identificação instantânea de \mathbf{Q} e \mathbf{R} foram propostos [GOD74] [MEH70]. Embora tais esquemas permitissem uma modelagem aceitável de \mathbf{R} , nem sempre conseguiam uma identificação eficiente de \mathbf{Q} . Nesta subseção apresentamos outra proposta para a estimação das estatísticas dos ruídos do sistema. Esta nova proposta baseia-se na utilização do algoritmo NLMS e no cálculo recursivo da média e variância de uma série temporal, conforme apresentado em [YOU84].

Inicialmente, considere o sistema descrito por meio das seguintes equações no espaço de estado:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \boldsymbol{\eta}_1(k) \quad (3.21)$$

$$y(k) = \mathbf{u}^T(k)\mathbf{w}(k) + \eta_2(k) \quad (3.22)$$

ou seja, \mathbf{A} é a matriz identidade e $\mathbf{C} = \mathbf{u}^T(k)$, nas equações (3.12) e (3.13). Caracterizando-se $\mathbf{w}(k)$ como o vetor de coeficientes de um filtro transversal (mostrado na Figura 3.2), $\boldsymbol{\eta}_1(k)$ pode ser visto como o ajuste dos coeficientes deste filtro por meio do algoritmo NLMS, conforme (3.11). Desta forma ,

$$\boldsymbol{\eta}_1(k) = \frac{\tilde{\mu}}{\|\mathbf{u}(k)\|^2} [\mathbf{u}(k)e(k)] \quad (3.23)$$

Conforme apresentado anteriormente, $\tilde{\mu}$ é o passo de adaptação fixo auxiliar, $\mathbf{u}^T(k)$ é o vetor de entradas do filtro e $e(k)$ é o erro de predição de $y(k)$. Já $\eta_2(k)$ pode ser modelado como o próprio $e(k)$, se considerarmos

$$\eta_2(k) = y(k) - \mathbf{u}^T(k)\mathbf{w}(k) = e(k) \quad (3.24)$$

A média $\bar{\eta}_2(k)$ e a variância $\mathbf{R}(k)$ de $\eta_2(k)$ podem ser estimadas recursivamente, da seguinte maneira [YOU84]:

$$\bar{\eta}_2(k) = \frac{k-1}{k}\bar{\eta}_2(k-1) + \frac{1}{k}\eta_2(k) \quad (3.25)$$

$$\mathbf{R}(k) = \frac{k-1}{k}\mathbf{R}(k-1) + \frac{1}{k}(\eta_2(k) - \bar{\eta}_2(k))^2 \quad (3.26)$$

Tomando-se como base estas duas equações, estamos aptos a estimar a covariância \mathbf{Q} .

Seja o vetor $\boldsymbol{\eta}_1(k) = [\eta_{1,1}(k), \eta_{1,2}(k), \dots, \eta_{1,N}(k)]^T$. Cada $\eta_{1,j}(k)$, com $j = 1, 2, \dots, N$, possui variância $q_j(k)$ estimada recursivamente conforme as equações (3.25) e (3.26). Dessa forma, estimamos a matriz de covariância $\mathbf{Q}(k)$.

$$\mathbf{Q}(k) = \begin{bmatrix} q_1(k) & \dots & 0 & 0 \\ \vdots & q_2(k) & \dots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & q_N(k) \end{bmatrix} \quad (3.27)$$

A estimação dinâmica das matrizes \mathbf{Q} e \mathbf{R} é feita durante a operação do filtro de Kalman, em um sistema descrito pelas equações (3.21) e (3.22). Esta é a nossa proposta para predição adaptativa.

3.3 Avaliação do preditor proposto

Proposto um novo preditor, nossa tarefa agora é avaliá-lo. Devido ao fato deste preditor ser baseado no filtro de Kalman, nossa primeira avaliação fundamenta-se na verificação das estatísticas dos ruídos do sistema no qual ele é aplicado. A segunda avaliação se resume em verificarmos se o preditor proposto possui, segundo algum critério estabelecido,

um desempenho de predição superior a outros preditores clássicos, ou mais simples. A terceira avaliação consiste em comprovarmos a eficiência da predição de uma série de amostras de expoentes de Hölder pontuais estimados a partir de um sinal de tráfego, em relação à predição das próprias amostras deste sinal de tráfego.

Para as três avaliações, quando necessário, utilizamos amostras das séries de tráfego *dec-pkt-1*, *dec-pkt-2* e *lbl-pkt-5*, assim como seus expoentes de Hölder pontuais estimados com uso das janelas 11, 12 e 13. Como o preditor proposto foi aplicado apenas a um passo adiante, a estratégia para uma predição a um passo mais longo foi a de aumentar o intervalo de tempo entre amostras das séries utilizadas. Desta forma, utilizamos os intervalos (escalas) de tempo de 100 ms e 200 ms.

Para a escolha da ordem do preditor proposto, considerou-se que uma menor quantidade de amostras na entrada do preditor resulta em um menor processamento de informações e armazenamento em memória. No entanto, caso a ordem escolhida para o preditor seja pequena demais, pode haver queda na qualidade do valor predito. Desta forma, a ordem do preditor que refletiu o melhor compromisso entre os dois fatores citados, foi de $N=7$. Este valor foi escolhido por meio de simulações com o preditor proposto. Em relação ao passo de adaptação $\tilde{\mu}$, sua escolha baseia-se inicialmente nos limites teóricos para a estabilidade do algoritmo NLMS. Estes limites são dados por $0 < \tilde{\mu} < 2$, conforme discutido na Subseção 3.2.1. Dentro destes limites, um valor demasiadamente pequeno deste parâmetro resulta em uma convergência lenta do algoritmo. No caso de um valor excessivamente grande, a convergência é atingida em poucas iterações. No entanto, o algoritmo pode ser incapaz de acompanhar as pequenas variações do sinal analisado [HAY91]. Desta forma, por meio de simulações, o valor de $\tilde{\mu}$ que resultou em um melhor comportamento do preditor proposto (segundo os critérios a serem apresentados na Subseção 3.3.2) foi de $\tilde{\mu} = 0,02$. As condições iniciais consideradas foram

$$\hat{\mathbf{w}}_0 \cong 0 \quad (3.28)$$

$$\mathbf{K}_0 = \hat{\mathbf{w}}_0 \hat{\mathbf{w}}_0^T \quad (3.29)$$

3.3.1 Avaliação dos ruídos do sistema

Conforme discutido na subseção 3.2.2, segundo a definição do filtro de Kalman, tanto o ruído de processo como o ruído de medida presentes no sistema analisado são considerados ruídos brancos, gaussianos, com média nula e descorrelacionados entre si. Nosso objetivo nesta subseção é mostrar que estes ruídos possuem as características estatísticas citadas, mesmo quando estimados de modo dinâmico. Para isto, utilizamos como sinal de entrada do preditor proposto (e conseqüentemente modelado conforme as equações 3.21 e 3.22) a série de amostras dos expoentes de Hölder pontuais, obtidas com uso da estimação com janela 12, a partir do sinal de tráfego Internet *dec1* na escala de tempo de 100 ms. As Figuras 3.4 e 3.5 nos exhibe tanto a função de autocorrelação como o histograma de um dos ruídos do vetor de ruídos de processo $\eta_1(k)$ e das amostras do ruído de medida $\eta_2(k)$. Utilizando-se as figuras, podemos visualizar, para um mesmo processo de ruído, as aparentes características: total descorrelacionamento temporal, distribuição gaussiana e média zero. Analisando conjuntamente amostras de dois ruídos distintos por meio da Figura 3.6, podemos verificar também um aparente descorrelacionamento temporal. Estas características puderam ser comprovadas utilizando-se os sinais de tráfego restantes e suas respectivas séries de amostras de expoentes de Hölder pontuais estimados. Desta forma, podemos dizer que os ruídos presentes no sistema analisado, mesmo que estimados de maneira dinâmica, apresentam as mesmas características estatísticas que o sistema sobre o qual o filtro de Kalman é definido.

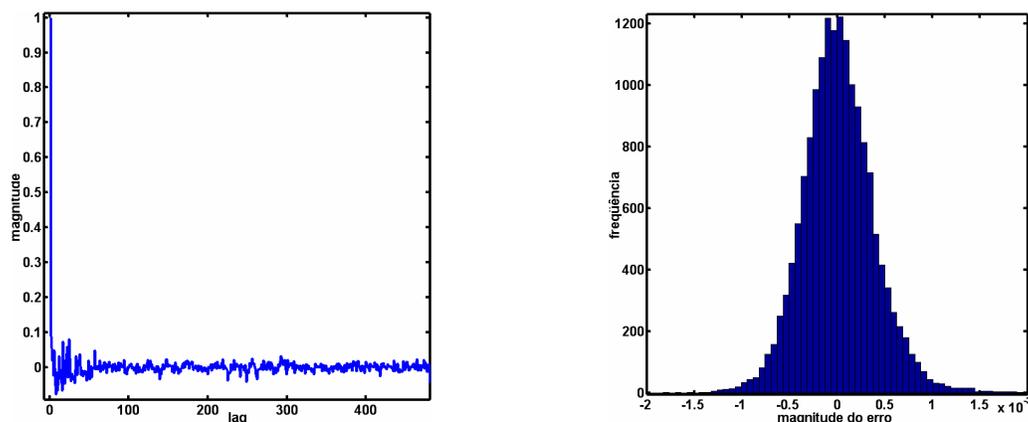


Figura 3.4 - Esquerda: função de autocorrelação das amostras de um dos ruídos contidos no vetor de ruído de processo $\eta_1(k)$. Direita: histograma dos valores das amostras do mesmo ruído.

