



**Paulo Jorge
Salvador Serra
Ferreira**

**Modelação de tráfego em redes de
telecomunicações: modelos Markovianos e
baseados em sistemas de Lindenmayer**

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Electrotécnica, realizada sob a orientação científica do Doutor Rui Jorge Morais Tomaz Valadas, Professor Associado do Departamento de Electrónica e Telecomunicações da Universidade de Aveiro.

o júri

presidente

Doutor José Carlos da Silva Neves

Professor Catedrático da Universidade de Aveiro

Doutor Carlos Alberto de Carvalho Belo

Professor Associado do Instituto Superior Técnico da Universidade Técnica de Lisboa

Doutor António Manuel Pacheco Pires

Professor Associado do Instituto Superior Técnico da Universidade Técnica de Lisboa

Doutor Rui Jorge Morais Tomaz Valadas

Professor Associado com Agregação da Universidade de Aveiro

Doutor Paulo Manuel Martins de Carvalho

Professor Auxiliar da Universidade do Minho

Doutor Amaro Fernandes de Sousa

Professor Auxiliar da Universidade de Aveiro

agradecimentos

Ao meu orientador, Prof. Doutor. Rui Valadas, cujos conhecimentos, compreensão e paciência, contribuíram em grande medida para o sucesso do meus trabalhos de doutoramento.

Ao Prof. Doutor António Pacheco pelas sugestões e apoio dado em algumas partes do trabalho.

À minha mãe, por todo o apoio, carinho, amor e por ter lutado para me proporcionar as condições que permitiram atingir os meus objectivos académicos e profissionais.

À minha namorada, Fernanda, pelo apoio e compreensão sempre presentes durante a elaboração deste trabalho.

A todos os meus colegas e amigos que me ajudaram, dia a dia, com os seus conhecimentos, e também com um grande companheirismo que tornou o trabalho mais agradável.

Esta tese foi elaborada com o apoio financeiro (bolsa BD/19781/99) da Fundação para a Ciência e a Tecnologia (FCT) e do Fundo Social Europeu (FSE) no âmbito do III Quadro Comunitário de Apoio.

resumo

A modelação de tráfego tem uma importância cada vez maior na gestão e dimensionamento das redes de telecomunicações. Os modelos de tráfego são utilizados, por exemplo, no dimensionamento de ligações e armazéns de pacotes tendo em conta os efeitos de multiplexagem estatística e em análise de desempenho de redes. Na Internet, a complexidade associada aos processos de geração e controle do tráfego, bem como a diversidade de aplicações, veio introduzir um conjunto de características peculiares no tráfego, de que são exemplos a auto-similaridade, a dependência longa e a multifractalidade. Estas características têm um impacto muito significativo no desempenho da rede e necessitam portanto de ser devidamente modeladas. Esta Tese propõe um conjunto de modelos de tráfego, capazes de descrever estes comportamentos peculiares, e que podem ser classificados em dois tipos: modelos Markovianos e modelos baseados em sistemas de Lindenmayer. Em ambos os tipos consideram-se modelos para tráfego com pacotes de comprimento fixo e pacotes de comprimento variável.

Na primeira parte da Tese são propostos dois modelos Markovianos e os respectivos procedimentos de inferência de parâmetros. O primeiro modelo é um processo de Poisson modelado à Markov em tempo discreto (dMMPP) que caracteriza as chegadas de pacotes. É obtido por sobreposição de um dMMPP sem memória com um número arbitrário de estados (M -MMPP) e de um conjunto de dMMPPs com dois estados (2-dMMPPs). Para inferir os parâmetros do processo, os 2-dMMPPs são utilizados para aproximar a função de autocovariância empírica e o M -MMPP é utilizado para modelar a função massa de probabilidade empírica tendo em conta as restrições impostas pelos 2-dMMPPs. O número de estados do processo pode ser ajustado de acordo com as características do tráfego a modelar. O segundo modelo é um processo de Markov com chegadas em rajada em tempo discreto (dBMAP) e corresponde a uma extensão do primeiro que permite incorporar o tamanho dos pacotes. Neste processo as chegadas de pacotes ocorrem de acordo com um dMMPP e os comprimentos seguem uma distribuição geral que depende da fase do dMMPP subjacente. Para inferir os parâmetros do dMMPP subjacente é utilizado o procedimento do primeiro modelo.

Na segunda parte da Tese são propostos modelos de tráfego baseados em sistemas de Lindenmayer (sistemas-L) e os respectivos procedimentos de inferência de parâmetros. Os sistemas-L foram introduzidos em 1968 por A. Lindenmayer para modelar o crescimento de plantas. Um sistema-L gera iterativamente sequências de símbolos progressivamente maiores, a partir de um símbolo inicial, por aplicação sucessiva de regras de produção. Para definir modelos de tráfego baseados em sistemas-L, os símbolos são interpretados como taxas de chegadas ou comprimentos médios de pacotes e cada iteração é associada a uma escala temporal do tráfego. São propostos um modelo para caracterizar as chegadas de pacotes e três modelos para caracterizar simultaneamente as chegadas e os comprimentos dos pacotes. Estes modelos conseguem capturar as características multiescalares e multifractais do tráfego.

Os modelos propostos nesta Tese foram testados com tráfego medido e foram avaliados comparando (i) as estatísticas de primeira e segunda ordem, (ii) o desempenho sobre uma fila de espera e (iii) as características de similaridade escalar, do tráfego medido com as mesmas do tráfego gerado a partir dos modelos inferidos. Os resultados obtidos mostram que os modelos propostos são, em geral, capazes de reproduzir de forma bastante rigorosa as principais características do tráfego.

abstract

Traffic modeling has an increasing importance in the management and dimensioning of telecommunications networks. Traffic models are used, for example, in the dimensioning of links and buffers while considering the effects of statistical multiplexing and in network performance analysis. In the Internet, the complexity associated with the generation and traffic control mechanisms, as well as the diversity of applications and services, introduced a set of peculiar traffic characteristics, such as self-similarity, long range dependence and multifractality. These characteristics have a strong impact on the network performance and, therefore, need to be properly modeled. This Thesis proposes a set of traffic models, which are able to describe these peculiar behaviors, and that can be classified in two classes: Markovian models and models based on Lindenmayer systems. In both cases, we propose models for traffic with fixed and variable packet size.

In the first part of the Thesis we propose two Markovian models and the respective parameter inference procedures. The first model is a Markov modulated Poisson process in discrete time (dMMPP) which characterizes the packet arrivals. It is obtained by superposing a memoryless dMMPP with an arbitrary number of states (M -dMMPP) and a set of dMMPPs with two states (2-dMMPPs). In order to infer the parameters, the 2-dMMPPs are used to fit the empirical autocovariance function and the M -dMMPP is used to fit the empirical probability mass function considering the restrictions imposed by the 2-dMMPPs. The number of states of the process can be adjusted according to the traffic characteristics. The second model is a batch Markovian arrival process in discrete time (dBMAP). It extends the first model by allowing the modeling of the packet size. In this process the packet arrivals occur according to a dMMPP and the packet sizes have a general distribution which depends on the phase of the subjacent dMMPP. The inference procedure of the first model is used to infer the parameters of the subjacent dMMPP.

In the second part of the Thesis we propose traffic models based on Lindenmayer systems (L-Systems) and the respective parameter inference procedures. L-Systems were introduced in 1968 by A. Lindenmayer as a method to model plant growth. Starting from an initial symbol, an L-System generates iteratively progressively longer sequences of symbols, by successive application of production rules. In order to define traffic models based on L-Systems, the symbols are interpreted as arrival rates or mean packet sizes and each iteration is associated with a time scale of the traffic. We proposed one model to characterize the packet arrivals and three models to characterize simultaneously the packet arrivals and the packet sizes with different levels of detail. These models are able to capture the multiscaling and multifractal characteristics of the traffic. The proposed models were tested using measured traffic and were evaluated by comparing (i) the first and second order statistics, (ii) the queuing behavior and (iii) scaling characteristics, of the measured traffic and of synthetic traffic generated according to the inferred models. The obtained results show that the proposed models are, in general, able to reproduce rigorously the main traffic characteristics.

Conteúdo

1	Introdução	1
1.1	Modelos de tráfego e suas aplicações	1
1.2	Motivações da Tese	3
1.3	Estrutura da Tese	4
1.4	Enquadramento do trabalho	6
1.5	Principais contribuições	7
2	Estado da arte	9
2.1	Introdução	9
2.2	O conceito de auto-similaridade	10
2.3	Similaridade escalar	11
2.3.1	Processos com dependência longa (ou memória longa)	14
2.3.2	Processos monofractais	15
2.3.3	Processos multifractais	15
2.3.4	Métodos de identificação e quantificação	16
2.3.5	Diagramas lineares multiescalares	19
2.4	Modelos de tráfego	20
2.4.1	Processos de renovamento	20
2.4.2	Modelos Markovianos	22
2.4.3	Processos modulados à Markov	23
2.4.4	Modelos de fluídos	24
2.4.5	Modelos auto-regressivos	24
2.4.6	Modelos TES	25
2.4.7	Modelos auto-similares	27
2.4.8	Modelos multifractais	27
2.5	Conclusão	29

3	Modelos Markovianos para o processo de chegadas de pacotes	31
3.1	Introdução	31
3.2	Processos de Poisson modulados à Markov em tempo discreto (dMMPP)	33
3.3	Modelo $M2^L$ -dMMPP	36
3.3.1	Descrição	36
3.3.2	Procedimento de inferência dos parâmetros	39
3.4	Resultados	46
3.4.1	$M2^L$ -dMMPP	47
3.4.2	$M2$ -dMMPP	47
3.4.3	Comparação e avaliação	47
3.5	Conclusões	51
4	Modelos Markovianos para os processos de chegadas e de tamanhos de pacotes	53
4.1	Introdução	53
4.2	Modelo dBMAP	54
4.2.1	Descrição	54
4.2.2	Procedimento de inferência de parâmetros	57
4.3	Resultados	58
4.4	Conclusões	61
5	Modelos com base em sistemas de Lindenmayer para o processo de chegadas de pacotes	63
5.1	Introdução	63
5.2	Sistemas de Lindenmayer	64
5.3	Modelo de tráfego para o ritmo de chegada de pacotes	68
5.3.1	Procedimento de inferência do modelo de sistema-L único	71
5.4	Relação com as cascatas conservativas	73
5.5	Resultados numéricos	74
5.6	Conclusões	78
6	Modelos com base em sistemas de Lindenmayer para os processos de chegadas e de tamanhos de pacotes	81
6.1	Introdução	81
6.2	Modelos de tráfego	82
6.2.1	Sistema-L duplo	82

6.2.2	Sistema-L conjunto	85
6.2.3	Sistema-L com FMP	87
6.3	Procedimentos de inferência de parâmetros	88
6.3.1	Sistema-L duplo	88
6.3.2	Sistema-L conjunto	91
6.3.3	Sistema-L com FMPs	92
6.4	Relação com cascatas conservativas	94
6.5	Resultados numéricos	94
6.5.1	Modelos inferidos	95
6.5.2	Resultados comparativos	97
6.6	Conclusões	106
7	Conclusões	111
7.1	Principais conclusões	111
7.2	Sugestões para trabalho futuro	113
A	Capturas de tráfego	117
A.1	Introdução	117
A.2	Medições efectuadas	117
A.2.1	Equipamento	117
A.2.2	Ferramentas de captura e análise	118
A.3	Apresentação e caracterização das capturas de tráfego usadas nesta Tese	118
	Referências	123

Lista de Figuras

2.1	Criação da poeira de Cantor.	11
2.2	Conjunto de Cantor uni-dimensional.	11
2.3	Conjunto de Cantor uni-dimensional com redistribuição de massa assimétrica (1/3 à esquerda, 2/3 à direita).	12
2.4	Realizações de processos fBm com diferentes valores de H	16
2.5	Exemplo de um diagrama de energias em escala logarítmica.	19
2.6	Exemplo de um diagrama linear multiescalar, que revela existência de similaridade multifractal (não trivial).	21
2.7	Representação da evolução de uma cascata aleatória.	29
3.1	Exemplo da sobreposição de 2 2-dMMPPs.	35
3.2	Sobreposição de um modelo M -dMMPP e L modelos 2-dMMPP.	37
3.3	Diagrama de fluxos do procedimento de inferência do modelo $M2^L$ -dMMPP.	40
3.4	Identificação das escalas temporais de interesse.	42
3.5	Algoritmo para o cálculo do número de estados e taxas de chegada do M -dMMPP.	44
3.6	Exemplo do algoritmo de ajuste de uma função a uma soma pesada de funções de probabilidade de Poisson. (À esquerda, de cima para baixo) função que se pretende ajustar e funções diferença após uma, $h^{(1)}$, e duas, $h^{(2)}$, iterações. (À direita, de cima para baixo) Funções de probabilidade de Poisson inferidas na primeira e segunda iteração, g_{φ_1} e g_{φ_2}	46
3.7	Ajuste da função massa de probabilidade, Bellcore.	48
3.8	Ajuste da função massa de probabilidade, UA.	48
3.9	Ajuste da função de autocovariância, Bellcore.	49
3.10	Ajuste da função de autocovariância, UA.	49
3.11	Rácio de perda de pacotes com uma utilização de 70%, Bellcore.	49