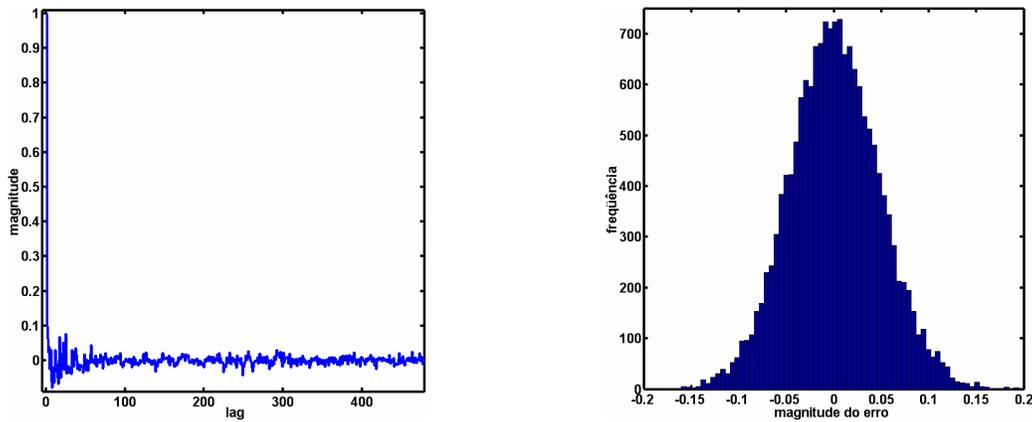


Figura 3.5 - Esquerda: função de autocorrelação das amostras do ruído de medida $\eta_2(k)$. Direita: histograma dos valores das amostras do mesmo ruído.

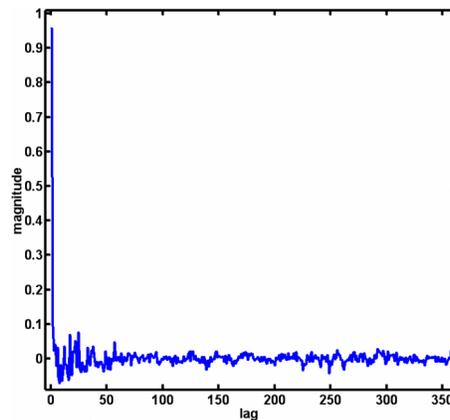


Figura 3.6 - Função de correlação cruzada entre amostras do ruído $\eta_2(k)$ e amostras de um dos ruídos do vetor $\eta_1(k)$.

3.3.2 Avaliação comparativa com outros preditores

Feita a primeira avaliação do preditor proposto, a segunda avaliação feita nesta seção consiste na comparação do desempenho deste preditor com o desempenho de outros dois preditores considerados clássicos ou mais simples, segundo algum critério.

Inicialmente, a qualidade da predição é avaliada comparativamente com uso de duas versões de erro quadrático médio normalizado (EQMN) como critério, definidas a seguir [SHA03]:

$$EQMN1 = \frac{E[(h_{pred} - h_{jan})^2]}{\sigma^2_{h_{jan}}} \quad (3.30)$$

$$EQMN2 = \frac{E[(h_{pred} - h_{jan})^2]}{E[(h_{ua} - h_{jan})^2]} \quad (3.31)$$

nos quais h_{jan} é o expoente de Hölder pontual estimado via janelamento, h_{pred} é o expoente de Hölder predito, $\sigma^2_{h_{jan}}$ é a variância dos expoentes de Hölder estimados e h_{ua} é o expoente de Hölder anterior a h_{jan} . Os EQMNs foram calculados a partir da metade das amostras preditas, quando se pôde supor que o preditor estaria estável.

Basicamente, o EQMN compara o EQM do preditor avaliado com o EQM de um preditor considerado mais simples. O EQMN1 considera a média amostral do sinal analisado como um preditor mais simples e o EQMN2 considera a última amostra disponível do sinal como outro preditor mais simples. O preditor avaliado pode ser considerado aceitável apenas se possuir ambos EQMNs menores que a unidade. Caso contrário, podemos concluir que seu desempenho será, na melhor das hipóteses, similar a algum dos preditores mais simples citados.

Além do preditor proposto, os outros dois preditores avaliados sob os mesmos critérios citados são o filtro NLMS (preditor 1) e o filtro de Kalman comum (preditor 2). A ordem do filtro e o valor do passo de adaptação $\tilde{\mu}$ do preditor 1 são iguais àqueles utilizados pelo preditor proposto. A única diferença entre o preditor proposto e o preditor2 é o fato de neste último, os ruídos do sistema possuírem média zero e variância constante. As (co)variâncias utilizadas no preditor 2 são aquelas encontradas após o término da simulação com o preditor proposto.

Tabela 3.1 - EQMNs para os expoentes de Hölder pontuais da série de tráfego *dec-pkt-1*, na escala de tempo de 100 ms, estimados usando janela 13.

	EQMN1	EQMN2
proposto	0,3246	0,6317
preditor1	0,3393	0,6605
preditor2	0,3247	0,6319

Tabela 3.2 - EQMNs para os expoentes de Hölder pontuais da série de tráfego *dec-pkt-1*, na escala de tempo de 200 ms, estimados usando janela 13.

	EQMN1	EQMN2
proposto	0,3128	0,5844
preditor1	0,3979	0,6125
preditor2	0,3130	0,5847

Tabela 3.3 - EQMNs para os expoentes de Hölder pontuais da série de tráfego *dec-pkt-1*, na escala de tempo de 200 ms, estimados usando janela 12.

	EQMN1	EQMN2
proposto	0,3104	0,5729
preditor1	0,3248	0,5994
preditor2	0,3106	0,5733

Tabela 3.4 - EQMNs para os expoentes de Hölder pontuais da série de tráfego *lbl-pkt-5*, na escala de tempo de 100 ms, estimados usando janela 13

	EQMN1	EQMN2
proposto	0,1328	0,7188
preditor1	0,1406	0,7611
preditor2	0,1329	0,7194

Tabela 3.5 - EQMNs para os expoentes de Hölder pontuais da série de tráfego *lbl-pkt-5*, na escala de tempo de 200 ms, estimados usando janela 13

	EQMN1	EQMN2
proposto	0,0980	0,5904
preditor1	0,1016	0,6125
preditor2	0,0981	0,5914

Tabela 3.6 - EQMNs para os expoentes de Hölder pontuais da série de tráfego *lbl-pkt-5*, na escala de tempo de 200 ms, estimados usando janela 12

	EQMN1	EQMN2
proposto	0,1372	0,5825
preditor1	0,1423	0,6042
preditor2	0,1373	0,5826

Por meio da Tabela 3.1 até a Tabela 3.6, podemos constatar um melhor desempenho do preditor proposto em relação aos outros preditores analisados. Isto se deve ao fato de além de possuir os EQMNs menores que a unidade, o preditor proposto também possui os EQMNs menores do que aqueles referentes aos preditores 1 e 2. Em alguns casos, o preditor proposto demonstrou desempenho praticamente semelhante ao preditor2 (filtro de Kalman comum). Isto se deve, conforme já relatado, ao fato de utilizarmos as (co)variâncias encontradas após a simulação com o preditor proposto, para amostras de um dado sinal, como (co)variâncias ótimas do preditor2 para as mesmas amostras. Isoladamente, como comentado na subseção 3.2.3, encontrar tais parâmetros não é algo trivial.

Outro parâmetro útil para a verificação da eficiência de um preditor é a velocidade de convergência do algoritmo de predição utilizado. Este parâmetro é mostrado na Figura 3.7 por meio do decaimento temporal do erro quadrático de predição. Esta figura nos mostra que o preditor proposto possui o decaimento mais íngreme deste erro, em comparação com os outros preditores analisados. Com o uso de poucas amostras iniciais, o erro quadrático do preditor proposto decai e se sustenta próximo a zero. Na mesma figura, percebe-se que para atingirmos o mesmo fim com os outros preditores analisados, necessita-se de mais amostras.

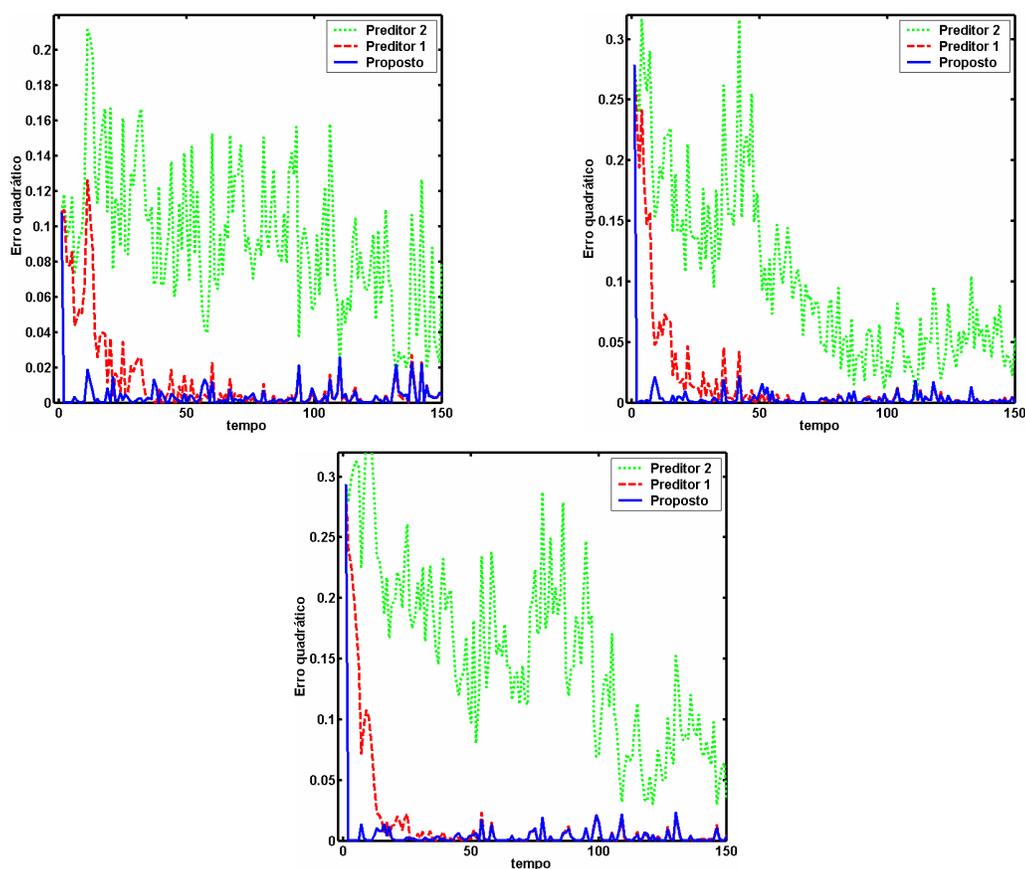


Figura 3.7 - Decaimento do erro quadrático de predição dos expoentes de Hölder de amostras da série *dec-pkt-1* na escala de tempo de 100 ms (acima à esquerda) e da série *lbt-pkt-5* nas escalas de tempo de 100 ms (acima à direita) e 200 ms (abaixo).

3.3.3 Predição das características do tráfego Internet

Feitas a primeira e a segunda avaliação do preditor proposto, nossa última avaliação consiste em comprovar a eficiência da predição de uma série de amostras de expoentes de Hölder pontuais estimados a partir de um sinal de tráfego, em relação à predição das próprias

amostras deste sinal de tráfego. Os critérios utilizados são o EQMN1 e o EQMN2, apresentados na subsecção 3.3.2.

Com uso da Tabela 3.7 até a Tabela 3.12, podemos verificar a viabilidade da predição das amostras de sinais de tráfego utilizando-se o preditor proposto, visto a presença de EQMNs menores que a unidade. Mas se compararmos estes EQMNs com aqueles encontrados na predição dos expoentes de Hölder pontuais referentes ao mesmo sinal, podemos observar que para estes últimos, os EQMNs possuem valores menores. Isto comprova uma maior facilidade na predição dos expoentes de Hölder de um sinal de tráfego do que a predição do próprio sinal.

Tabela 3.7 - EQMNs para predição de amostras da série de tráfego *dec-pkt-1*, na escala de tempo de 100 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.

	EQMN1	EQMN2
série <i>dec-pkt-1</i>	0,7474	0,7129
expoentes de Hölder pontuais (janela 13)	0,3246	0,6317
expoentes de Hölder pontuais (janela 12)	0,3757	0,6347
expoentes de Hölder pontuais (janela 11)	0,3991	0,6350

Tabela 3.8 - EQMNs para predição de amostras da série de tráfego *dec-pkt-1*, na escala de tempo de 200 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.

	EQMN1	EQMN2
série <i>dec-pkt-1</i>	0,6416	0,7729
expoentes de Hölder pontuais (janela 13)	0,3128	0,5844
expoentes de Hölder pontuais (janela 12)	0,3104	0,5729
expoentes de Hölder pontuais (janela 11)	0,3274	0,5660

Tabela 3.9 - EQMNs para predição de amostras da série de tráfego *dec-pkt-2*, na escala de tempo de 100 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.

	EQMN1	EQMN2
série <i>dec-pkt-2</i>	0,7369	0,7150
expoentes de Hölder pontuais (janela 13)	0,4299	0,6465
expoentes de Hölder pontuais (janela 12)	0,5056	0,6396
expoentes de Hölder pontuais (janela 11)	0,5400	0,6355

Tabela 3.10 - EQMNs para predição de amostras da série de tráfego *dec-pkt-2*, na escala de tempo de 200 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.

	EQMN1	EQMN2
série <i>dec-pkt-2</i>	0,5942	0,8086
expoentes de Hölder pontuais (janela 13)	0,3533	0,5799
expoentes de Hölder pontuais (janela 12)	0,3855	0,5529
expoentes de Hölder pontuais (janela 11)	0,3947	0,5533

Tabela 3.11 - EQMNs para predição de amostras da série de tráfego *lbl-pkt-5*, na escala de tempo de 100 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.

	EQMN1	EQMN2
série <i>lbl-pkt-5</i>	0,6943	0,7236
expoentes de Hölder pontuais (janela 13)	0,1328	0,7188
expoentes de Hölder pontuais (janela 12)	0,1626	0,7119
expoentes de Hölder pontuais (janela 11)	0,1965	0,7181

Tabela 3.12 - EQMNs para predição de amostras da série de tráfego *lbl-pkt-5*, na escala de tempo de 200 ms, e para predição de seus respectivos expoentes de Hölder pontuais estimados por meio de janelas de tempo.

	EQMN1	EQMN2
série <i>lbl-pkt-5</i>	0,5645	0,7845
expoentes de Hölder pontuais (janela 13)	0,0980	0,5904
expoentes de Hölder pontuais (janela 12)	0,1372	0,5825
expoentes de Hölder pontuais (janela 11)	0,1828	0,5923

3.4 Avaliação conjunta da estimação e predição dos expoentes de Hölder pontuais

Esta seção tem como proposta a comparação dos expoentes de Hölder pontuais preditos, utilizando-se o filtro proposto, com os expoentes de Hölder pontuais de referência (definidos na subseção 2.5.1). Como pudemos perceber anteriormente, a estimação destes expoentes com uso de janelas de tempo, resulta em imprecisões, à medida que o tamanho desta janela é diminuído. Além deste tipo de imprecisão, devemos considerar as imprecisões conseqüentes dos erros de predição. Assim, o objetivo desta seção é a obtenção de resultados mais conclusivos acerca da eficiência da estimação dos expoentes em janelas conjuntamente com a predição dos mesmos.

A metodologia utilizada para a avaliação é semelhante àquelas utilizadas nas subseções 2.5.2 e 3.3.2. Ela consiste na utilização do erro quadrático médio normalizado

(EQMN) como um meio de análise dos resultados. A versão do EQMN considerado nesta seção é a seguinte:

$$EQMN = \frac{E[(h_{ref} - h_{pred})^2]}{\sigma^2_{h_{ref}}} \quad (3.32)$$

em que h_{pred} é o valor predito do expoente de Hölder pontual, h_{ref} é o expoente de Hölder pontual de referência e $\sigma^2_{h_{ref}}$ é a variância dos expoentes de Hölder pontuais de referência. Conforme abordado nas subseções 2.5.2 e 3.3.2, para podermos considerar a predição aceitável, idealmente os valores dos EQMNs devem ser menores que a unidade. No entanto, quanto menor seu valor, mais correta é a predição em relação ao expoente de referência.

Os resultados exibidos a seguir são referentes às predições dos expoentes de Hölder pontuais estimados a partir de amostras das séries de tráfego que utilizamos anteriormente. Para cada série, consideramos a estimação dos expoentes com uso da janela 13, janela 12 e janela 11. Utilizamos também duas escalas diferentes para cada série: 100 e 200 ms.

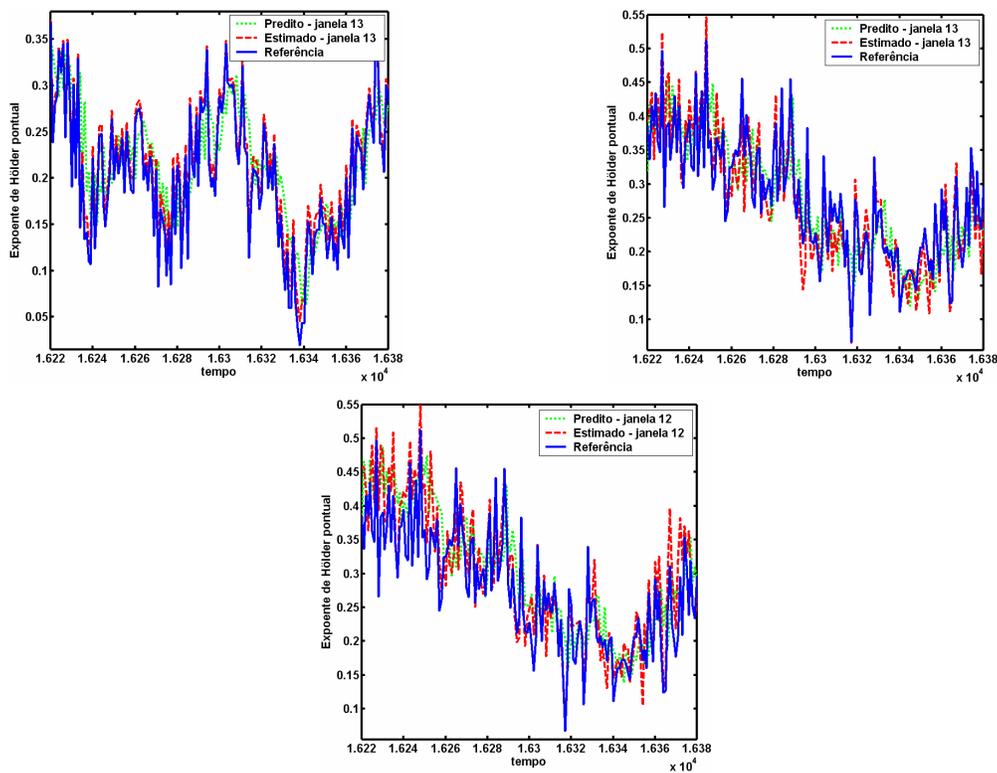


Figura 3.8 - Expoentes de Hölder pontuais de referência, expoentes de Hölder pontuais estimados e predição destes últimos, todos a partir da série *dec-pkt-1*. Acima à esquerda: escala de tempo de 100 ms e janela 13. Acima à direita: escala de tempo de 200 ms e janela 13. Abaixo: escala de tempo de 200 ms e janela 12.

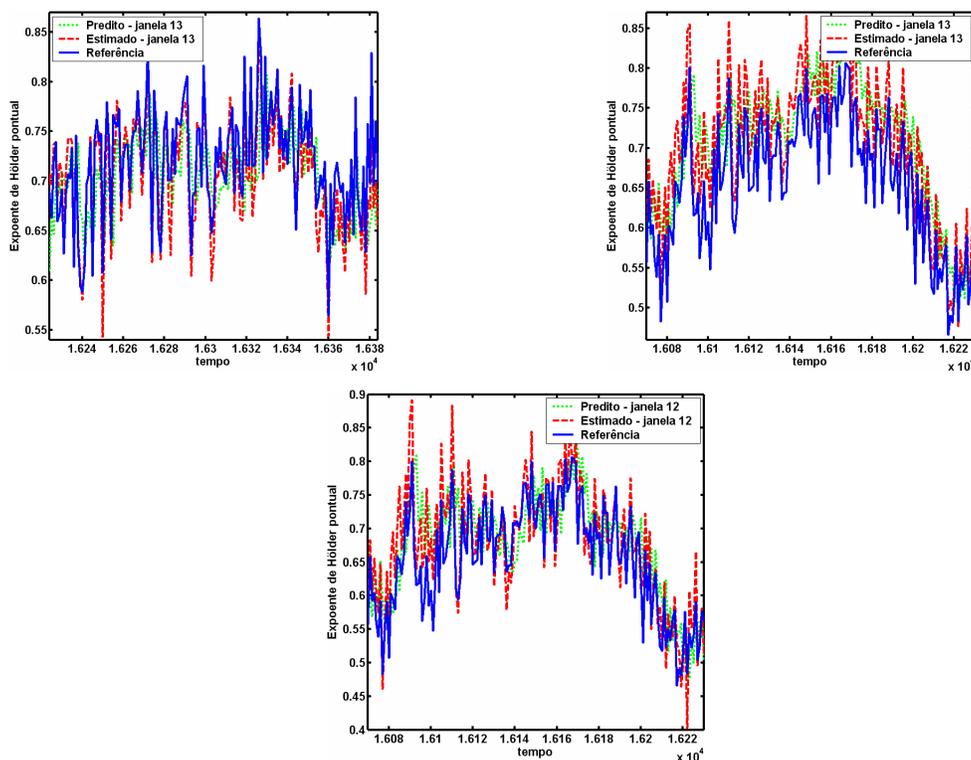


Figura 3.9 - Expoentes de Hölder pontuais de referência, expoentes de Hölder pontuais estimados e predição destes últimos, todos a partir da série *lbl-pkt-5*. Acima à esquerda: escala de tempo de 100 ms e janela 13. Acima à direita: escala de tempo de 200 ms e janela 13. Abaixo: escala de tempo de 200 ms e janela 12.