3.12	Atraso médio na fila de espera com uma utilização de 70%, Bellcore. . .	49
3.13	Rácio de perda de pacotes com uma utilização de 90%, Bellcore.	50
3.14	Atraso médio na fila de espera com uma utilização de 90%, Bellcore. . .	50
3.15	Rácio de perda de pacotes com uma utilização de 90%, UA.	50
3.16	Atraso médio na fila de espera com uma utilização de 90%, UA.	50
3.17	Rácio de perda de pacotes com uma utilização de 98%, UA.	51
3.18	Atraso médio na fila de espera com uma utilização de 98%, UA.	51
4.1	Exemplo de um dBMAP com 4 estados.	56
4.2	Ajuste da função massa de probabilidade, processo TP, Bellcore.	59
4.3	Ajuste da função massa de probabilidade, processo TP, UA.	59
4.4	Ajuste da função massa de probabilidade, processo BPS, Bellcore.	60
4.5	Ajuste da função massa de probabilidade, processo BPS, UA.	60
4.6	Ajuste da função de autocovariância, processo BPS, Bellcore.	60
4.7	Ajuste da função de autocovariância, processo BPS, UA.	60
4.8	Rácio de perda de pacotes com uma utilização de 90%, Bellcore.	61
4.9	Atraso médio na fila de espera com uma utilização de 90%, Bellcore. . .	61
4.10	Rácio de perda de pacotes com uma utilização de 70%, Bellcore.	61
4.11	Atraso médio na fila de espera com uma utilização de 70%, Bellcore. . .	61
4.12	Rácio de perda de pacotes com uma utilização de 90%, UA.	62
4.13	Atraso médio na fila de espera com uma utilização de 90%, UA.	62
4.14	Rácio de perda de pacotes com uma utilização de 98%, UA.	62
4.15	Atraso médio na fila de espera com uma utilização de 98%, UA.	62
5.1	Exemplo de um sistema-L e das similaridades entre escalas, curva de Koch.	65
5.2	Exemplo de sistemas-L aplicado ao crescimento de plantas, erva.	66
5.3	Exemplo de sistema-L aplicado ao crescimento de plantas, arbusto. . . .	66
5.4	Exemplo de sistemas-L, labirinto.	67
5.5	Exemplos de sistemas-L estocásticos aplicados ao crescimento de plantas, arbustos.	67
5.6	Exemplos de construção de sistemas-L estocásticos, rectângulos.	69
5.7	Representação da evolução do modelo de sistema-L único.	70
5.8	Diagrama de fluxo do procedimento de inferência do modelo de sistem-L único.	72

5.9	Média de $W_{a_j l}$ em função de λ_l e j , ritmo de chegada de pacotes da captura Bellcore.	75
5.10	Variância de $W_{a_j l}$ em função de λ_l e j , ritmo de chegada de pacotes da captura Bellcore.	75
5.11	diagrama de energias em escala logarítmica, Bellcore.	75
5.12	diagrama de energias em escala logarítmica, UA.	75
5.13	Diagramas lineares multiescalares, UA.	77
5.14	Diagramas lineares multiescalares, Bellcore.	77
5.15	Ajuste da função massa de probabilidade, UA.	77
5.16	Ajuste da função massa de probabilidade, Bellcore.	77
5.17	Ajuste da função de autocovariância, UA.	78
5.18	Ajuste da função de autocovariância, Bellcore.	78
5.19	Rácio de perda de pacotes com uma utilização de 70%, Bellcore.	79
5.20	Atraso médio na fila de espera com uma utilização de 70%, Bellcore.	79
5.21	Rácio de perda de pacotes com uma utilização de 90%, Bellcore.	79
5.22	Atraso médio na fila de espera com uma utilização de 90%, Bellcore.	79
5.23	Rácio de perda de pacotes com uma utilização de 90%, UA.	80
5.24	Atraso médio na fila de espera com uma utilização de 90%, UA.	80
5.25	Rácio de perda de pacotes com uma utilização de 98%, UA.	80
5.26	Atraso médio na fila de espera com uma utilização de 98%, UA.	80
6.1	Representação da evolução do modelo SLD.	83
6.2	Representação da evolução do modelo SLC.	86
6.3	Diagrama de fluxo do procedimento de inferência do modelo SLD.	89
6.4	Diagrama de fluxo do procedimento de inferência do modelo SLC.	91
6.5	Diagrama de fluxo do procedimento de inferência do modelo de sistema-L com FMPs.	93
6.6	Média de $W_{b_j m}$ em função de γ_m e j , tamanho médio dos pacotes da captura Bellcore.	95
6.7	Variância de $W_{b_j m}$ em função de γ_m e j , tamanho médio dos pacotes da captura Bellcore.	95
6.8	diagrama de energias em escala logarítmica, processo TMP, Bellcore.	96
6.9	diagrama de energias em escala logarítmica, processo TMP, UA.	96
6.10	diagrama de energias em escala logarítmica, processo BPS, Bellcore.	97
6.11	diagrama de energias em escala logarítmica, processo BPS, UA.	97

6.12	Diagramas lineares multiescalares com sistemas-L, processo BPS, UA. .	99
6.13	Diagramas lineares multiescalares com cascatas conservativas, processo BPS, UA.	99
6.14	Diagramas lineares multiescalares com sistemas-L, processo TMP, UA.	99
6.15	Diagramas lineares multiescalares com cascatas conservativas, processo TMP, UA.	99
6.16	Diagramas lineares multiescalares com sistemas-L, processo BPS, Bellcore.	100
6.17	Diagramas lineares multiescalares com cascatas conservativas, processo BPS, Bellcore.	100
6.18	Diagramas lineares multiescalares com sistemas-L, processo TMP, Bellcore.	100
6.19	Diagramas lineares multiescalares com cascatas conservativas, processo TMP, Bellcore.	100
6.20	Ajuste da função massa de probabilidade com sistemas-L, processo BPS, Bellcore.	102
6.21	Ajuste da função de autocovariância com sistemas-L, processo BPS, Bellcore.	102
6.22	Ajuste da função massa de probabilidade com cascatas conservativas, processo BPS, Bellcore.	102
6.23	Ajuste da função de autocovariância com cascatas conservativas, processo BPS, Bellcore.	102
6.24	Ajuste da função massa de probabilidade com sistemas-L, processo TMP, Bellcore.	103
6.25	Ajuste da função de autocovariância com sistemas-L, processo TMP, Bellcore.	103
6.26	Ajuste da função massa de probabilidade com cascatas conservativas, processo TMP, Bellcore.	103
6.27	Ajuste da função de autocovariância com cascatas conservativas, processo TMP, Bellcore.	103
6.28	Ajuste da função massa de probabilidade com sistemas-L, processo BPS, UA.	104
6.29	Ajuste da função de autocovariância com sistemas-L, processo BPS, UA.	104

6.30	Ajuste da função massa de probabilidade com cascatas conservativas, processo BPS, UA.	104
6.31	Ajuste da função de autocovariância com cascatas conservativas, processo BPS, UA.	104
6.32	Ajuste da função massa de probabilidade com sistemas-L, processo TMP, UA.	105
6.33	Ajuste da função de autocovariância com sistemas-L, processo TMP, UA.	105
6.34	Ajuste da função massa de probabilidade com cascatas conservativas, processo TMP, UA.	105
6.35	Ajuste da função de autocovariância com cascatas conservativas, processo TMP, UA.	105
6.36	Rácio de perda de pacotes com sistemas-L e uma utilização de 90%, UA.	106
6.37	Atraso médio na fila de espera com sistemas-L e uma utilização de 90%, UA.	106
6.38	Rácio de perda de pacotes com cascatas conservativas e uma utilização de 90%, UA.	106
6.39	Atraso médio na fila de espera com cascatas conservativas e uma utilização de 90%, UA.	106
6.40	Rácio de perda de pacotes com sistemas-L e uma utilização de 98%, UA.	107
6.41	Atraso médio na fila de espera com sistemas-L e uma utilização de 98%, UA.	107
6.42	Rácio de perda de pacotes com cascatas conservativas e uma utilização de 98%, UA.	107
6.43	Atraso médio na fila de espera com cascatas conservativas e uma utilização de 98%, UA.	107
6.44	Rácio de perda de pacotes com sistemas-L e uma utilização de 70%, Bellcore.	108
6.45	Atraso médio na fila de espera com sistemas-L e uma utilização de 70%, Bellcore.	108
6.46	Rácio de perda de pacotes com cascatas conservativas e uma utilização de 70%, Bellcore.	108
6.47	Atraso médio na fila de espera com cascatas conservativas e uma utilização de 70%, Bellcore.	108

6.48	Rácio de perda de pacotes com sistemas-L e uma utilização de 90%, Bellcore.	109
6.49	Atraso médio na fila de espera com sistemas-L e uma utilização de 90%, Bellcore.	109
6.50	Rácio de perda de pacotes com cascatas conservativas e uma utilização de 90%, Bellcore.	109
6.51	Atraso médio na fila de espera com cascatas conservativas e uma utilização de 90%, Bellcore.	109
A.1	Esquema das medições na rede da UA.	119
A.2	Variação temporal do ritmo de chegadas em pacotes por segundo da captura da Bellcore.	120
A.3	Variação temporal do ritmo de chegadas em bytes por segundo da captura da Bellcore.	120
A.4	Variação temporal do ritmo de chegadas em pacotes por segundo da captura da UA.	120
A.5	Variação temporal do ritmo de chegadas em bytes por segundo da captura da UA.	121
A.6	Representação do processo de chegadas em pacotes por segundo para diferentes níveis de agregação, Bellcore.	121
A.7	Representação do processo de chegadas em bytes por segundo para diferentes níveis de agregação, Bellcore.	121
A.8	Representação do processo de chegadas em pacotes por segundo para diferentes níveis de agregação, UA.	122
A.9	Representação do processo de chegadas em bytes por segundo para diferentes níveis de agregação, UA.	122

Capítulo 1

Introdução

A modelação de tráfego tem uma importância cada vez maior na gestão e dimensionamento das redes de telecomunicações. Os modelos de tráfego são utilizados, por exemplo, no dimensionamento de ligações e armazéns de pacotes tendo em conta os efeitos de multiplexagem estatística e em análise de desempenho de redes. Na Internet, a complexidade associada aos processos de geração e controle do tráfego, bem como a diversidade de aplicações, veio introduzir um conjunto de características peculiares no tráfego, de que são exemplos a auto-similaridade, a dependência longa e a multifractalidade. Estas características têm um impacto muito significativo no desempenho da rede e necessitam portanto de ser devidamente modeladas.

1.1 Modelos de tráfego e suas aplicações

A modelação de tráfego, aqui entendida enquanto modelação paramétrica, permite uma descrição relativamente detalhada do tráfego. A noção de modelo de tráfego tem sempre subjacente duas componentes: a estrutura do modelo e o método de inferência dos parâmetros. Um modelo de tráfego só existe como tal quando a par de uma estrutura matemática existe um procedimento de inferência eficiente e rigoroso para os parâmetros dessa mesma estrutura matemática. Um bom modelo de tráfego deve, por um lado, permitir na sua estrutura matemática a replicação das características do tráfego com maior impacto no desempenho da rede e, por outro, possuir um procedimento de inferência de parâmetros que consiga capturar de forma eficiente essas características. Em geral, estas características incluem: (i) as estatísticas de primeira e segunda ordem, (ii) as características de similaridade escalar (tanto ao nível do processo de chegadas como do processo do tamanho dos pacotes) e (iii) a correlação entre

os processos de chegadas e dos tamanhos dos pacotes.