A Figura 3.8 e a Figura 3.9 nos mostram a diferença visual entre os expoentes de Hölder pontuais de referência, expoentes de Hölder pontuais estimados com janelas de tempo e a predição deste último, para duas séries, duas escalas de tempo e dois tamanhos de janelas diferentes. Por meio das Tabelas 3.13 e 3.14, podemos verificar em todas as séries utilizadas, uma deterioração da qualidade do valor do expoente em relação à diminuição do tamanho da janela. Estes resultados eram esperados, pois nesta situação, ambos erros de estimação e de predição estão sendo considerados. Isto pode ser verificado, se compararmos as Tabelas 3.13 e 3.14 com as Tabelas 2.1 e 2.2; estas últimas consideram apenas os erros referentes à estimação. Apesar do acréscimo do valor dos EQMNs da seção atual em relação aos EQMNs exibidos na subseção 2.5.2, não há valor superior à unidade; o maior valor presente é 0,7434. Desta forma pode-se concluir que a janela 11 ainda possui quantidade de amostras suficientes, para a obtenção de resultados aceitáveis do valor do expoente predito em relação ao expoente de referência.

Tabela 3.13 -EQMNs dos expoentes de Hölder pontuais das séries de tráfego na escala de tempo de 100ms.

	Janela 11	Janela 12	Janela 13
série <i>dec1</i>	0,6503	0,4688	0,3493
série <i>dec2</i>	0,7434	0,6373	0,4951
série <i>lbl5</i>	0,5795	0,5289	0,3375

Tabela 3.14 -EQMNs dos expoentes de Hölder pontuais das séries de tráfego na escala de tempo de 200ms.

	Janela 11	Janela 12	Janela 13
série <i>dec1</i>	0,6782	0,5313	0,4804
série <i>dec2</i>	0,6398	0,5219	0,4699
série <i>lbl5</i>	0,5982	0,3868	0,1767

3.5 Conclusões do capítulo

Neste capítulo, propusemos um novo esquema adaptativo de predição de amostras de um dado sinal, cujas características estatísticas não são totalmente conhecidas (como a série de expoentes de Hölder pontuais estimados a partir de amostras de um sinal de tráfego). O preditor proposto é baseado em dois filtros clássicos (filtro de Kalman e filtro NLMS) e possui como vantagem à estimação dinâmica das (co)variâncias dos ruídos do sistema.

Este preditor mostrou-se coerente com a definição do filtro de Kalman, devido à presença de ruídos com distribuição gaussiana, média zero e descorrelacionados entre si, nos sistemas nos quais houve aplicação deste preditor. O uso do preditor proposto resultou em erros quadráticos médios normalizados menores e uma queda mais rápida do erro quadrático em relação aos outros dois preditores adaptativos citados, para amostras de expoentes de Hölder pontuais estimados a partir de um sinal de tráfego Internet. Em relação a estes expoentes, os mesmos demonstraram possuir uma característica mais predictível do que o próprio sinal de tráfego que os origina (caracterizado, por exemplo, pela quantidade de bytes transmitidos em um instante de tempo). Isto pode ser verificado não somente em relação ao fato de a série de expoentes de Hölder pontuais possuir a função de autocorrelação com um decaimento mais rápido, mas também devido ao EQMN desta série ser menor do que o EQMN relativo às amostras do sinal que origina tais expoentes.

Finalmente, pudemos nos certificar que, embora os expoentes de Hölder pontuais preditos apresentem erros referentes ao processo de estimação e predição, os mesmos situam-

se dentro de uma margem aceitável. Esta constatação foi feita até para expoentes estimados utilizando-se janelas de tempo tão pequenas quanto a janela 11 (2048 amostras). O impacto causado por estes erros será melhor avaliado no próximo capítulo, ao aplicarmos os expoentes de Hölder pontuais preditos em um esquema de escalonamento.

4 ESCALONAMENTO UTILIZANDO O EXPOENTE DE HÖLDER PONTUAL

4.1 Introdução do capítulo

Este capítulo propõe um novo esquema de escalonamento de fluxos de tráfego Internet, baseado no *Generalized Processor Sharing* (GPS). A inovação deste esquema está na utilização do expoente de Hölder pontual como parâmetro de decisão da prioridade de cada fluxo, em cada intervalo de tempo considerado. Nosso objetivo é obter uma melhor distribuição da taxa de transmissão de um enlace aos fluxos presentes e uma menor perda de dados.

Para atingirmos nosso objetivo, inicialmente na seção 4.2 introduzimos o conceito de escalonamento de fluxos de dados, assim como a disciplina de escalonamento GPS. Na seção 4.3, propomos uma disciplina de escalonamento baseada no GPS e que utiliza os expoentes de Hölder pontuais de cada fluxo presente. Na mesma seção, apresentamos esta disciplina em um esquema de escalonamento, assim como os resultados das simulações. Por último, na seção 4.4, apresentamos as conclusões obtidas.

4.2 Escalonamento de fluxos de dados

4.2.1 Conceito de escalonamento

As redes de computadores, em sua essência, permitem que as sessões ou fluxos de dados compartilhem recursos, tais como a taxa de transmissão de um enlace e espaços em *buffers* nos pontos de multiplexagem. Entretanto, a competição pelo uso desses recursos pode originar situações de **contenção**.

Dada uma fila de usuários requisitando um determinado recurso, uma tarefa de escalonamento é útil para um servidor na decisão da ordem de atendimento destas requisições. Uma disciplina de escalonamento normalmente realiza duas funções:

- decide a ordem de atendimento das requisições de serviço
- gerencia a fila de requisições de serviço

As disciplinas de escalonamento são utilizadas em qualquer sistema ou camada do protocolo de redes passível de contenção de recursos. Elas desempenham um papel fundamental na provisão de qualidade de serviço às aplicações, permitindo controle diferenciado de atraso, taxa de transmissão e taxa de perda dos dados.

Em geral, os seguintes requisitos são esperados de um algoritmo de escalonamento: facilidade em sua implementação, proteção aos fluxos presentes e alocação dos recursos de modo justo. O conceito de proteção refere-se ao isolamento dos fluxos de entrada do escalonador, para que um fluxo mal-comportado não afete o desempenho dos demais fluxos. Em relação ao conceito de justiça na alocação de recursos, costuma-se utilizar como base o critério **max-min** (*max-min fairness*), definido da seguinte maneira [KES01] :

- recursos são alocados em ordem crescente de valor solicitado, normalizado pela ponderação do fluxo.
- nenhum fluxo obtém uma quota de utilização superior à solicitada .
- fluxos cujos requisitos não possam ser satisfeitos, compartilham os recursos disponíveis na proporção de suas ponderações.

Na concepção de um mecanismo de escalonamento, é interessante citarmos dois graus de liberdade que podem ser considerados: a **prioridade** de cada fluxo e a **conservação** do serviço pelo escalonador. Em relação a esta última, dizemos que um escalonador é conservativo quando torna-se inativo apenas na ausência de demanda de serviço.

Entre as disciplinas de escalonamento mais conhecidas estão a *First-Come First-Served* (FCFS) e a *Generalized Processor Sharing* (GPS) [KES01]. A disciplina FCFS é caracterizada pela simplicidade na implementação. Basicamente esta disciplina é constituída por apenas uma fila e as requisições são atendidas na ordem em que chegam. A desvantagem desta disciplina está no fato de não oferecer proteção aos fluxos, nem alocação justa de recursos. Estas duas últimas características são desejáveis para a diferenciação de serviços e estão presentes na disciplina GPS, apresentada a seguir.

4.2.2 Generalized Processor Sharing

O *Generalized Processor Sharing* (GPS) é uma das disciplinas de escalonamento mais estudadas atualmente [PAR93]. Entre suas propriedades mais importantes estão a proteção aos diferentes fluxos e o compartilhamento da taxa de serviço. Devido a estas características, o GPS é recomendado como disciplina de escalonamento em redes que necessitam de suporte à diferenciação de serviços [BEN04].

O GPS é considerado um modelo teórico e não-realístico, devido ao fato de os fluxos de entrada serem modelados como fluidos (ou pacotes infinitamente divisíveis). No entanto, existem muitas implementações práticas para tráfego em pacotes, desde as mais simples como o *Round-Robin* [KES01], até as mais elaboradas como o *Weighted Fair Queueing* (WFQ) [DEM90] e o *Worst-case Fair Weighted Fair Queueing* (WF²Q) [BEN96]. Geralmente a diferença entre uma implementação realística e o GPS teórico não é muito distante.

Considere n fluxos compartilhando um servidor GPS com taxa c . Cada fluxo i , $1 \leq i \leq n$, possui sua própria fila (veja a figura 4.1) e parâmetro φ_i . As duas regras que definem a disciplina de escalonamento GPS são as seguintes:

- o mecanismo de escalonamento é conservativo
- caso o fluxo i , durante o intervalo de tempo $[s,t]$, possua dados enfileirados à espera da transmissão³, então

$$\frac{S(i,s,t)}{S(j,s,t)} \geq \frac{\varphi_i}{\varphi_j} \text{ para todo } j = 1, 2, \dots, n \quad (4.1)$$

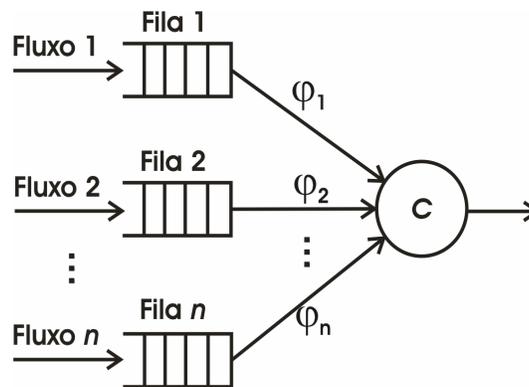


Figura 4.1 - Sistema GPS com n filas.

³ *Backlogged flow*, em inglês.

O termo $S(i, s, t)$ é a quantidade de dados do fluxo i servida no intervalo $[s, t]$. Os parâmetros $\{\varphi_i\}_{1 \leq i \leq n}$, reais e positivos, são conhecidos como **atribuições GPS** (*GPS assignments*) e determinam a porcentagem que cada fluxo compartilha da taxa de serviço c .

A desigualdade (4.1) permite que um determinado fluxo i obtenha um compartilhamento da taxa de serviço maior do que aquele determinado pelo seu φ_i . A mesma desigualdade também garante uma taxa mínima de serviço g_i a cada fluxo i dada por

$$g_i = \frac{\varphi_i}{\sum_{j=1}^n \varphi_j} c \quad (4.2)$$

Intuitivamente, um fluxo sem dados enfileirados já obtém uma taxa g_i maior que sua necessidade. Um servidor GPS certifica-se que fluxos com dados enfileirados, com quantidade desta taxa ainda insuficiente para suas necessidades, compartilhem a taxa de serviço remanescente de outros fluxos, em proporção a seus pesos. Desta forma, pode-se dizer que a disciplina de escalonamento GPS é justa em relação ao critério de justiça max-min [KES01].

4.3 Proposta: esquema de escalonamento com estimação e predição dos expoentes de Hölder pontuais

Esta seção é dedicada à modelagem e simulação de um esquema de escalonamento constituído não apenas por uma disciplina de escalonamento, mas também por um estimador dinâmico e um preditor dos expoentes de Hölder pontuais dos fluxos de dados que constituem o tráfego. Nosso interesse está na verificação da eficiência deste esquema, em relação a outros esquemas semelhantes, considerando as possíveis imprecisões de estimação e predição discutidas na seção 3.4. A disciplina de escalonamento utilizada baseia-se na disciplina GPS e é apresentada a seguir.

4.3.1 GPS com expoentes de Hölder pontuais

Nesta subseção, propomos uma nova disciplina de escalonamento, com base no modelo GPS apresentado na seção anterior. Nosso objetivo é a obtenção dinâmica das

atribuições GPS, que agora nos servem para a indicação da prioridade de cada fluxo i presente no escalonador. Em particular, isto é relativamente interessante para a alocação dinâmica da taxa de transmissão de um enlace para diferentes classes de serviço [SHI02] e para uma multiplexagem estatística mais eficiente do tráfego de redes de computadores [PAR01].

Considere n fluxos compartilhando um servidor GPS com taxa c . Cada fluxo i possui sua própria fila e um parâmetro $\varphi_i(t)$. Definimos $\varphi_i(t)$ como o grau de prioridade do fluxo i em função do tempo t e válido para o intervalo de tempo $[t, t + \Delta t)$. Para cada fluxo i , consideramos sua taxa de serviço garantida g_i dada por

$$g_i = \frac{tm_i}{\sum_{j=1}^n tm_j} c \quad (4.3)$$

em que tm_i é a taxa média do fluxo i e c a taxa total de serviço. Basicamente substituímos φ_i por tm_i na equação (4.2). Esta taxa garantida é importante para evitarmos o fenômeno da negação de serviço (*starvation*) a fluxos com menores prioridades. Nesse caso, um fluxo mais prioritário monopolizaria a taxa total de serviço. Caso algum fluxo i não esteja mais com dados enfileirados no intervalo de tempo Δt após a definição de $\varphi_i(t)$, sua taxa de serviço remanescente será redistribuída para algum fluxo ainda com dados na fila e com maior prioridade em relação a outros na mesma situação. Isto se repete até a utilização total da soma das taxas remanescentes ou até o momento que todos os fluxos não possuam mais dados a serem transmitidos neste intervalo. Ao fim deste intervalo de tempo, no instante $t + \Delta t$, novos parâmetros de prioridade $\varphi_i(t + \Delta t)$ são determinados, agora válidos para o intervalo de tempo $[t + \Delta t, t + 2\Delta t)$. Generalizando, sempre há a determinação dos parâmetros de prioridade $\varphi_i(t + k\Delta t)$ dos fluxos i para o intervalo de tempo $[t + k\Delta t, t + (k + 1)\Delta t)$, com $k \in \mathbb{Z}$.

Os parâmetros de prioridade $\{\varphi_i(t)\}_{1 \leq i \leq n}$ podem ser derivados a partir de parâmetros de tráfego tais como a taxa instantânea do tráfego e, inclusive, os expoentes de Hölder pontuais. Na realidade, como estamos lidando com alocação de taxa aos fluxos, em um intervalo de tempo futuro, faz-se necessário um método de predição. No capítulo 3, em especial na subseção 3.3.3, pudemos mostrar que os expoentes de Hölder pontuais, estimados

a partir de um sinal de tráfego Internet, possuem previsibilidade maior que o próprio sinal de tráfego. Isto nos sugere que o expoente de Hölder pontual é um parâmetro de tráfego interessante para derivarmos cada $\varphi_i(t)$.

4.3.2 Composição do escalonador proposto

O escalonador proposto é composto por três tipos de unidades distintas: unidade de estimação dinâmica dos expoentes de Hölder pontuais, unidade de previsão dos expoentes de Hölder pontuais e unidade de escalonamento dos fluxos de dados.

O primeiro tipo de unidade do escalonador trata da estimação dinâmica dos expoentes de Hölder. Especificamente ela realiza as seguintes tarefas, conforme a Figura 4.2:

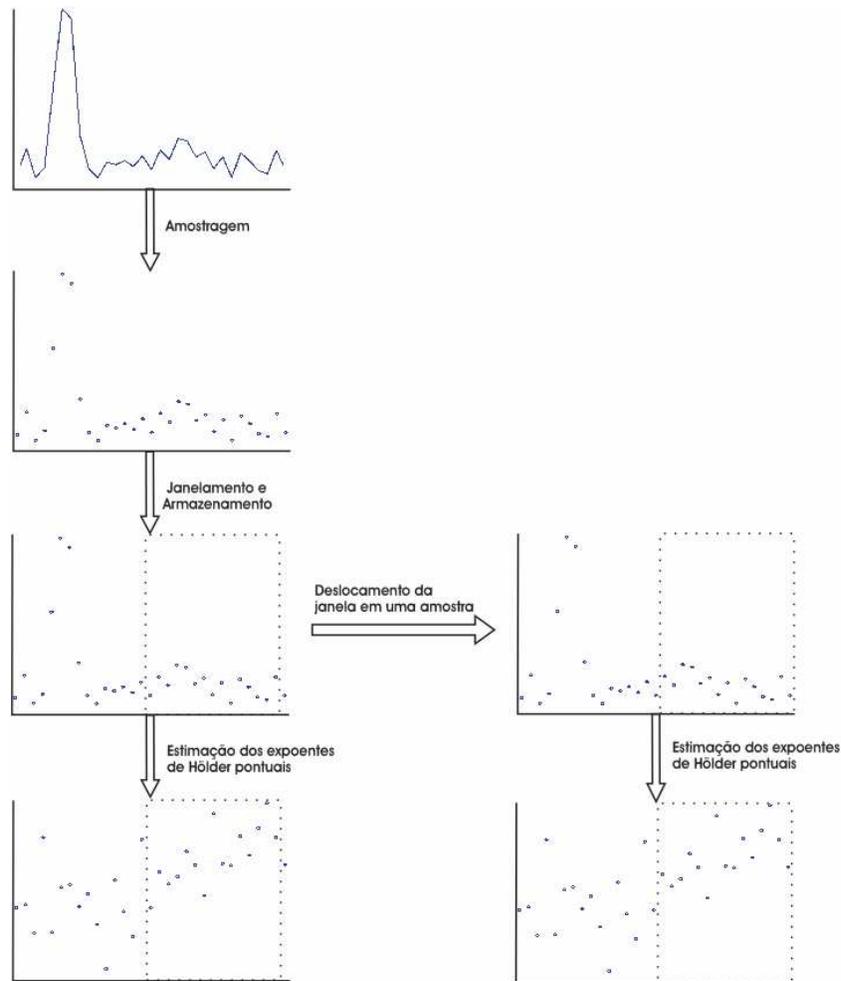


Figura 4.2 - Tarefas realizadas pela unidade de previsão dos expoentes de Hölder pontuais.

- 1 – Amostragem :** cada fluxo é amostrado em intervalos de tempo uniformes.
- 2 – Janelamento e armazenamento:** as 2^j últimas amostras que compõem uma janela j são armazenadas em um *buffer* à parte.
- 3 – Estimação dos expoentes de Hölder pontuais:** os expoentes de Hölder pontuais das amostras da janela são estimados conforme apresentado na subseção 2.3.4.

É importante esclarecermos que nesta situação, embora as janelas sejam seqüenciais, elas possuem amostras em comum. Com a chegada de uma amostra mais atual no próximo intervalo de tempo, a janela é deslocada adiante e a amostra mais antiga é excluída. Em seguida, realiza-se uma nova estimação dos expoentes de Hölder pontuais.

O segundo tipo de unidade do escalonador trata da predição do próximo expoente de Hölder pontual de cada fluxo. Este tipo de unidade depende daquela apresentada anteriormente. A unidade de predição realiza as seguintes tarefas, exibidas na Figura 4.3:

- 1 – Armazenamento dos expoentes de Hölder pontuais estimados:** as $N+1$ estimativas deste expoente, para cada fluxo, são armazenadas
- 2 – Predição do próximo expoente de Hölder pontual :** o próximo expoente de Hölder de cada fluxo é predito por meio do preditor proposto na seção 3.2.

Note que, devido ao fato de a ordem do preditor ser igual a N , necessita-se de uma amostra atual e N amostras imediatamente anteriores dos expoentes de Hölder pontuais estimados. Neste trabalho, conforme mencionado na seção 3.3, a ordem do preditor considerada foi $N = 7$.