Os modelos de tráfego podem ser ferramentas preciosas na concepção, na gestão e no dimensionamento de uma rede de telecomunicações. As vantagens que resultam de um uso eficiente dos mesmos são muito variadas. Os modelos de tráfego permitem fazer uma previsão relativamente rigorosa do crescimento e comportamento do tráfego a médio e longo prazo. O operador da rede pode então planear o crescimento da rede de forma faseada no tempo, garantindo em cada momento uma utilização eficiente dos recursos. Esta possibilidade de investimento faseado, permite ao operador garantir, por um lado, a satisfação dos utilizadores ao nível da qualidade do serviço e, por outro, diminuir a percentagem de capital investido sem retorno imediato. Esta vantagem económica para o operador pode resultar num menor custo do serviço oferecido aos clientes. Neste contexto os modelos de tráfego podem também ser utilizados para definir políticas de tarifação que, por exemplo, promovam a utilização da rede nos períodos com menor congestionamento.

Os modelos de tráfego permitem ainda ao operador gerir de forma mais eficiente os recursos da rede já instalados (dos comutadores e das ligações). Por exemplo, é através de uma modelação adequada do tráfego que é possível dimensionar a capacidade dos armazéns de pacotes e das ligações por forma a beneficiar dos efeitos de multiplexagem estatística. Em geral, a modelação de tráfego permite levar a cabo estudos de desempenho da rede, com base por exemplo em simulação de eventos discretos, que permitem otimizar as diferentes funções de controle da rede (como por exemplo o encaminhamento, o controle de admissão e o controle de fluxos), por forma a obter a melhor qualidade de serviço para todos os clientes. Uma outra possibilidade de utilização de modelos de tráfego neste contexto é a previsão do tráfego gerado por grupos de clientes com um determinado perfil. Esta caracterização permite, por exemplo, escolher o ponto ideal de acesso à rede para um novo cliente com um perfil de utilização específico. Ao agrupar no, mesmo ponto de acesso à rede, clientes com solicitações de largura de banda elevadas e clientes cujas necessidades de largura de banda sejam menores, é possível minimizar o número de troços de acesso à rede congestionados. Esta metodologia resulta numa qualidade de serviço média superior para todos os clientes. A previsão do tráfego gerado por um grupo de clientes com um determinado comportamento pode ser feita recorrendo a um arquivo de modelos de tráfego previamente inferidos. Seria assim possível, a partir do perfil de utilização da rede, inferir qual o tráfego total gerado por um cliente (ou grupo de clientes), sobrepondo diferentes

modelos existentes no arquivo.

As capacidades de previsão dos modelos de tráfego a curto prazo podem ser exploradas no controle da rede, concretamente em algoritmos de controle de admissão, em métodos de atribuição de recursos e em estratégias de encaminhamento adaptativas.

Em resumo os modelos de tráfego podem ser usados para estudar o desempenho de uma rede em operação ou para efectuar previsões de tráfego para efeitos de controle ou de planeamento. No capítulo 2 desta Tese será feita uma descrição dos modelos mais importantes existentes na literatura.

1.2 Motivações da Tese

Recentemente foram descobertas no tráfego da Internet características de similaridade escalar, com um impacto significativo na qualidade de serviço oferecida pela rede, e que vieram colocar em causa as metodologias de gestão de tráfego e de dimensionamento até então adoptadas. São exemplos de comportamentos do tráfego que têm em comum a propriedade da similaridade escalar a auto-similaridade, a dependência longa e a multifractalidade. Estes comportamentos são induzidos por vários factores: (i) um grande número de aplicações gera tráfego em rajada; (ii) alguns mecanismos de controle de rede, como por exemplo o controle de congestionamento usado no TCP, introduzem correlação entre chegadas de pacotes, (iii) os ficheiros descarregados pelos utilizadores possuem, em alguns casos, tamanhos com distribuição de cauda longa e (iv) a própria interacção humano-computador revela ter propriedades de cauda longa (por exemplo, no intervalo entre cliques, período de utilização e número de aplicações activas simultaneamente). Devido ao impacto que estes comportamentos peculiares do tráfego podem ter no desempenho da rede torna-se necessário desenvolver modelos adequados para os descrever.

Para além disto, verifica-se que os modelos propostos na literatura para caracterização do tráfego Internet têm, com raras excepções, esquecido a modelação dos tamanhos dos pacotes, considerando apenas o processo de chegadas. Este facto pode explicar-se por uma tentativa de reutilizar os modelos anteriormente concebidos para tráfego ATM, em que a modelação dos tamanhos dos pacotes é desnecessária uma vez que os pacotes têm comprimento fixo. No âmbito do tráfego Internet, surgiram alguns modelos que caracterizavam tanto o tamanho como a chegada de pacotes, mas assumindo independência entre os dois processos [GR01]. De facto, à data de início desta

Tese não existiam na literatura quaisquer modelos que caracterizassem as correlações existentes entre o processo de chegadas e os tamanhos dos pacotes.

Na Tese irão ser abordados dois tipos de modelos: modelos Markovianos e modelos fractais. Em cada tipo são ainda incluídos modelos que consideram o tamanho dos pacotes fixo e modelos que caracterizam o tamanho dos pacotes.

Os modelos Markovianos propostos são ou têm por base processos de Poisson modulados à Markov (MMPP). O método de inferência de MMPPs proposto nesta Tese revela-se capaz de capturar de forma precisa as estatísticas de primeira e segunda ordem e de replicar algumas das características de similaridade escalar (nomeadamente a dependência longa) até um determinado limite temporal. Os MMPPs não são intrinsecamente processos com dependência longa, mas quando são construídos com uma arquitectura apropriada conseguem replicar muito eficazmente essa propriedade do tráfego, dentro da gama de escalas temporais relevantes.

Os modelos fractais propostos são baseados nos chamados sistemas de Lindenmayer, que foram introduzidos em 1968 pelo biólogo A. Lindenmayer para descrever o crescimento das plantas. Os sistemas de Lindenmayer constituem um método recursivo de construção de sequências e caracterizam-se por um alfabeto, um axioma e um conjunto de regras de produção. Estes sistemas, são capazes de replicar estruturas com propriedades fractais, o que motivou a sua adaptação à modelação de tráfego. Neste contexto foram utilizados sistemas de Lindenmayer com regras de produção estocásticas, por isso mesmo designados por sistemas de Lindenmayer estocásticos. Para além disto, introduziu-se a possibilidade de as regras de produção serem diferentes para diferentes gamas de escalas temporais, permitindo que o modelo proposto capture as características multiescalares do tráfego.

1.3 Estrutura da Tese

A Tese encontra-se organizada em sete capítulos.

No capítulo 2 são apresentados inicialmente os conceitos e definições matemáticas das características de similaridade escalar do tráfego, nomeadamente a auto-similaridade, a dependência longa, a monofractalidade e a multifractalidade. São ainda apresentados os métodos de identificação e quantificação destas características. Na segunda parte do capítulo é feita uma descrição dos modelos de tráfego mais importantes propostos na literatura.

Os quatro capítulos seguintes propõem seis modelos de tráfego distintos e os procedimentos de inferência de parâmetros respectivos. Os capítulos seguem uma estrutura idêntica: após a descrição dos modelos e dos procedimentos de inferência de parâmetros, são apresentados um conjunto de resultados de avaliação dos mesmos. No Apêndice A é feita uma descrição dos dados reais utilizados na avaliação dos modelos, e da forma como os mesmos foram medidos.

No capítulo 3 é descrito um novo modelo de tráfego e o procedimento de inferência de parâmetros respectivo. O modelo proposto é um processo de Poisson modulado à Markov em tempo discreto (dMMPP) que caracteriza o número de chegadas em intervalos de tempo fixo. Aplica-se a tráfego onde o tamanho dos pacotes seja fixo. O modelo é obtido por sobreposição de L 2-dMMPPs e de um M -dMMPP sem memória. Os 2-dMMPP modelam a função de autocovariância e o M -dMMPP modela a função massa de probabilidade, dentro das restrições impostas pela modelação da autocovariância. Este modelo será apelidado de $M2^L$ -dMMPP. Uma característica importante deste modelo é o facto de o número de estados do dMMPP não ser fixado à priori, permitindo uma adaptação ao tipo de tráfego que se está a caracterizar.

No capítulo 4 é apresentada uma extensão do modelo anterior e do procedimento de inferência de parâmetros respectivo que inclui a possibilidade de modelar o tamanho dos pacotes. O modelo proposto é um processo Markoviano de chegadas em rajada em tempo discreto (dBMAP) onde as chegadas de pacotes ocorrem de acordo com um processo discreto de Poisson modulado à Markov (dMMPP) e cada pacote tem um tamanho definido de acordo com uma distribuição geral que depende da fase do dMMPP subjacente. O procedimento de inferência de parâmetros do dMMPP subjacente é idêntico ao descrito no capítulo 3. A caracterização do tamanho dos pacotes é feita independentemente para cada um dos estados do dMMPP subjacente.

No capítulo 5 é feita inicialmente uma apresentação dos sistemas de Lindenmayer (sistemas-L), sendo de seguida descrito um modelo de tráfego e o procedimento de inferência de parâmetros respectivo, baseado nestes sistemas. Neste capítulo o modelo está restringido à caracterização de tráfego em que os pacotes tenham um comprimento fixo.

No capítulo 6 são apresentadas três extensões ao modelo anterior, que possuem a capacidade de modelar tanto o ritmo de chegadas de pacotes como o tamanho destes. O primeiro modelo é formado por dois sistemas-L independentes, em que um modela o ritmo de chegadas de pacotes e o outro o tamanho dos pacotes, sendo por isso apelidado

de sistema-L duplo (SLD). O segundo modelo é um sistema-L bi-dimensional onde o ritmo de chegada de pacotes e o tamanho dos mesmos são modulados conjuntamente, o que originou a escolha do nome sistema-L conjunto (SLC). O terceiro modelo conjuga um sistema-L único, que modela a chegada dos pacotes, com uma caracterização do tamanho dos pacotes descrita por um conjunto de funções de probabilidade, e é chamado de sistema-L com funções massa de probabilidade (SLFMP).

Finalmente no capítulo 7 são apresentadas as principais conclusões do trabalho e as sugestões para trabalho futuro.

1.4 Enquadramento do trabalho

O trabalho desenvolvido no âmbito desta Tese foi parcialmente enquadrado em diversos projectos de Investigação e Desenvolvimento:

- CESTA (Caracterização Estatística e Simulação de Tráfego ATM), PRIQUAS (Planeamento de Redes IP com Qualidade de Serviço) e TELEMAT (Eng. de Tráfego para Redes DiffServ/MPLS), financiados pela Portugal Telecom Inovação.
- POSI/1999/CPS/34826 SCALE (Caracterização Estatística de Tráfego de Telecomunicações) e POSI/42069/CPS/2001 SCALE2 (Modelação de Tráfego e Avaliação de Desempenho em Redes Multiserviço), financiados pela Fundação para a Ciência e Tecnologia.
- EURESCOM P1112 - NEW DIMENSIONS (Network Dimensioning based on Modelling of Internet Traffic), financiado pela Comunidade Europeia.
- Projecto "Transporte de Tráfego IP para Agregados de Acessos xDSL", financiado pela ANACOM.
- COST 279 (Analysis and Design of Advanced Multiservice Networks supporting Mobility, Multimedia, and Internetworking), financiado pela Comunidade Europeia.
- Rede de excelência EURO-NGI (European Next Generation Internet) , financiada pela Comunidade Europeia.

1.5 Principais contribuições

- Proposta e análise de um modelo de tráfego baseado num processo de Poisson modulado à Markov e do procedimento de inferência de parâmetros respectivo, capaz de capturar de forma precisa as estatísticas de primeira e segunda ordem e a dependência longa do processo de chegada de pacotes, e com um número de estados que se adapta às características particulares do tráfego [SV01a, SV01b, SPV02, SPV03, SNVP04].
- Proposta e análise de um modelo de tráfego baseado num processo Markoviano de chegadas em rajada em tempo discreto e do procedimento de inferência de parâmetros respectivo, que caracteriza simultaneamente o processo de chegada de pacotes, o processo do tamanho dos pacotes e a correlação entre os dois processos, capaz de capturar de forma precisa as estatísticas de primeira e segunda ordem e a dependência longa do processo de chegada de pacotes, e com um número de estados que se adapta às características particulares do tráfego [SPV04, SNVP04].
- Proposta e análise de um modelo de tráfego baseado nos sistemas de Lindemayer e do procedimento de inferência de parâmetros respectivo, capaz de capturar as características de similaridade escalar do processo de chegada de pacotes [SNV02b, SNVP04, SNV04].
- Proposta e análise de três modelos de tráfego baseados em sistemas de Lindemayer, que caracterizam simultaneamente o processo de chegada de pacotes e o processo do tamanho dos pacotes [SNV02a, SNV03a, SNV03b, SNVP04, SNV04]. O primeiro modelo caracteriza a similaridade escalar dos processos de chegada e do tamanho dos pacotes. O segundo captura também a correlação entre as chegadas e o tamanho dos pacotes. O terceiro modelo caracteriza a similaridade escalar do processo de chegadas e a correlação deste processo com o tamanho dos pacotes.