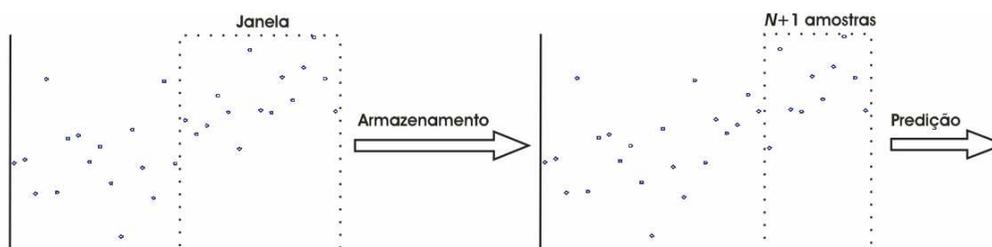


Figura 4.3 - Tarefas realizadas pela unidade de predição dos expoentes de Hölder pontuais.

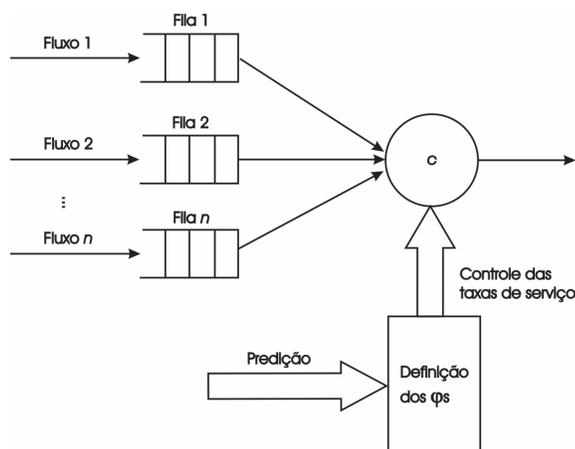


Figura 4.4 - Tarefas realizadas pela unidade de escalonamento dos fluxos.

O terceiro e último tipo de unidade do módulo de escalonamento realiza o próprio escalonamento dos fluxos. Para isso, esta unidade necessita das duas unidades apresentadas anteriormente. Em especial, esta unidade realiza as duas seguintes tarefas, ilustradas na Figura 4.4:

- 1 - Definição dos níveis de prioridade:** os parâmetros de prioridade dos fluxos, para o próximo intervalo de tempo, são determinados utilizando-se os expoentes de Hölder preditos.
- 2 – Distribuição das taxas de transmissão:** as taxas de transmissão são (re)distribuídas dentro do intervalo de tempo, tomando-se como base os parâmetros de prioridade de cada fluxo.

É importante lembrar que quanto menor for o expoente de Hölder pontual, em um instante de tempo para um determinado sinal, maior é sua variação local. Esta variação pode ser tanto em relação ao aumento instantâneo como em relação à diminuição instantânea do tráfego. Considere a situação mais crítica quando há uma rajada de dados para um determinado fluxo em um dado intervalo de tempo. Nesta situação, caso o fluxo apresente baixa prioridade, possivelmente não haverá taxa de transmissão suficiente para o mesmo, resultando em perda de dados. Desta forma, convencionamos associar um parâmetro de prioridade com valor **maior** a um expoente de Hölder pontual com valor **menor**.

4.3.3 Simulações com o escalonador proposto

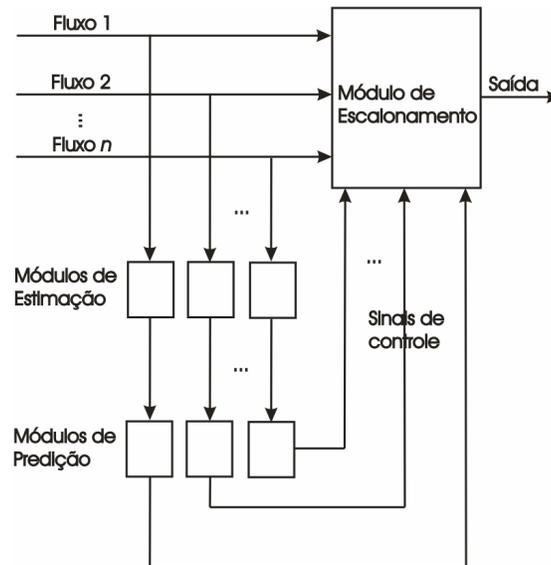


Figura 4.5 - Esquema completo do escalonador proposto.

O esquema completo de simulação está ilustrado na Figura 4.5. As simulações foram realizadas utilizando-se três fluxos formados por traços das séries *dec-pkt-1*, *dec-pkt-2* e *lbl-pkt-5*, utilizados em capítulos anteriores. Cada fluxo foi modelado como um fluido de bytes. Consideramos diferentes tamanhos de escala de tempo (100 e 200 ms) e de janelas para estimação dos expoentes de Hölder pontuais (janelas 11, 12, 13 e 14), assim como diferentes valores para a taxa de transmissão total do enlace (referente à taxa total de serviço do escalonador) e tamanhos de *buffers*. Em uma mesma simulação, os fluxos possuem *buffers* com capacidades iguais e não-variáveis.

Com objetivo de comparação, consideramos outros três critérios para a definição dos parâmetros de prioridade $\varphi(t)$ de cada fluxo:

- 1 - $\varphi_i(t)$ diretamente proporcional à taxa média de chegada de cada fluxo (definido como GPS convencional).
- 2 - $\varphi_i(t)$ inversamente proporcional ao expoente de Hölder pontual médio de cada fluxo.
- 3 - $\varphi_i(t)$ diretamente proporcional à taxa de chegada predita de cada fluxo, para o próximo instante de tempo

Para definirmos o critério 1 e 2, utilizamos o conceito de média recursiva, apresentada na subseção 3.2.3.

$$\bar{x}(k) = \frac{k-1}{k} \bar{x}(k-1) + \frac{1}{k} x(k) \quad (4.4)$$

na qual $\bar{x}(k)$ é a média temporal do processo $x(k)$ até o instante de tempo discreto k . Neste caso, $x(k)$ poderia ser tanto uma série dos expoentes de Hölder pontuais estimados (critério 1) como uma série de taxas instantâneas de chegada de bytes de um dado fluxo (critério 2).

Em relação ao critério 3, utilizamos o preditor proposto para a predição da taxa de chegada de cada fluxo no próximo instante de tempo. Conforme analisado na subseção 3.3.3, os resultados do uso do preditor proposto para as séries de tráfegos utilizadas foram considerados viáveis.

O intervalo de alocação de uma nova taxa a cada fluxo equivale ao intervalo de tempo entre duas amostras (100 ou 200 ms). Nesse ínterim, alguns fluxos podem ter uma taxa excessiva às suas necessidades momentâneas, sendo redistribuída segundo os níveis de prioridade de cada fluxo; outra consequência é o fato de alguns fluxos terem uma taxa insuficiente às suas necessidades naquele intervalo, mesmo após a realocação das taxas excedentes oriundas de outros fluxos. Desta forma, há a possibilidade de perda de pacotes devido ao transbordamento do *buffer* de alguns fluxos e, inclusive, má utilização da taxa de transmissão total do enlace.

4.3.4 Resultados das simulações

Conforme mencionamos brevemente na subseção anterior, as simulações foram realizadas em duas escalas de tempo distintas (100 e 200 ms). Inicialmente, para cada escala de tempo, obtivemos para cada um dos fluxos de dados utilizados sua respectiva média e desvio padrão. As Tabelas 4.1 e 4.5 apresentam características estatísticas dos fluxos na escala de tempo de 100 ms e 200 ms, respectivamente.

As Tabelas 4.2, 4.3 e 4.4 apresentam os resultados obtidos em três simulações, realizadas na escala de tempo de 100 ms. Nestas simulações, os tamanhos do *buffer* são 16 Kbytes e 25 Kbytes e os valores da taxa de transmissão do enlace são 2,25 Mbps e 2,75 Mbps.

Utilizando estes mesmos valores do tamanho do buffer e da taxa de transmissão do enlace, as tabelas 4.6, 4.7 e 4.8 apresentam os resultados para três simulações na escala de tempo de 200 ms.

Cada tabela retrata a porcentagem de perda de bytes, o ganho de desempenho sobre o GPS convencional e a utilização do enlace em função do parâmetro de tráfego utilizado no esquema de escalonamento. A porcentagem de perda de bytes retrata a quantidade total de bytes perdida em relação à quantidade total de bytes transmitida pelos fluxos. O ganho de desempenho sobre o GPS convencional é um indicador da eficiência do uso de um determinado parâmetro de tráfego no escalonamento em relação ao uso da taxa média (considerada para o GPS convencional). A eficiência citada está relacionada com uma menor porcentagem de perda de bytes, discutida acima. Por último, a utilização do enlace é a razão da quantidade média da taxa total utilizada pelos fluxos em relação à taxa total de transmissão do enlace.

Ainda em relação às tabelas, os parâmetros de tráfego estão referenciados de maneira abreviada, conforme apresentado a seguir:

- taxa média
- taxa instantânea predita (taxa inst.)
- expoente de Hölder pontual médio (H. médio)
- expoente de Hölder pontual de referência predito (H. pont. ref.)
- expoente de Hölder pontual estimado por janelamento, predito (H. pont. j_)

4.3.4.1 Resultados para a escala de tempo de 100 ms

Tabela 4.1 - Características dos fluxos

	Média	Desvio Padrão
fluxo 1	500 Kbps	297,12 Kbps
fluxo 2	1 Mbps	465,00 Kbps
fluxo 3	250 Kbps	348,89 Kbps

Tabela 4.2 - Performance do escalonador proposto para taxa de 2,25 Mbps e *buffer* de 16 Kbytes

Parâmetro de tráfego	Porcentagem de perda de Bytes	Ganho sobre o GPS convencional	Utilização do enlace
Taxa média	0,0146	-	0,7664
Taxa inst.	0,0135	7,35 %	0,7673
H Médio	0,0146	0 %	0,7664
H Pont Ref	0,0125	14,38 %	0,7681
H Pont J13	0,0127	13,01 %	0,7679
H Pont J12	0,0131	10,27 %	0,7676
H Pont J11	0,0136	6,85 %	0,7672

Tabela 4.3 - Performance do escalonador proposto para taxa de 2,25 Mbps e *buffer* de 25 Kbytes