Capítulo 2

Estado da arte

2.1 Introdução

Actualmente a Internet é uma rede fortemente heterogénea onde um número crescente de utilizadores usa uma vasta gama de aplicações, muitas das quais geram tráfego em rajada. A interacção dos indivíduos com a rede conduz a propriedades estatísticas de cauda longa, por exemplo, no intervalo entre cliques, no período de utilização, no número de aplicações activas simultaneamente e no tamanho dos ficheiros descarregados. Estas características, em conjugação com alguns mecanismos de rede, como por exemplo o controle de congestionamento usado no TCP, são responsáveis pelo aparecimento de comportamentos peculiares no tráfego, nomeadamente de características de similaridade escalar, isto é, de invariância das estatísticas com a escala temporal. São exemplos de similaridade escalar a auto-similaridade, a dependência longa e a multifractalidade.

Os comportamentos de similaridade escalar têm um impacto significativo na qualidade de serviço oferecida pela rede e vieram trazer novos desafios à modelação de tráfego da Internet. Uma das consequências foi a proposta de modelos de tráfego que incorporam estas características na sua estrutura, como por exemplo, o movimento Browniano fraccionário e as cascatas multiplicativas. No entanto, como se verá ao longo desta Tese, modelos mais clássicos, como por exemplo o processo de Poisson modulado à Markov, podem fornecer boas aproximações do comportamento de similaridade escalar, constituindo-se por vezes em soluções de modelação de tráfego superiores pois, em geral, conseguem também capturar outras características estatísticas importantes do tráfego para além da similaridade escalar.

Este capítulo tem como objectivo fazer uma descrição do estado da arte da mo-

delação de tráfego que enquadre o trabalho descrito ao longo da Tese. Está organizado do seguinte modo. Na secção 2.2 é introduzido o conceito de auto-similaridade. Na secção 2.3 é feita uma caracterização precisa do fenómeno da similaridade escalar, sendo apresentados os conceitos e definições matemáticas de auto-similaridade, dependência longa e multifractalidade; são também descritos os métodos de identificação e quantificação destas características. Na secção 2.4 é feita uma resenha dos principais modelos de tráfego, desde os processos de renascimento até os modelos intrinsecamente auto-similares e multifractais, passando pelos modelos Markovianos.

2.2 O conceito de auto-similaridade

As noções de auto-similaridade e de fractal foram introduzidas por Benoit Mandelbrot [Man77]. Estes, descrevem o fenómeno segundo o qual uma determinada propriedade de um objecto ou série temporal é preservada ao longo de várias escalas (espaciais ou temporais). Uma parte de um objecto que seja auto-similar ou fractal, quando ampliada, tem semelhanças com o objecto no seu todo. Existem na literatura alguns trabalhos que abordam estes conceitos, no âmbito da análise de tráfego, e os analisam em pormenor [PW00, ABF⁺02, RV97]. Nesta secção pretende-se apenas apresentar as noções gerais do fenómeno de auto-similaridade.

Um exemplo de estrutura com propriedades auto-similares é a chamada poeira de Cantor, que é o caso bi-dimensional de um conjunto de Cantor. A poeira de Cantor é gerada partindo de um quadrado negro. Em cada iteração: (i) cada quadrado negro é transformado em 4 quadrados negros com $1/9$ da área do quadrado que lhes deu origem e (ii) estes quadrados mais pequenos são colocados nos quatro cantos do quadrado que lhes deu origem. Se este processo de redimensionamento e translação for efectuado infinitas vezes obtém-se, no limite, o que se designa por conjunto de Cantor bi-dimensional (ou poeira de Cantor). Este processo de construção está ilustrado na figura 2.1 A auto-similaridade da poeira de Cantor pode ser observada ampliando qualquer um dos seus cantos e verificando que esta ampliação é similar ao objecto no seu todo. O conjunto de Cantor uni-dimensional (figura 2.2) pode ser interpretado como uma série temporal de tráfego, $\{X(t), t \in \mathbb{R}\}$ onde $X(t)$ apenas pode tomar os valores 0 ou 1. Esta série temporal é por vezes também apelidada de tráfego Cantor. Neste caso, $X(t) = 1$ significa que está a ser transmitido um pacote (ou byte) no instante temporal t . Se o processo de construção terminar na n -ésima iteração, então os

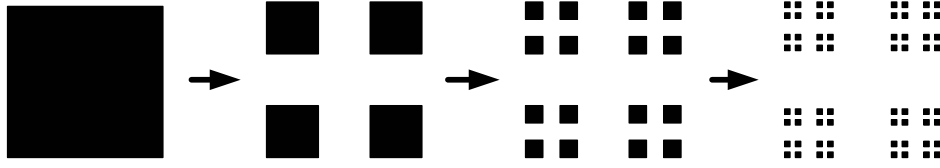


Figura 2.1: Criação da poeira de Cantor.



Figura 2.2: Conjunto de Cantor uni-dimensional.

segmentos de recta, de comprimento $1/3^n$, podem ser interpretados com o sendo os períodos de actividade com duração $1/3^n$, e os espaços entre os segmentos de recta serão os períodos de inactividade. Tráfego com uma intensidade não uniforme pode ser obtido introduzindo no processo de construção propriedades assimétricas. Por exemplo, para o conjunto de Cantor uni-dimensional em vez de as componentes à esquerda e à direita terem massa idêntica, podem ter massas diferentes. Na definição das regras de produção de alguns fractais usa-se o termo "massa", o qual representa um atributo do objecto sobre o qual as regras são aplicadas. O conceito de massa pode ser associado a qualquer quantidade numérica adicionável (área, volume, quantidade de tráfego, etc). Neste caso, o conceito de massa está associado à área de cada objecto, sendo a área o resultado do produto do tempo de actividade pela quantidade de tráfego nesse período. De notar, que a atribuição de massas está sujeita à restrição que impõe que o total de massa seja preservado em cada etapa do processo iterativo. A amplitude de cada barra poderá representar o número de chegadas de pacotes (ou bytes) no período de actividade em questão. Este tipo de construção pode então ser interpretado como uma série temporal de tráfego $\{X(t), t \in \mathbb{R}\}$ onde $X(t) \in \mathbb{R}_0^+$. Agora $X(t) = x$ significa que no instante t as chegadas ocorrem a um ritmo de x chegadas por segundo, sendo que as chegadas podem ser interpretadas como pacotes ou bytes. Um exemplo deste tipo de construção de séries temporais está ilustrado na figura 2.3, onde a distribuição de massa é feita com um peso de $1/3$ à esquerda e de $2/3$ à direita.

2.3 Similaridade escalar

A similaridade escalar de uma série temporal define-se como sendo a invariância quanto à escala das características estatísticas desta, e a consequente impossibilidade de distinguir estatisticamente a série temporal no seu todo de partes da mesma. Um pro-

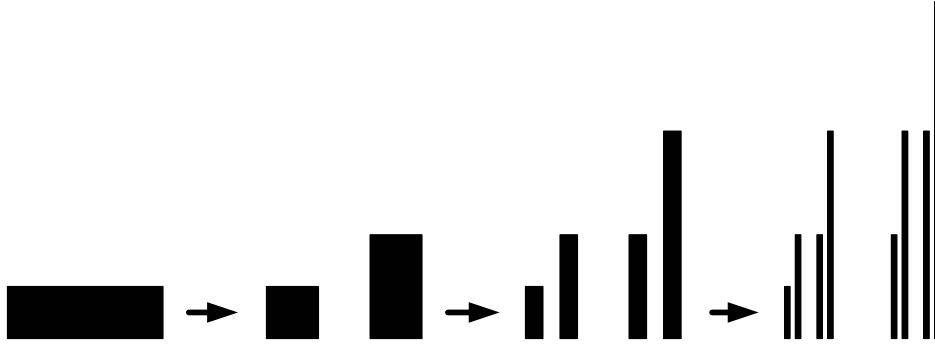


Figura 2.3: Conjunto de Cantor uni-dimensional com redistribuição de massa assimétrica (1/3 à esquerda, 2/3 à direita).

cesso $X(t)$ é exactamente auto-similar (ou auto-similar em sentido distribucional), com parâmetro de auto-similaridade $H > 0$, se

$$\{X(t), t \in \mathbb{R}\} \stackrel{d}{=} \{c^H X(t/c), t \in \mathbb{R}\}, \forall c > 0 \quad (2.1)$$

onde $\stackrel{d}{=}$ representa a igualdade em termos distribucionais para todas as dimensões finitas. Num processo exactamente auto-similar uma parte ampliada de uma realização do processo é indistinguível (estatisticamente) do todo inicial. O parâmetro de auto-similaridade H é habitualmente designado por parâmetro de Hurst. A definição (2.1) implica que os momentos de ordem q de X (caso existam) se comportam de acordo com uma lei polinomial no tempo

$$\mathbb{E}[|X(t)|^q] = \mathbb{E}[|X(1)|^q] |t|^{qH}, \forall q \in \mathbb{N}^+. \quad (2.2)$$

Um exemplo de um processo com características de auto-similaridade exacta é o movimento Browniano fraccionário (fBm).

Um processo exactamente auto-similar nunca pode ser estacionário, pois a estacionaridade requer $X(t) \stackrel{d}{=} X(t/c)$ o que é inconsistente com (2.1). É assim usual restringir a classe dos processos auto-similares aos processos auto-similares com incrementos estacionários. Considera-se que um processo X tem incrementos estacionários $Y_\delta(t)$ se

$$\{Y_\delta(t) := X(t + \delta) - X(t), t \in \mathbb{R}\} \stackrel{d}{=} \{X(\delta) - X(0)\}, \forall \delta \quad (2.3)$$

ou seja, se a distribuição do processo $Y_\delta(t)$ for independente de t . Assumindo $\mathbb{E}[Y_\delta(t)] = 0$ e $X(0) = 0$, verifica-se que $\mathbb{E}[X(t)] = 0$. Para um processo X auto-similar com incrementos estacionários, o parâmetro de auto-similaridade pertence ao intervalo entre 0 e 1, e a sua função de autocovariância caso exista tem a seguinte forma:

$$\mathbb{E}[X(t)X(s)] = \frac{\sigma^2}{2} (|t|^{2H} + |s|^{2H} - |t - s|^{2H}), \sigma^2 = \mathbb{E}[X(1)^2]. \quad (2.4)$$

A auto-similaridade exacta da expressão (2.1) pressupõe a existência de similaridade escalar para todas as escalas de 0 ao ∞ . No entanto com dados reais a similaridade escalar está logo à partida restrita a uma gama finita de escalas, devido ao tamanho limitado das amostras de dados. Além disso, a similaridade escalar pode ser observada apenas em regiões assintóticas, ou seja, para escalas muito grandes (ou muito pequenas). Por outro lado, da expressão (2.2) pode verificar-se que a auto-similaridade exacta implica que a similaridade escalar é regida apenas pelo expoente qH , para todos os momentos de ordem q . Mas quando se analisam dados reais verifica-se que momentos de ordens diferentes podem ter expoentes que são regidos por mais que um parâmetro, e alguns momentos podem nem sequer exibir similaridade escalar. Foi assim necessário desenvolver modelos que permitissem desvios da auto-similaridade exacta, considerando apenas invariância quanto à escala nas estatísticas de segunda ordem (funções de autocovariância e espectros de densidade de potência) e em regiões assintóticas das escalas.

O processo estacionário de incrementos $Y_\delta(t)$ diz-se exactamente auto-similar de segunda ordem se a sua autocovariância tomar a seguinte forma para valores de H compreendidos entre $1/2$ e 1 :

$$\mathbb{E}[Y_\delta(t)Y_\delta(t+s)] = \frac{\sigma^2}{2} (|s + \delta|^{2H} + |s - \delta|^{2H} - 2|s|^{2H}). \quad (2.5)$$

A similaridade pode ser apenas observada em regiões assintóticas das escalas. O processo de incrementos Y_δ será assintoticamente auto-similar de segunda ordem com parâmetro de Hurst H ($1/2 < H < 1$) se

$$\lim_{m \rightarrow \infty} \mathbb{E}[Y_\delta^{(m)}(t)Y_\delta^{(m)}(t+s)] = \frac{\sigma^2}{2} (|s + \delta|^{2H} + |s - \delta|^{2H} - 2|s|^{2H}) \quad (2.6)$$

onde $Y_\delta^{(m)}$ é o processo agregado de nível m de Y_δ definido por

$$Y_\delta^{(m)}(i) = \frac{1}{m} \sum_{t=m(i-1)+1}^{mi} Y_\delta(t), \quad i = 1, 2, 3, \dots \quad (2.7)$$

Verifica-se que a auto-similaridade do processo X é transferida para o processo dos incrementos, isto é,

$$Y_\delta(t) \stackrel{d}{=} c^H Y_{\delta/c}(t/c) \quad (2.8)$$

$$\mathbb{E}[Y_\delta(t)^2] = \mathbb{E}[(X(t+\delta) - X(t))^2] = \sigma^2 |\delta|^{2H}. \quad (2.9)$$

Um processo pode ainda apresentar características de similaridade escalar distintas em diferentes gamas de escalas. Um processo deste tipo será designado por processo multiescalar.

Os processos cujo espectro obedece a uma lei polinomial numa gama de frequências (ou escalas) limitada (mas suficientemente grande) são chamados habitualmente de processos $1/f$. A densidade espectral de um processo $1/f$ tem a seguinte forma:

$$\Gamma_X(v) \sim c_f |v|^{-\alpha}, v_m \leq |v| \leq v_M. \quad (2.10)$$

onde c_f é uma constante. Existem dois casos especiais de auto-similaridade assintótica onde a gama de escalas é semi-infinita: (i) os processos com dependência longa (LRD) onde se verifica similaridade para frequências baixas, $v_m \rightarrow 0$ (escalas temporais largas) e (ii) os processos monofractais para os quais se verifica similaridade para frequências altas, $v_M \rightarrow \infty$ (escalas temporais finas). Estes processos serão analisados mais em pormenor nas secções 2.3.1 e 2.3.2 respectivamente.

2.3.1 Processos com dependência longa (ou memória longa)

A dependência longa ou memória longa (LRD) de um processo é a formalização da similaridade escalar no limite das escalas mais largas, e é definido com base nas estatísticas de segunda ordem. A LRD caracteriza-se por um decaimento polinomial da função de autocovariância no limite das escalas temporais maiores.

Se $\{X(t), t \in \mathbb{R}\}$ designar um processo estocástico estacionário de segunda-ordem e r_X e Γ_X forem respectivamente a autocovariância e densidade espectral desse processo, então o processo X terá memória longa se

$$r_X(\delta) \sim c_r |\delta|^{\alpha-1}, \delta \rightarrow +\infty, \alpha \in]0, 1[\quad (2.11)$$

ou

$$\Gamma_X(v) \sim c_f |v|^{-\alpha}, v \rightarrow 0, \alpha \in]0, 1[\quad (2.12)$$

onde c_f e c_r são constantes e $\alpha = 2H - 1$. Pela restrição $\alpha \in]0, 1[$ o parâmetro de Hurst H fica restrito ao intervalo $]1/2, 1[$. Na maior parte dos casos práticos a função de autocovariância é assintoticamente monotónica, o que torna as expressões (2.11) e (2.12) equivalentes.

Um processo pode possuir LRD e não ser auto-similar no sentido distribucional. Também um processo auto-similar pode não possuir LRD. Por exemplo, o movimento Browniano (com $H = 1/2$) é auto-similar no sentido distribucional, tendo o ruído branco como o seu processo dos incrementos, e este último não possui LRD. Como a definição de LRD exige $1/2 < H < 1$ nenhum processo auto-similar com $0 < H \leq 1/2$

tem memória longa. No entanto, os processos assintoticamente auto-similares de segunda ordem são sempre LRD, e vice-versa, uma vez que nestes processos $1/2 < H < 1$.

2.3.2 Processos monofractais

A noção de processo monofractal e multifractal recorre à chamada regularidade de Hölder que descreve a regularidade das realizações dos processos estocásticos, fazendo uma comparação local com uma lei polinomial. Está assim estreitamente relacionada com a similaridade escalar no limite das escalas mais pequenas. O expoente desta lei polinomial, $h(t)$, é chamado de expoente (local) de Hölder e depende tipicamente tanto do tempo como da realização do processo estocástico.

Se $\{X(t), t \in \mathbb{R}\}$ designar um processo estocástico estacionário, com função de autocovariância $\mathbb{E}[X(t)X(t + \delta)] \sim (\sigma^2/2C)(1 - C|\delta|^{2h})$ (com $h > 0$) quando $\delta \rightarrow 0$, então para incrementos pequenos de x

$$\mathbb{E}[(X(t + \delta) - X(t))^2] \approx |\delta|^{2h}, \delta \rightarrow 0 \quad (2.13)$$

Esta relação dá informação sobre a regularidade de X já que a condição $h > 0$ garante a continuidade. No entanto a diferenciabilidade só existe caso $h > 1$. Ou seja, com $0 < h < 1$, as realizações de X são contínuas mas não diferenciáveis em qualquer ponto.

Os processos para os quais o expoente de Hölder $h(t)$ é o mesmo para qualquer t exibem uma regularidade constante para as suas realizações e são geralmente chamados de processos monofractais. O expoente de Hölder $h(t)$ proporciona a quantificação da regularidade local de uma função: uma realização de um processo estocástico exhibe menor variabilidade quando h aumenta de 0 até 1. Na Figura 2.4 é possível observar este mesmo fenómeno para o caso do movimento Browniano fraccionário (fBm), o qual é um exemplo de um processo monofractal com $h(t) = H$.

2.3.3 Processos multifractais

Quando o expoente de Hölder $h(t)$ é ele próprio uma função altamente irregular, possivelmente um processo estocástico, o processo X diz-se multifractal. A grande variabilidade de $h(t)$ torna extremamente complexa a tarefa de caracterizar as variações temporais desta função, sendo por isso necessário caracterizá-la estatisticamente. A caracterização estatística pode ser feita recorrendo ao chamado espectro multifractal da função $h(t)$. São usados dois tipos de análise espectral: o espectro multifractal de

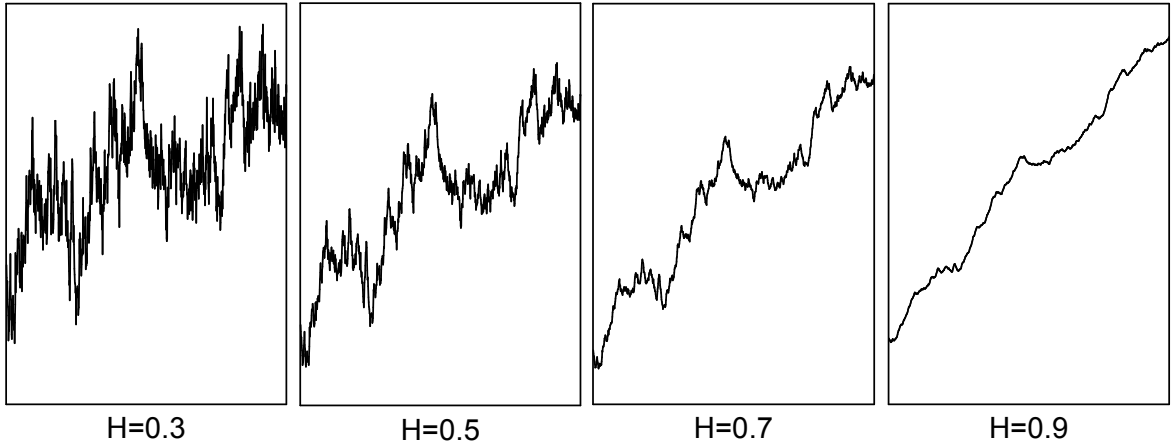


Figura 2.4: Realizações de processos fBm com diferentes valores de H .

Hausdorff e o espectro multifractal de Legendre. O espectro multifractal de Hausdorff é a função $D(a)$ que corresponde à dimensão de Hausdorff do conjunto de pontos $t \in \mathbb{R}$ onde $h(t) = a$, ou seja,

$$D(a) = \dim(t \in \mathbb{R} : h(t) = a) \quad (2.14)$$

onde $\dim(S)$ designa a dimensão de Hausdorff do conjunto S .

O espectro de Legendre recorre a uma das características dos processos multifractais segundo a qual o comportamento das chamadas funções de partição segue uma lei polinomial com expoente $\zeta(q)$:

$$\int |X(t + \tau) - X(t)|^q dt \approx |\tau|^{\zeta(q)}, \tau \rightarrow 0, q \in \mathbb{R} \quad (2.15)$$

O espectro multifractal de Legendre pode ser obtido efectuando a transformada de Legendre de $\zeta(\tau)$, ou seja,

$$D(h) = \min_q (qh - \zeta(q) + 1) \quad (2.16)$$

Exemplos de processos multifractais são as cascatas multiplicativas [FGW98, VAFC00, Rie02]. Alguns dos casos especiais das cascatas multiplicativas serão discutidos mais à frente neste capítulo na Secção 2.4.8.

2.3.4 Métodos de identificação e quantificação

Nesta secção serão apresentados alguns métodos usados na caracterização da similaridade escalar de uma série temporal.

2.3.4.1 Gráfico variância-escala

O gráfico variância-escala é um dos métodos mais antigos para a identificação da auto-similaridade e para a estimação do parâmetro de Hurst [CB97]. Tem como base o decaimento lento da variância ao longo das escalas de uma série com propriedades auto-similares.

O processo agregado de nível m (ou escala m) de uma série temporal discreta $Y(k), k \in \mathbb{N}$, é dado por

$$Y^{(m)}(i) = \frac{1}{m} \sum_{k=m(i-1)+1}^{mi} Y(k). \quad (2.17)$$

Se $Y(k)$ for um processo estacionário e exactamente (ou assintoticamente) auto-similar então, para valores de m onde a similaridade escalar se verifique, a variância do processo agregado é dada por

$$\text{Var}(Y^{(m)}) = \sigma^2 m^{2H-2}. \quad (2.18)$$

Aplicando o logaritmo a ambos os termos da expressão obtém-se,

$$\log(\text{Var}(Y^{(m)})) = \log(\sigma^2) + \beta \log m, \quad \beta = 2H - 2. \quad (2.19)$$

A expressão (2.19) mostra que, no caso em que $Y(k)$ seja auto-similar existe uma relação linear entre o logaritmo da variância de $Y^{(m)}$ e o logaritmo de m . O declive da recta obtida (β) permite inferir o valor estimado do parâmetro de Hurst, através de $H = 1 + \beta/2$.

2.3.4.2 Diagrama das energias em escala logarítmica

O diagrama das energias em escala logarítmica é um método de identificação de similaridade escalar baseado em onduletas e foi proposto em [VA99].

Um sinal $X(t)$ pode ser construído recorrendo a uma soma pesada de funções de escalamento $\phi_{j_0,k}$ e de onduletas $\psi_{j,k}(t)$

$$X(t) = \sum_k c_X(j_0, k) \phi_{j_0,k} + \sum_{j \leq j_0} \sum_k d_X(j, k) \psi_{j,k}(t) \quad (2.20)$$

onde $c_X(j_0, k)$ são os coeficientes de escalamento e os parâmetros $d_X(j, k)$ são os coeficientes de uma transformada discreta por onduletas. A família de funções base $\{\psi_{j,k}(t) = 2^{-j/2} \psi_0(2^{-j}t - k)\}, j = 1, \dots, J, k \in \mathbb{Z}$ é gerada a partir da onduleta mãe

ψ_0 . Note-se que se a onduleta mãe for uma função passa-banda com frequência central de f_0 , então as onduletas $\{\psi_{j,k}(t)\}$ têm a frequência central deslocada para $2^{-j}f_0$ e o instante temporal central deslocado de $2^j k$. Assim, a análise por onduletas permite investigar os comportamentos de uma função simultaneamente no tempo e na frequência (ou escala), constituindo-se como uma ferramenta natural para verificar comportamentos de invariância quanto à escala.