Parâmetro de tráfego	Porcentagem de perda de Bytes	Ganho sobre o GPS convencional	Utilização do enlace
Taxa média	0,0099	-	0,7701
Taxa inst.	0,0087	12,12 %	0,7710
H Médio	0,0099	0 %	0,7701
H Pont Ref	0,0079	20,20%	0,7716
H Pont J13	0,0080	19,19%	0,7716
H Pont J12	0,0083	16,16%	0,7713
H Pont J11	0,0088	11,11%	0,7709

Tabela 4.4 - Performance do escalonador proposto para taxa de 2,75 Mbps e *buffer* de 25 Kbytes

Parâmetro de tráfego	Porcentagem de perda de Bytes	Ganho sobre o GPS convencional	Utilização do enlace
Taxa média	0,0017	-	0,6353
Taxa inst.	0,0013	23,53 %	0,6355
H Médio	0,0017	0 %	0,6353
H Pont Ref	0,0010	41,18 %	0,6357
H Pont J13	0,0012	29,41 %	0,6356
H Pont J12	0,0012	29,41 %	0,6356
H Pont J11	0,0012	29,41 %	0,6356

4.3.4.2 Resultados para a escala de tempo de 200 ms

Tabela 4.5 - Características dos fluxos

	Média	Desvio Padrão
fluxo 1	500 Kbps	272,93 Kbps
fluxo 2	1 Mbps	411,54 Kbps
fluxo 3	250 Kbps	306,03 Kbps

Tabela 4.6 - Performance do escalonador proposto para taxa de 2,25 Mbps e *buffer* de 16 Kbytes

Parâmetro de tráfego	Porcentagem de perda de Bytes	Ganho sobre o GPS convencional	Utilização do enlace
Taxa média	0,0221	-	0,7606
Taxa inst.	0,0214	3,17 %	0,7611
H Médio	0,0221	0 %	0,7606
H Pont Ref.	0,0210	4,98 %	0,7614
H Pont J13	0,0210	4,98 %	0,7615
H Pont J12	0,0211	4,52 %	0,7614
H Pont J11	0,0212	4,07 %	0,7613

Tabela 4.7 - Performance do escalonador proposto para taxa de 2,25 Mbps e *buffer* de 25 Kbytes

Parâmetro de tráfego	Porcentagem de perda de Bytes	Ganho sobre o GPS convencional	Utilização do enlace
Taxa média	0,0179	-	0,7639
Taxa inst.	0,0170	5,03 %	0,7645
H Médio	0,0179	0 %	0,7639
H Pont Ref.	0,0165	7,82 %	0,7649
H Pont J13	0,0165	7,82 %	0,7649
H Pont J12	0,0167	6,70 %	0,7648
H Pont J11	0,0169	5,59 %	0,7646

Tabela 4.8 - Performance do escalonador proposto para taxa de 2,75 Mbps e *buffer* de 25 Kbytes

Parâmetro de tráfego	Porcentagem de perda de Bytes	Ganho sobre o GPS convencional	Utilização do enlace
Taxa média	0,0033	-	0,6342
Taxa inst.	0,0029	12,12%	0,6345
H Médio	0,0033	0 %	0,6342
H Pont Ref.	0,0027	18,18 %	0,6347
H Pont J13	0,0028	15,15 %	0,6346
H Pont J12	0,0028	15,15 %	0,6346
H Pont J11	0,0029	12,12 %	0,6345

4.3.5 Análise dos resultados das simulações

Por meio das tabelas 4.2, 4.3, 4.4, 4.6, 4.7 e 4.8 podemos observar as seguintes tendências nos resultados das simulações:

1. A diminuição do tamanho da janela na estimação do expoente de Hölder pontual aumenta a porcentagem de perda de bytes e diminui o ganho de desempenho do escalonador proposto em relação ao GPS convencional. Isto se deve às imprecisões na própria estimação do expoente (conforme mostrado no capítulo 2).
2. Considerando-se um mesmo tamanho do *buffer*, o aumento da taxa de transmissão total proporciona uma diminuição da perda de dados e um aumento no ganho de desempenho do escalonador proposto em relação ao GPS convencional. No entanto uma quantidade maior da própria taxa de transmissão total torna-se inutilizada, de maneira geral.
3. Considerando-se uma mesma quantidade de taxa de transmissão total, o aumento do tamanho do *buffer* proporciona uma diminuição da perda de dados, um aumento no ganho de desempenho do escalonador proposto em relação ao GPS convencional e uma melhor utilização da taxa de transmissão do enlace.

4. A diminuição da escala de tempo entre amostras proporciona uma diminuição da perda de dados, um aumento no ganho de desempenho do escalonador proposto em relação ao GPS convencional e um melhor aproveitamento da taxa de transmissão do enlace. Isto se deve a uma maior quantidade de ajustes dentro de um mesmo intervalo de tempo, obtendo uma alocação mais coerente de taxa a cada fluxo, apesar do aumento de informações a serem processadas.
5. O uso da média dos expoentes de Hölder pontuais como parâmetro de tráfego demonstrou desempenho similar ao GPS convencional, nos fluxos analisados.
6. O uso da taxa de chegada predita com o preditor proposto, para caracterização local de cada fluxo, demonstrou um desempenho superior ao GPS comum. No entanto, seu desempenho apresentou-se inferior, na maioria dos casos, ao do escalonador proposto. Isto ocorre nas situações em que o tamanho da janela de tempo utilizada no módulo de estimação dos expoentes de Hölder pontuais é suficientemente grande.

4.4 Conclusões do capítulo

Este capítulo apresentou um esquema de escalonamento que utiliza a estimação e a predição do expoente de Hölder pontual para cada fluxo presente. Podemos concluir que este esquema, em grande parte dos casos, possui melhor desempenho em relação à perda de bytes e melhor uso da taxa de transmissão do enlace, em comparação com os outros esquemas avaliados. Para estes, utilizamos a taxa média (no GPS convencional), o expoente de Hölder pontual médio e a taxa instantânea estimada com o preditor proposto, como critério para a definição da prioridade dos fluxos.

O esquema proposto obteve seu melhor desempenho nas situações em que havia a possibilidade de um maior processamento das informações (para o uso de um tamanho de janela de tempo maior para a estimação dos expoentes) e também nas situações em que havia alta taxa de transmissão do enlace e *buffers* com maior capacidade de armazenamento.

5 CONCLUSÕES

A integração de serviços nas redes de comunicações traz consigo a necessidade de mecanismos de gerenciamento e controle de congestionamento do tráfego. O uso de mecanismos dinâmicos e preventivos requer que as características locais do tráfego sejam previstas. A predição (e assim o controle e gerenciamento) do tráfego encontra como obstáculo o fato de o mesmo possuir uma alta irregularidade e complexidade nas leis de escala em pequenas escalas de tempo.

Desta forma, este trabalho contribui com um esquema de escalonamento que incorpora o conhecimento da regularidade local do tráfego, por meio dos expoentes de Hölder pontuais. Este esquema inicialmente requer a estimação dinâmica destes expoentes para cada fluxo de tráfego. Nossa proposta apresentada para este caso, além de considerar o caráter oscilatório local, utiliza o método das janelas de tempo durante a estimação dos expoentes das amostras de cada fluxo, para operar dinamicamente. Em situações em que há a necessidade de um menor processamento de informações, o tamanho da janela utilizada pode ser diminuído, às custas de uma menor precisão das estimativas.

Os expoentes de Hölder pontuais estimados não possuem características estatísticas e temporais totalmente determináveis a priori. Como intencionamos a alocação antecipada dos recursos aos fluxos, contribuimos com um preditor adaptativo. Este preditor, em comparação com outros preditores mais simples ou considerados clássicos, demonstrou uma maior precisão nas estimativas realizadas e uma taxa de convergência mais rápida, ambas consideradas no sentido de mínimo erro quadrático médio.

Para o esquema de escalonamento, a predição dos expoentes de Hölder pontuais mostrou-se útil para a determinação das prioridades instantâneas de cada fluxo. Por meio de simulações, este esquema exibiu uma menor porcentagem de perda de bytes e uma maior utilização da taxa de serviço disponível no enlace, em comparação com outros indicadores de prioridade utilizados, mesmo com a presença das imprecisões resultantes da estimação e predição desses expoentes. Este esquema demonstrou seu melhor desempenho nas situações em que havia a possibilidade de um maior processamento das informações, de altas taxas de transmissão do enlace e *buffers* com maior capacidade de armazenamento.

Uma das possíveis extensões para este trabalho é a modelagem do esquema de escalonamento proposto para fluxos de tráfego compostos por pacotes. A disciplina de escalonamento neste caso poderia ter como base, por exemplo, o *Weighted Fair Queueing* (WFQ).

REFERÊNCIAS

- [ABR02] ABRY, P.; BARANIUK, R.; FLANDRIN, P.; RIEDI, R.; VEITCH, D. The multiscale nature of network traffic: Discovery, analysis, and modeling. *IEEE signal processing magazine*, v. 19, n. 3, p. 28-46, maio 2002.
- [ALA05] ALARCON-AQUINO, V.; BARRIA, J. A. Multiresolution FIR neural-network-based learning algorithm applied to network traffic prediction. *IEEE transactions on systems, man, and cybernetics – part C: applications and reviews*, v. PP, n. 99, p. 1-13, 2005.
- [ARN98] ARNEODO, A.; BACRY, E.; JAFFARD, S.; MUZY J. F. Singularity spectrum of multifractal functions involving oscillating singularities. *Journal of Fourier analysis and applications*, v. 4, n. 2, p. 159-174, 1998.
- [BEN04] BENSAOU, B.; SHIXIN, Z.; XIREN, C. Statistical bounds on the drop probability of assured forwarding services in DiffServ interior nodes under the processor sharing scheduling discipline. Proceedings of the 23rd *IEEE international conference on performance, computing and communications – IPCCC'04*, Phoenix, p. 223-230, abril 2004.
- [BEN96] BENNETT, J. C. R.; ZHANG, H. WF²Q: Worst-case fair weighted fair queueing. *Proceedings of IEEE INFOCOM'96*, San Francisco, p. 120-128, mar. 1996.
- [BOR00] BOR-SEN, C.; SEN-CHUEH, P.; KU-CHEN, W. Traffic modeling, prediction, and congestion control for high-speed networks: a fuzzy AR approach. *IEEE transactions of fuzzy systems*, v. 8, n.5, p. 491-508, out. 2000.
- [CRO97] CROVELLA, M. E.; BESTAVROS, A. Self-similarity in world wide web traffic: evidence and possible causes. *IEEE/ACM transactions on networking*, v. 6, n. 5, p. 835-846, dez.1997.
- [DAO95] DAOUDI, K.; LÉVY-VÉHEL, J. Speech modeling based on regularity analysis. *Proceedings of the IASTED/IEEE international conference on signal and image processing*, Las Vegas, nov. 1995.
- [DAO98] DAOUDI, K.; LÉVY-VÉHEL, J.; MEYER, Y. Construction of continuous functions with prescribed local regularity. *Journal of constructive approximation*, v. 14, n. 3, p. 349-385, 1998.
- [DAU92] DAUBECHIES, I. *Ten lectures on wavelets*. Philadelphia: SIAM, 1992.