No caso de processos $X(t)$ estacionários de segunda ordem com memória longa verifica-se que [AFTV00]

$$\mathbb{E}[d_X(j, k)^2] \approx \int \Gamma_X(v) 2^j |\Psi_0(2^j v)|^2 dv \quad (2.21)$$

$$\approx 2^{j\alpha} c_f \int |v|^{-\alpha} |\Psi_0(v)|^2 dv \quad (2.22)$$

$$\approx 2^{j\alpha} c_f C, j \rightarrow \infty \quad (2.23)$$

onde $C = \int |v|^{-\alpha} |\Psi_0(v)|^2 dv$, c_f é uma constante não nula e o processo dos coeficientes $\{d_X(j, k), k \in \mathbb{Z}\}$ não apresenta memória longa desde que $N > \alpha/2$, em que N é o número de momentos nulos da onduleta mãe, isto é,

$$\int t^k \psi_0(t) dt \equiv 0, \quad k = 0, 1, \dots, N - 1. \quad (2.24)$$

Esta última propriedade justifica as vantagens de se fazer este tipo de análise estatística recorrendo a onduletas. A reescrita da equação (2.23) como

$$\log_2(\mathbb{E}[d_X(j, k)^2]) = j\alpha + \log_2(c_f C) \quad (2.25)$$

sugere uma regressão linear para a estimação de α e c_f . Veitch e Abry em [VA99] introduziram

$$\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} (d_X(j, k))^2 \quad (2.26)$$

e

$$y_j = \log_2(\mu_j) - g_j \quad (2.27)$$

onde μ_j é um estimador não enviesado da variância do processo $\{d_X(j, k), k \in \mathbb{Z}\}$, n_j representa o número de coeficientes na escala j disponíveis para análise e g_j é um parâmetro introduzido de modo a corrigir o enviesamento da estimação da variância de $d_X(j, k)$ resultante da não-linearidade introduzida pela função \log_2 . Neste caso,

$$\mathbb{E}[y_j] = j\alpha + \log_2(c_f C) \quad (2.28)$$

$$\text{Var}(y_j) = \xi(2, n_j/2) / \ln^2(2) \quad (2.29)$$

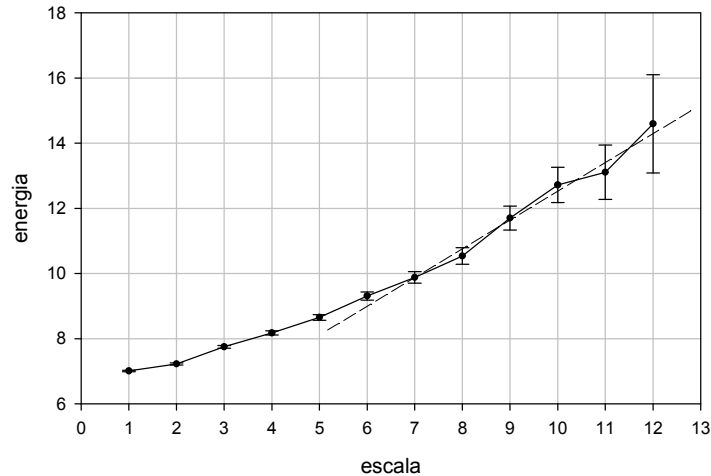


Figura 2.5: Exemplo de um diagrama de energias em escala logarítmica.

onde $\xi(z, \nu)$ é a função de Riemann Zeta generalizada. A representação gráfica (em escala logarítmica) de $\mathbb{E}[y_j]$ em função de j , em conjugação com os intervalos de confiança para $\mathbb{E}[y_j]$ obtidos a partir de $\text{Var}(y_j)$, permite obter uma estimativa do factor de escala α . Esta representação é chamada de diagrama de energias em escala logarítmica de segunda ordem. Diz-se de segunda ordem porque resulta da estimação da variância do processo $d_X(j, k)$. Fazendo uma representação gráfica dos valores médios y_j em função da escala j , em conjugação com os intervalos de confiança de y_j , é possível identificar as gamas de escalas onde existe similaridade. Existirá similaridade escalar em conjunto de escalas sucessivas tais que, dentro dos limites dos intervalos de confiança, os valores de y_j respectivos estejam alinhados segundo uma recta. A similaridade escalar é caracterizada pelo factor de escala α que corresponde ao declive da recta.

Na Figura 2.5 é possível observar um diagrama de energias em escala logarítmica. Neste diagrama é possível verificar que existe similaridade escalar na gama de escalas superiores (de 7 a 12) pois nesta gama existe (dentro dos limites dos intervalos de confiança a 95%) um alinhamento dos pontos. Deste modo é possível concluir que a captura analisada possui características de dependência longa.

2.3.5 Diagramas lineares multiescalares

Abry *et al.* propuseram uma generalização dos diagrama das energias em escala logarítmica para o estudo das estatísticas de ordem superior à segunda [AFTV00]. Para a classe dos processos multifractais e para j pequenos o estimador do momento de

ordem q do processo $\{d_X(j, k), k \in \mathbb{Z}\}$, $\mu_j^{(q)}$, comporta-se de acordo com

$$\mu_j^{(q)} = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j, k)|^q \approx 2^{j(\zeta(q)+q/2)} = 2^{j\alpha_q}, q \in \mathbb{R}. \quad (2.30)$$

onde α_q é o factor de escala de ordem q .

Para um processo auto-similar o factor de escala varia linearmente com a ordem q , isto é, $\alpha_q = q(H+1/2)$. A falta da linearidade de α_q (em função de q) permite identificar a presença de fenómenos de similaridade multifractal não trivial num determinado processo. É assim necessário estimar α_q para uma variedade de valores de q , o que pode ser feito a partir dos diagramas de energias em escala logarítmica de ordem q . Estas estimativas são efectuadas através de

$$y_j^{(q)} \approx \log_2(\mu_j^{(q)}), \quad q \in \mathbb{R} \quad (2.31)$$

$$\mathbb{E}[y_j^{(q)}] \approx j\alpha_q, \quad q \in \mathbb{R} \quad (2.32)$$

À semelhança do que é feito para as energias de segunda ordem, α_q é dado pelo declive que caracteriza o alinhamento das energias $y_j^{(q)}$ em função da escala temporal j .

Os diagramas lineares multiescalares consistem na representação gráfica de $h(q) = \alpha_q/q - 1/2$ em função de q (ver Figura 2.6). Um alinhamento horizontal dentro dos limites de confiança de h revela um comportamento linear de $\alpha_q = q(H + 1/2)$, o que significa estarmos na presença de auto-similaridade ou de similaridade multifractal trivial. Por outro lado a falta de alinhamento horizontal de h revela um comportamento não linear de α_q mostrando estarmos na presença de similaridade multifractal (não trivial), como é o caso representado na Figura 2.6.

2.4 Modelos de tráfego

Ao longo dos anos foram vários os modelos de tráfego sugeridos para diferentes aplicações: processos de renascimento, Markovianos, fluídos, auto-regressivos, TES, autosimilares e multifractais [JMW96, ABF⁺02]. Nesta secção será feita uma breve descrição de cada um destes tipos de modelos.

2.4.1 Processos de renascimento

Os modelos de tráfego com base em processos de renascimento são os mais antigos. Num processo de renascimento os intervalos entre chegadas são independentes e identicamente distribuídos segundo uma distribuição geral. Estes modelos, apesar da sua

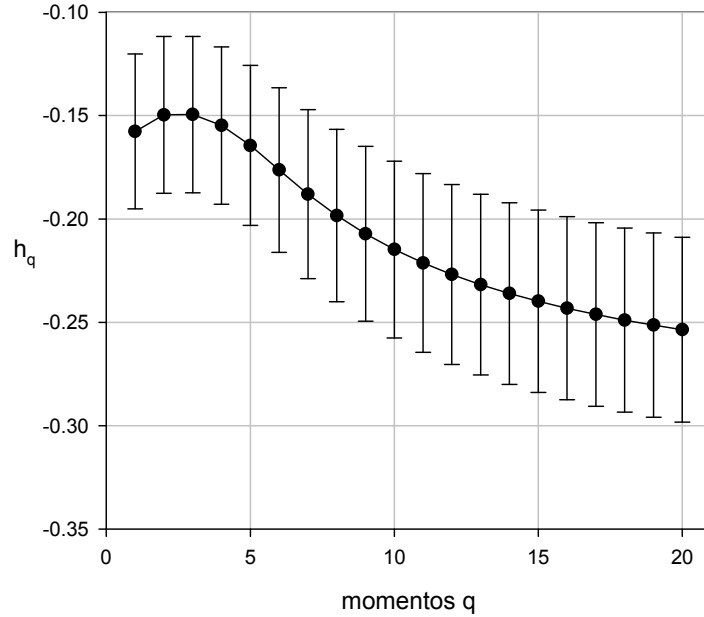


Figura 2.6: Exemplo de um diagrama linear multiescalar, que revela existência de similaridade multifractal (não trivial).

simplicidade analítica, possuem como grande desvantagem o facto de não conseguirem capturar a correlação entre chegadas. Os principais processos de renovamento usados como modelos de tráfego são os processos de Poisson, os processos de Bernoulli e os processos de renovamento do tipo fase.

Os processos de Poisson são os modelos de tráfego mais antigos, e surgiram no início da rede telefónica propostos por A. K. Erlang para modelar chegadas de chamadas telefónicas. Um processo de Poisson com taxa $\lambda > 0$ pode ser definido como um processo de contagem $\{N(t), t \geq 0\}$ em que $N(0) = 0$, o número de chegadas em intervalos de tempo disjuntos são independentes e o número de chegadas em qualquer intervalo de comprimento t segue uma distribuição de Poisson com média λ , isto é

$$P\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad n = 0, 1, 2, \dots \quad (2.33)$$

Num processo de Poisson os intervalos entre chegadas (A_n) são independentes e identicamente distribuídos segundo uma distribuição exponencial, isto é

$$P\{A_n \leq t\} = 1 - e^{-\lambda t}. \quad (2.34)$$

Esta classe de processos possui duas características importantes: (i) a sobreposição de processos de Poisson independentes resulta também num processo de Poisson cuja taxa de chegadas é a soma das taxas dos processos sobrepostos e (ii) sendo um processo de renovamento é um processo sem memória. Estas características simplificam a análise

dos sistemas de filas de espera cujas chegadas sejam de Poisson, de que são exemplos os sistemas M/M/m, M/M/m/m e M/G/1.

Os processos de Bernoulli são a versão discreta dos processos de Poisson: em cada intervalo de tempo ocorre uma chegada de acordo com uma distribuição de Bernoulli (com probabilidade p) o que resulta em intervalos entre chegadas geometricamente distribuídas, ou seja,

$$P\{A_n = j\} = p(1-p)^j, j \in \mathbb{N}_0^+. \quad (2.35)$$

O número de chegadas num determinado intervalo k segue uma distribuição binomial:

$$P\{N_k = n\} = C_n^k p^n (1-p)^{k-n}, \quad 0 \leq n \leq k. \quad (2.36)$$

Num processo de renovamento do tipo fase o intervalo entre chegadas A_n é definido como sendo o tempo que um processo de Markov em tempo contínuo (com um número de estados finito) demora até chegar a um estado absorvente. Após a absorção o processo de Markov é reinicializado segundo uma determinada distribuição e repetido o processo de modo a obter A_{n+1} , o qual é independente de A_n . Assim, o processo dos intervalos entre chegadas terá uma distribuição definida por uma soma pesada de exponenciais. Em relação aos processos de Poisson, os processos do tipo fase têm a vantagem de conseguir aproximar qualquer distribuição dos intervalos de tempo entre chegadas.

2.4.2 Modelos Markovianos

Uma outra classe de modelos são os modelos Markovianos. Ao contrário dos processos de renovamento estes processos conseguem introduzir dependência entre as chegadas. Em consequência deste facto, estes modelos conseguem capturar as características de chegadas em rajada do tráfego, pois possuem autocorrelação não nula.

O modelos Markovianos são caracterizados por um processo de Markov com vários estados, também designado de cadeia de Markov. A transição entre estados é feita de acordo com uma probabilidade que depende apenas do estado de origem e do estado de destino. Estes processos podem ser definidos em tempo discreto ou contínuo. No caso dos modelos Markovianos em tempo discreto, o tempo de permanência em cada estado segue uma distribuição geométrica com um parâmetro que depende apenas do estado. Num modelo de Markov simples, cada transição entre estados sinaliza uma chegada, resultando assim em chegadas com intervalos geometricamente distribuídos. No caso de um processo de Markov em tempo contínuo o tempo de permanência num

estado segue uma distribuição exponencial. Para um processo simples de Markov em tempo contínuo, onde cada transição do processo sinaliza uma chegada, o intervalo entre chegadas é exponencial com parâmetros que dependem apenas do estado de onde ocorreu a transição.

Um processo Markoviano de renovamento é uma generalização de um processo de Markov, onde o tempo de permanência em cada estado segue uma distribuição genérica cujos parâmetros são dependentes do estado da cadeia de Markov subjacente.

Uma das subclasses dos modelos Markovianos mais versáteis, e por isso mais usadas e estudadas, são os processos de chegadas Markovianos (MAPs) os quais gozam de grande facilidade de tratamento analítico [LMHN90]. Num MAP os intervalos entre chegadas são um processo de renovamento onde as chegadas ocorrem quando o processo atinge um estado absorvente. No entanto ao contrário dos processos de renovamento do tipo fase a distribuição usada para reinicializar o processo passa a ser dependente do estado de onde o processo transitou para o estado absorvente, deixando de ser uma distribuição única. Este modelo possui a propriedade de que a sobreposição de MAPs resulta ainda num MAP. Esta propriedade facilita o estudo da multiplexagem estatística de várias fontes MAP numa fila de espera.

Uma generalização do MAP simples é o processo de Markov com chegadas em rajada (BMAP), o qual foi apresentado em [Luc93]. O BMAP é um processo idêntico ao MAP mas onde as chegadas podem ocorrer em rajada, isto é, podem existir várias chegadas no mesmo instante de tempo. Também os BMAPs são fechados quanto à sobreposição de processos, ou seja, o processo resultante da sobreposição de dois BMAPs é também um BMAP. No capítulo 4 será apresentada a formalização matemática desta classe de modelos.

2.4.3 Processos modulados à Markov

Os processos modulados à Markov são talvez a classe de modelos de tráfego mais estudada. Nestes modelos existe um processo de Markov subjacente que evolui no tempo e cujo estado actual modula as características estatísticas das chegadas. Enquanto a cadeia de Markov subjacente estiver num determinado estado as chegadas ocorrem segundo determinada regra; quando o processo Markoviano comuta para um novo estado as chegadas passam a ocorrer de acordo com uma regra definida pelo novo estado.

Talvez uma das mais importantes classes de modelos de tráfego seja a dos processos de Poisson modulados à Markov, também chamados de MMPP. Parte dos modelos

propostos neste Tese pertencem a esta classe. Num MMPP, as chegadas ocorrem, em cada estado, segundo um processo de Poisson com uma taxa que pode ser diferente para cada estado da cadeia de Markov subjacente. Um MMPP é completamente descrito por duas matrizes, uma que define as taxas de chegadas nos estados e outra que, no caso de um MMPP em tempo discreto, é uma matriz de probabilidades de transição e, no caso de um MMPP em tempo contínuo, é uma matriz de taxas de transição entre estados. No capítulo 3 será feita uma apresentação dos modelos baseados em MMPPs existentes na literatura bem como a formalização matemática desta classe de modelos.

2.4.4 Modelos de fluídos

Os modelos de fluídos não tentam modelar as chegadas individuais mas descrevem o tráfego como um fluído caracterizado apenas por uma taxa de chegadas. O modelo de fluídos caracteriza-se pela sua simplicidade, o que é uma vantagem no que respeita à inferência e análise matemática do mesmo, mas torna-se uma desvantagem quando se pretende modelar características do tráfego que vão além do volume de tráfego e duração dos períodos de actividade.

2.4.5 Modelos auto-regressivos

Os modelos auto-regressivos definem um novo elemento de uma sequência aleatória em função dos valores anteriores dessa mesma sequência que estejam dentro de uma determinada janela temporal. Podem identificar-se três classes de modelos auto-regressivos: lineares auto-regressivos (AR), auto-regressivos de média móvel (ARMA) e auto-regressivos de média móvel integrados (ARIMA).

Nos modelos auto-regressivos lineares de ordem p [BJ76], $AR(p)$, a sequência de tráfego $\{X_n, n > 0\}$ é construída iterativamente através de

$$X_n = a_0 + \sum_{r=1}^p a_r X_{n-r} + \epsilon_n, \quad n > 0 \quad (2.37)$$

onde a_r , $0 \leq r \leq p$, são constantes reais e ϵ_n representa ruído branco. Usando o operador B^r definido como $X_{n-r} = B^r X_n$ o processo $AR(p)$ pode ser descrito como

$$a(B)X_n = \epsilon_n \quad (2.38)$$

onde $a(B) = 1 - \sum_{r=1}^p a_r B^r$. A sequência $\{X_n\}$ será estacionária se o polinómio $a(B)$ tiver os seus zeros dentro do círculo unitário. Um modelo simples do tipo $AR(2)$ tem

sido usado para modelar vídeo codificado com ritmo de bits variável (VBR) [HTL92]. Os modelos auto-regressivos são tipicamente usados para ajustar a função de autocovariância não sendo em regra capazes de ajustar a distribuição marginal empírica.

Um modelo auto-regressivo de média móvel de ordem (p, q) , $\text{ARMA}(p, q)$, é construído adicionando a um modelo AR uma componente de médias móveis [Ada97]. Um modelo $\text{ARMA}(p, q)$ tem a seguinte forma

$$X_n = a_0 + \sum_{r=1}^p a_r X_{n-r} + \sum_{r=0}^q b_r \epsilon_{n-r}, \quad n > 0 \quad (2.39)$$

onde a_r , $0 \leq r \leq p$, e b_r , $0 \leq r \leq q$, são constantes reais do modelo e ϵ_r representa ruído branco. A expressão (2.39) pode ser igualmente representada como

$$a(B)X_n = b(B)\epsilon_n \quad (2.40)$$

onde $a(B) = 1 - \sum_{r=1}^p a_r B^r$ e $b(B) = 1 - \sum_{r=1}^q b_r B^r$. No modelo ARMA, a componente adicional permite um melhor ajuste da função de autocorrelação. A estimação dos parâmetros do modelo ARMA é mais complexa do que no modelo AR, pois as componentes b_r , $r = 1, \dots, q$ têm de ser estimadas recorrendo a um sistema de equações não-linear.

O modelo auto-regressivo de média móvel integrado de ordem (p, d, q) , $\text{ARIMA}(p, d, q)$, é obtido a partir do modelo $\text{ARMA}(p, q)$ permitindo que o polinómio $a(B)$ tenha d raízes iguais à unidade e as restantes raízes fora do círculo unitário. O modelo $\text{ARIMA}(p, d, q)$ tem a seguinte forma

$$a(B)\nabla^d X_n = b(B)\epsilon_n \quad (2.41)$$

onde ∇ é um operador diferença definido por $\nabla X_n = X_n - X_{n-1}$ e $\nabla^d = (1 - B)^d$. O modelo ARIMA é utilizado para modelar tráfego não estacionário.

2.4.6 Modelos TES

Os modelos TES, cujo designação advém das três operações efectuadas aos dados (transformação, expansão e amostragem), permitem o ajuste simultâneo da distribuição empírica e da função de autocovariância até um deslocamento temporal razoável [MS92, JM92a, JM92b].

Existem duas variantes dos processos TES, TES^+ e TES^- , os quais possuem respectivamente autocorrelação positiva e negativa. Os modelos TES resultam da conjugação

de dois processos estocásticos, uma sequência de *background* e uma sequência de *foreground*. A sequência de *background* do TES⁺ tem a forma

$$U_n^+ = \begin{cases} U_0, & n = 0 \\ < U_{n-1}^+ + V_n >, & n > 0 \end{cases} \quad (2.42)$$

e a sequência de *background* do TES⁻ tem a forma

$$U_n^- = \begin{cases} U_n^+, & n \text{ par} \\ 1 - U_n^+, & n \text{ ímpar} \end{cases} \quad (2.43)$$

onde U_0 é uma variável uniformemente distribuída entre 0 e 1, $\{V_n, n = 1, \dots, \infty\}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas, independente de U_0 e chamada de sequência de inovação, e a operação $< x >$ representa $x - \max\{n \leq x, n \in \mathbb{N}\}$. As sequências de *foreground* do TES⁺ e do TES⁻ são respectivamente

$$X_n^+ = D(U_n^+) \quad (2.44)$$

$$X_n^- = D(U_n^-) \quad (2.45)$$

onde D representa uma função de distorção. O processo de inferência do modelo TES recorre a uma distorção D de dois níveis definida por

$$D(x) = F^{-1}(S_\xi(x)), \quad x \in [0, 1] \quad (2.46)$$

onde F^{-1} representa a função de distribuição empírica inversa e S_ξ é uma operação de "alisamento" dada por

$$S_\xi(y) = \begin{cases} y/\xi, & 0 \leq y < \xi \\ (1 - y)/(1 - \xi), & \xi \leq y < 1 \end{cases} \quad (2.47)$$

que preserva a uniformidade. Para $0 < \xi < 1$ o efeito da função S_ξ é fazer com que as realizações do processo TES de *background* tenham uma aparência mais contínua.

Uma vantagem importante da metodologia TES, que também é perseguida nesta Tese, é o desacoplamento entre o ajuste da distribuição empírica e da autocorrelação empírica. O ajuste da distribuição empírica está garantido à partida. O ajuste da autocorrelação empírica é baseado numa procura heurística dos pares (ξ, f_V) onde f_V representa a função densidade da sequência de inovação.

2.4.7 Modelos auto-similares

Englobam-se nesta classe os processos que possuem intrinsecamente características auto-similares, de que são exemplos o movimento browniano fraccionário (fBm) [MN68] e os processos ARIMA fraccionários. Estes modelos conseguem modelar a auto-similaridade e a média do tráfego, mas são incapazes de capturar outras características importantes do tráfego, nomeadamente a distribuição e a função de autocovariância.

O movimento Browniano simples é um processo estocástico $\{B(t), t > 0\}$ onde os incrementos $B(t + \delta) - B(t)$ têm uma distribuição normal com média zero e variância $\sigma^2\delta$ e os incrementos em intervalos de tempo disjuntos são independentes. O movimento Browniano fraccionário é um processo auto-similar com parâmetro de Hurst $H \in [0.5, 1[$ e difere do movimento Browniano simples por ter incrementos com variância $\sigma^2\delta^{2H}$. O fBm, $\{fB(t), t > 0\}$, pode ser deduzido a partir do movimento Browniano simples através da seguinte expressão

$$fB(t) = \int_0^t (t-u)^{H-0.5} dB(u). \quad (2.48)$$

A sequência de incrementos é chamada de ruído Gaussiano fraccionário.

Os processos ARIMA fraccionários, F-ARIMA(p, d, q), com $0 < d < 1/2$ são uma extensão dos modelos ARIMA(p, d, q) e são definidos por

$$a(B)\nabla^d X_n = b(B)\epsilon_n \quad (2.49)$$

onde d toma valores entre 0 e 1/2, B é um operador definido por $BX_n = X_{n-1}$, $a(B) = 1 - \sum_{r=1}^p a_r B^r$ e $b(B) = 1 - \sum_{r=1}^q b_r B^r$. O que diferencia os processos ARIMA e F-ARIMA, é que num processo F-ARIMA o operador $\nabla^d = (1 - B)^d$ é descrito por uma expansão binomial dada por

$$(1 - B)^d = \sum_{k=0}^{\infty} C_k^d (-1)^k B^k \quad (2.50)$$

onde $C_k^d = d!/k!(d-k)!$. O parâmetro de Hurst de um modelo F-ARIMA é dado por $H = d + 1/2$.

2.4.8 Modelos multifractais

Os primeiros processos multifractais conhecidos foram as chamadas cascatas multiplicativas [FGW98, Rie02] e continuam a ser a classe de modelos multifractais mais

divulgada. A cascata mais simples é a chamada de cascata binomial que surgiu no contexto da modelação de processos financeiros [FCM97, Rie02]. Esta pode ser descrita por uma estrutura binária em árvore. Uma cascata binomial pode ser determinística, caso não exista nenhuma componente aleatória no seu processo de construção, ou aleatória. O modelo mais genérico de cascatas aleatórias foi proposto por Mandelbrot [Man74] como um modelo para a turbulência. Uma cascata aleatória é caracterizada por uma massa inicial, uniformemente distribuída ao longo de um intervalo de tempo, o qual é subdividido em dois (ou mais) subintervalos idênticos e onde a massa é aleatoriamente atribuída a cada um desses subintervalos, de acordo com uma variável aleatória W chamada de gerador. No caso de um processo de tráfego a massa pode ser interpretada como sendo o número de chegadas (pacotes ou bytes) num determinado intervalo de tempo. Seja $W_{i,j}, i = 1, \dots, N_j, j = S, \dots, 1$, uma variável aleatória independente, com a mesma distribuição de W , que redistribui a massa da escala temporal j para o intervalo de tempo i (o qual pertence à escala temporal seguinte $j - 1$). A construção do processo de tráfego começa na escala de tempo superior S com uma massa inicial M distribuída uniformemente ao longo de um intervalo de tempo unitário. Esta discussão será restrita ao caso em que cada intervalo (pai) é subdividido em dois subintervalos (filhos). Na primeira iteração, é produzida uma escala mais fina ($S - 1$) dividindo o intervalo em dois subintervalos de comprimento $1/2$, e é atribuída uma massa $MW_{1,S-1}$ ao subintervalo da esquerda e uma massa $MW_{2,S-1}$ ao subintervalo da direita. Na segunda iteração é produzida a escala temporal $S - 2$ a partir da escala temporal $S - 1$, em que cada um dos intervalos anteriores gera novos subintervalos, um à esquerda e outro à direita, dando origem a quatro subintervalos de comprimento $1/4$ com massas $MW_{1,S-1}W_{1,S-2}$, $MW_{1,S-1}W_{2,S-2}$, $MW_{2,S-1}W_{3,S-2}$ e $MW_{2,S-1}W_{4,S-2}$, respectivamente. Repetindo este processo iterativamente, a massa do intervalo pai i da escala temporal j é redistribuída pelos subintervalos $2i - 1$ e $2i$ da escala $j - 1$ com pesos $W_{2i-1,j-1}$ e $W_{2i,j-1}$, respectivamente. De referir que o modo como a massa é redistribuída é independente da quantidade de massa, pois $W_{i,j}$ é independente da massa. A construção do processo está ilustrado na Figura 2.7. Combinando este processo com o movimento Browniano fraccionário é possível definir uma nova classe de modelos designada por movimento Browniano fraccionário em tempo multifractal [VR97]. Riedi *et al.* [RCRB99] propuseram um modelo multifractal com base em onduletas (MWM), onde utilizam as onduletas de Haar e uma estrutura multiplicativa aplicada aos coeficientes da onduleta. Mais recentemente foi apresentado um modelo que combina uma