- [DEM90] DEMMERS, A.; KESHAV, S.; SHENKER, S. Analysis and simulation of a fair queueing algorithm. *Journal of internetworking research and experience*, p. 3-26, out. 1990.
- [DIM91] DIMITRATOS, J.; GEORGAKIS, C.; EL-AASSER, M. S.; KLEIN, A. An experimental study of adaptive Kalman filtering in emulsion copolymerization. *Chemical engineering science*, v. 46, n. 12, p. 3203-3218, 1991.
- [FAL90] FALCONER, K. *Fractal geometry: mathematical foundations and applications*, Nova York: John Wiley & Sons, 1990.
- [FEL98] FELDMANN, A.; GILBERT, A. C.; WILLINGER, W.; KURTZ, T. G. The changing nature of network traffic: scaling phenomena. *Computer communications review*, v. 28, n. 2, abril 1998.
- [GEO04] GEORGIEV, A.; ALLEN, P. K. Localization methods for a mobile robot in urban environments. *IEEE transactions on robotics*, v. 20, n. 5, p. 851- 864, out. 2004.
- [GOD74] GODBOLE, S. S. Kalman filtering with no a priori information about noise – white noise case: identification of covariances. *IEEE transactions on automatic control*, v. AC-19, n. 5, p. 561-563, out. 1974.
- [GRO84] GROSSMAN, A.; MORLET, J. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM journal of mathematical analysis*, v. 15, n. 4, p. 723-736, jul. 1984.
- [GRA05] GRAVELLE, T.; MORLEY, J. C. A Kalman filter approach to characterizing the Canadian term structure of interest rates. *Applied financial economics*, v. 15, n. 10, p. 691–705, jun. 2005.
- [GUI98] GUILHENEUF, B.; LÉVY-VÉHEL, J. 2-microlocal analysis and application in signal processing. *Proceedings of international wavelets conference*, Tangier, 1998.
- [HAL00] HALL, J.; MARS, P. Limitations of artificial neural networks for traffic prediction in broadband networks. *IEE proceedings - communications*, v. 147, n. 2, p. 114-118, abr. 2000.
- [HAY91] HAYKIN, S. *Adaptive filter theory*. Nova Jersey: Prentice Hall, 1991.
- [HEF66] HEFFES, H. The effects of erroneous models on the Kalman filter response. *IEEE transactions on automatic control* (short papers), v. AC-11, p. 541-543, jul. 1966.
- [HOB58] HOBSON, E. W. *The theory of functions of a real variable and the theory of Fourier's series. Vol. I*. Nova York: Dover Publications Inc., 1958.

- [HOU05] HOUTEKAMER, P. L.; MITCHELL, H. L.; PELLERIN, G.; BUEHNER, M.; CHARRON, M.; SPACEK, L.; HANSEN, B. Atmospheric data assimilation with an ensemble Kalman filter: results with real observations. *Monthly weather review*, v. 133, n. 3, p. 604–620, mar. 2005.
- [JAF89] JAFFARD, S. Exposants de Hölder en des points donnés et coefficients d'ondelettes, *C. R. Acad. Sci*, Paris, v. 308, p. 79-81, 1989.
- [KES01] KESHAV, S. *An engineering approach to computer networking: ATM networks, the internet, and the telephone network*. Boston: Addison-Wesley, 2001.
- [KOM02] KOMNINAKIS, C.; FRAGOULI, C.; SAYED, A. H.; WESEL, R. D. Multi-input multi-output fading channel tracking and equalization using Kalman estimation. *IEEE transactions on signal processing*, v. 50, n. 5, maio 2002.
- [KRI02] KRISHNA P., M.; GADRE, V. M.; DESAI, U. B. Kalman filter based prediction of broadband traffic burstiness. *Proceedings of the international conference on acoustics, speech, and signal processing*, v. 3, p. 2137-2140, maio 2002.
- [LEL94] LELAND, W. E.; TAQQU, M. S.; WILLINGER, W.; WILSON, D. V. On the self-similar nature of Ethernet traffic (extended version), *IEEE/ACM transactions on networking*, v. 2, n. 1, p. 1-15, fev. 1994.
- [LIM00] LIM, A. O.; AB-HAMID, K. Kalman prediction method for congestion avoidance in ATM networks. *Proceedings of the IEEE TENCN conference*, Kuala Lumpur, v.1, p. 346-351, set. 2000.
- [MAL92] MALLAT, S.; HWANG, W. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, v. 38, n. 8, p. 617-643, mar. 1992.
- [MEH70] MEHRA, R. K. On the identification of variances and adaptive Kalman filtering. *IEEE transactions on automatic control*, v. AC-15, n. 2, p. 175-184, abril 1970.
- [MEY90] MEYER, Y. *Ondelettes et operateurs*. Paris: Hermann, 1990.
- [NAG67] NAGUMO, J. J.; NODA, A. A learning method for system identification". *IEEE transactions on automation control*, v. AC-12, p.283-287, jun. 1967.
- [NIS67] NISHIMURA, T. Error bounds of continuous Kalman filters and the application to orbit determination problems. *IEEE transactions on automatic control*, v. AC-12, p. 268-275, jun. 1967.
- [OST01] OSTRING, S. A. N.; SIRISENA, H. The influence of long-range dependence on traffic prediction. *Proceedings of the international communications conference*, Helsinki, p. 1092-1101, jun. 2001.

- [PAR00] PARK, K.; WILLINGER, W. *Self-similar network traffic and performance evaluation*. New York: Wiley, 2000.
- [PAR01] PARK, Y. K.; KIM, G. E. An efficient statistical multiplexing scheme for ATM real-time multimedia services. *IEEE transactions on consumer electronics*, v. 47, n. 1, p. 127 – 133, fev. 2001.
- [PAR93] PAREKH, A. K.; GALLAGER, G. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. *IEEE/ACM transactions on networking*, v. 1, n. 3, p. 344-357, jun. 1993.
- [PAX95] PAXSON, V.; FLOYD, S. Wide-area traffic: the failure of Poisson modeling. *IEEE/ACM transactions on networking*, v. 3, n. 3, p. 226-244, jun. 1995.
- [QIA04] QIAO, Y., SKICEWICZ, J.; DINDA, P. An empirical study of the multiscale predictability of network traffic. *IEEE proceedings of the international symposium on high performance distributed computing*, v. 0, p. 66-76, 2004.
- [RIE97] RIEDI, R. H.; LÉVY-VÉHEL, J. TCP traffic is multifractal: a numerical study. *Technical report*, INRIA Rocquencourt, n. 3129, mar. 1997.
- [RUP93] RUPP, M. The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes. *IEEE transactions on signal processing*, v. 41, p. 1149-1160, mar. 1993.
- [SAM94] SAMORODNITSKY, G.; TAQQU, M. *Stable non-gaussian random processes: stochastic models with infinite variance*. Nova York: Chapman and Hall, 1994.
- [SAN00] SANG, A.; LI, S. Q. A predictability analysis of network traffic. *Computer Networks*, v. 39, n. 4, p. 329-345, jul. 2002.
- [SEU00] SEURET, S.; GILBERT, A. C. Pointwise Hölder exponent estimation in data network traffic. *ITC Specialist Seminar*, Monterey, set. 2000.
- [SHA03] SHAH, K.; BOHACEK, S.; JONCKHEERE, E. On the predictability of data network traffic. *Proceedings of the 2003 american control conference*, v. 2, p. 1619-1624, 2003.
- [SHI02] SHIMONISHI, H.; MAKI, I.; MURASE, T.; MURATA, M. Dynamic fair bandwidth allocation for DiffServ classes. *Proceedings of the IEEE international conference on communications – ICC'02*, v. 4, p. 2348 – 2352, maio 2002.
- [SIR03] SIRIPONGWUTIKORN, P.; BANERJEE, S.; TIPPER, D. A survey of adaptive bandwidth control algorithms. *IEEE communications surveys and tutorials*, v. 5, n. 1, p. 14-26, 2003.

- [SLO93] SLOCK, D. T. M. On the convergence behavior of the LMS and the normalized LMS algorithms. *IEEE transactions on signal processing*, v. 41, p. 2811-2825, set. 1993.
- [SOL95] SOLO, V.; KONG, X. *Adaptive signal processing algorithms: stability and performance*. Nova Jersey: Prentice Hall, 1995.
- [STR00] STRUZIK, Z. R. Determining local singularity strengths and their spectra with the wavelet transform. *Fractals*, v. 8, n. 1, p. 163-179, mar. 2000.
- [VOO03] VOORONS, M.; GERMAIN, M.; BENIE, G.B.; KO, F. Segmentation of high resolution images based on the multifractal analysis. *Proceedings of the international geoscience and remote sensing symposium*, Toulouse, v.6, p. 353-3533, jul. 2003.
- [WES04] WEST, B. J.; SCAFETTA, N.; COOKE, W. H.; BALOCCHI, R. Influence of progressive central hypovolemia on hölder exponent distributions of cardiac interbeat intervals. *Annals of biomedical engineering*, v. 32, n. 8, p. 1077-1087, aug. 2004.
- [YOU84] YOUNG, P. *Recursive estimation and time-series analysis*, Nova York: Springer-Verlag, 1984.

ARTIGOS PUBLICADOS

- [JO105] JORGE, C.; VIEIRA, F. H. T.; LEE, L. L. Escalonamento GPS baseado na regularidade local do tráfego Internet. *XXII Simpósio brasileiro de telecomunicações – SBrT'05*, Campinas, set. 2005.
- [JO205] JORGE, C.; VIEIRA, F. H. T.; LEE, L. L. Predição adaptativa do expoente de Hölder para tráfego multifractal de redes. *XXVIII Congresso nacional de matemática aplicada e computacional*, São Paulo, set. 2005.