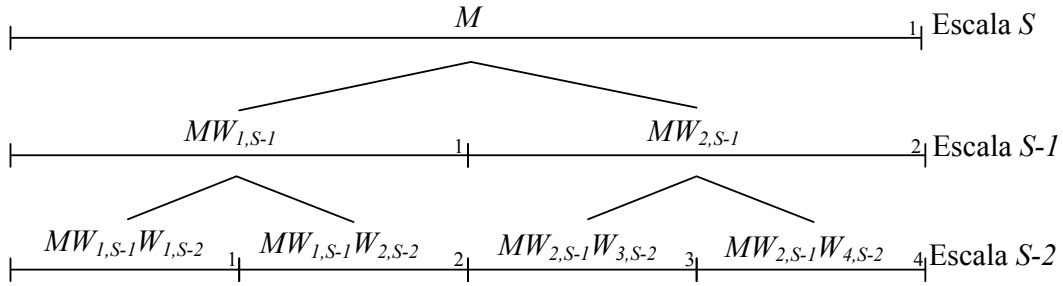


Figura 2.7: Representação da evolução de uma cascata aleatória.

cascata multiplicativa com um processo log-normal independente [DMM02]. Todos estes modelos conseguem capturar em parte ou totalmente as características multifractais do tráfego.

Feldmann *et al.* [FGW98] propuseram o uso de cascatas conservativas como um modelo para tráfego IP. Nas cascatas conservativas, o gerador W toma valores entre 0 e 1, tem média $1/2$ e é simétrico em relação à média. Além disso, a massa é redistribuída de tal modo que o total de massa atribuído aos subintervalos filhos da esquerda e da direita é igual à massa do intervalo pai. Assim, a massa é preservada durante o processo de divisão: se a massa do i -ésimo intervalo pai for M , a massa do intervalo filho da esquerda será $MW_{2i-1,j-1}$ e a massa do intervalo filho da direita será $M(1 - W_{2i-1,j-1})$. Como resultado, a massa em todas as escalas será exactamente M , se a massa inicial for M . Esta é a chamada propriedade de preservação de massa. De referir, que é possível o uso de diferentes geradores para cada escala temporal, levando assim ao aparecimento de características multiescalares.

Foi ainda proposto um modelo baseado em cascatas duplas independentes, em que uma cascata modela as chegadas dos pacotes e a outra modela os tamanhos dos pacotes [GR01, SNV03b].

2.5 Conclusão

Neste capítulo foi feita uma descrição do estado da arte da modelação de tráfego em redes de comunicações.

Foram apresentados os conceitos e definições matemáticas das características de similaridade escalar do tráfego: auto-similaridade, dependência longa e multifractalidade. Foram também descritos os modelos de tráfego mais importantes propostos na literatura.

A modelação das características de similaridade escalar não está restrita aos pro-

cessos que as possuem de forma intrínseca, como por exemplo o movimento Browniano fraccionário para o caso da auto-similaridade. Modelos mais clássicos, como por exemplo os Markovianos, podem fornecer boas aproximações e apresentam como vantagem a existência de um leque de resultados teóricos que facilitam a análise do comportamento em fila de espera e o cálculo da largura de banda efectiva. Além disso, estes modelos permitem incorporar outras características estatísticas relevantes do tráfego que, em geral, não são inteiramente capturadas pelos processos intrinsicamente invariantes quanto à escala (por exemplo as estatísticas de primeira e segunda ordem).

Neste capítulo foram ainda apresentados alguns métodos de identificação e quantificação dos comportamentos de similaridade escalar, que serão utilizados ao longo da Tese, quer na caracterização das capturas usadas, quer na aferição da capacidade dos modelos propostos para capturarem e replicarem estas características.

A importância de modelar as características de similaridade escalar do tráfego a par com a escassez de modelos suficientemente flexíveis que permitam capturar, não apenas estas características, mas também outras com impacto significativo no desempenho da rede motivou o desenvolvimento dos modelos de tráfego apresentados nos capítulos seguintes desta Tese.

Capítulo 3

Modelos Markovianos para o processo de chegadas de pacotes

3.1 Introdução

Neste capítulo será apresentado um novo modelo de tráfego e respectivo processo de inferência de parâmetros. O modelo proposto é um processo de Poisson modulado à Markov em tempo discreto (dMMPP) com um procedimento de inferência de parâmetros que se pretende versátil no que respeita ao tipo de tráfego que consegue modelar.

Um modelo de tráfego só está completo quando existe um procedimento de inferência de parâmetros associado à estrutura matemática do modelo. No caso dos MMPPs a complexidade do procedimento de inferência de parâmetros está ligada ao número de estados do MMPP ao qual se pretende ajustar o tráfego. A maior parte dos procedimentos de inferência de parâmetros de MMPPs disponíveis na literatura aplica-se apenas a modelos que tenham por base uma cadeia de Markov com apenas 2 estados (2-MMPP) [HL86, GR94, KS95, MH87, NP98]. Apesar dos 2-MMPPs conseguirem capturar as características fundamentais de algumas fontes de tráfego (por exemplo, tráfego de voz), o número de estados em geral não é suficiente para permitir um bom ajuste da distribuição marginal quando o tráfego exhibe uma variabilidade ao longo de uma extensa gama de valores de ritmos de chegada. Skelly *et al.* [SSD93] propuseram um método para a estimação dos parâmetros de um MMPP genérico que apenas ajusta as estatísticas de primeira ordem. Os ritmos de chegada de Poisson são inferidos a partir da função marginal empírica e as taxas de transição entre estados são inferidas fazendo a contagem das transições ao longo da sequência empírica do número de chegadas. Uma limitação deste método é a necessidade de pré-definir o

número de estados do MMPP, pois o número de estados está relacionado com a resolução usada na construção da função de distribuição empírica. No final obtêm-se ritmos de chegada de Poisson igualmente espaçados, resultando num procedimento que não é adaptável às características particulares do tráfego que se pretende modelar. O trabalho de Li e Hwang [LH97], foi um passo em frente no ajuste dos parâmetros de um MMPP. Apesar do procedimento se aplicar a um caso particular de MMPPs, os CMPPs (processo de Poisson modulado por uma cadeia de Markov circulante), permite o ajuste da distribuição marginal e da função de autocovariância. Num CMPP as probabilidades em regime estacionário são idênticas para todos os estados do modelo. A estrutura do CMPP permite ultrapassar o problema usualmente chamado de inversão dos valores próprios, o qual está associado à necessidade de inverter a exponencial da matriz que define as taxas de transição dos estados da cadeia de Markov de modo a obter os elementos desta matriz a partir da função de autocovariância. Este procedimento permite capturar as componentes pseudo-periódicas presentes nos dados, já que a matriz das taxas de transição entre estados de um CMPP pode ter valores próprios complexos. No entanto, este procedimento possui uma reduzida flexibilidade no ajuste da distribuição marginal dos ritmos de chegada, porque os estados do CMPP são equiprováveis. Em particular, com este procedimento é difícil capturar rajadas de chegadas que ocorram com uma baixa probabilidade. Estas chegadas em rajada podem afectar significativamente o comportamento do tráfego na passagem por uma fila de espera. No procedimento proposto em [LH97] a precisão do ajuste da distribuição marginal apenas pode ser melhorada à custa do aumento do número de estados do CMPP. Andersen e Nielsen [AN98] usam a sobreposição de 2-MMPPs de modo a modelar as diferentes escalas temporais da autocovariância. Cada uma dessas escalas temporais é ajustada por uma função exponencial, usando um procedimento semelhante ao usado em [FW97]. O procedimento de Andersen e Nielsen tem duas grandes desvantagens; (i) as escalas temporais são definidas à partida e (ii) o ajuste das estatísticas de primeira ordem é pobre, já que apenas a média é ajustada.

O modelo e respectivo procedimento de inferência de parâmetros propostos neste capítulo da Tese foram desenvolvidos com o objectivo de conseguir modelar de forma precisa as estatísticas de primeira e de segunda ordem do tráfego, e consequentemente a memória longa do processo (caso exista). O modelo proposto é obtido por sobreposição de L 2-dMMPPs e de um M -dMMPP sem memória, sendo por isso designado de $M2^L$ -dMMPP. Os 2-dMMPP modelam a função de autocovariância e o M -dMMPP

modela a função massa de probabilidade, tomando em conta as restrições impostas pela modelação da autocovariância. Uma característica importante deste modelo é o facto de o número de estados do dMMPP não ser fixado à priori, permitindo uma adaptação ao tipo de tráfego que se está a caracterizar.

Este capítulo está organizado do seguinte modo. Na secção 3.2 faz-se uma introdução aos processos de Poisson modulados à Markov. Na secção 3.3 é apresentado em pormenor o modelo M^L -dMMPP e o respectivo procedimento de inferência de parâmetros. Na secção 3.4 são apresentados e discutidos os resultados da aplicação do modelo proposto a tráfego real. Finalmente na secção 3.5 são apresentadas algumas conclusões.

3.2 Processos de Poisson modulados à Markov em tempo discreto (dMMPP)

Um processo de Poisson modulado à Markov em tempo discreto (dMMPP) pode ser visto como um passeio aleatório Markoviano cujos incrementos em cada instante têm uma distribuição de Poisson com uma média que depende do estado da cadeia de Markov moduladora. Mais precisamente, a cadeia de Markov (homogénea) $(X, J) = \{(X_k, J_k), k = 0, 1, \dots\}$ com um espaço de estados $\mathbb{N}_0 \times S$ é um dMPPP se e só se para $k = 0, 1, \dots$,

$$P(X_{k+1} = m, J_{k+1} = j | X_k = n, J_k = i) = \begin{cases} 0, & m < n \\ p_{ij} e^{-\lambda_i} \frac{\lambda_i^{m-n}}{(m-n)!}, & m \geq n \end{cases} \quad (3.1)$$

para todo $m, n \in \mathbb{N}_0$ e $i, j \in S$, em que $\lambda_i, i \in S$, é constante real e não-negativo, e $\mathbf{P} = (p_{ij})$ é uma matriz estocástica irredutível. De referir que a distribuição de $X_{k+1} - X_k$ dado que $J_k = j$ é de Poisson com média λ_j , sendo que λ_j representa o incremento médio do processo X quando a cadeia de Markov moduladora está no estado j . O dMMPP é um caso particular de um dBMAP (processo em tempo discreto de chegadas Markovianas em rajada). Num dBMAP a distribuição dos incrementos do processo em cada instante de tempo pode depender do estado visitado no instante de tempo anterior, para além de depender do estado actual. O dBMAP será descrito em mais pormenor no capítulo 4.

Sempre que (3.1) se verificar, diz-se que (X, J) é um dMPPP com um conjunto de

estados moduladores S e com parâmetros (matrizes) \mathbf{P} e $\mathbf{\Lambda}$, e pode-se escrever

$$(X, J) \sim \text{dMMPP}_S(\mathbf{P}, \mathbf{\Lambda}) \quad (3.2)$$

onde $\mathbf{\Lambda} = (\lambda_{ij}) = (\lambda_i \delta_{ij})$, $\delta_{ij} = 1$ quando $i = j$ e $\delta_{ij} = 0$ quando $i \neq j$. A matriz \mathbf{P} é a matriz das probabilidades de transição entre os estados da cadeia de Markov moduladora J , e $\mathbf{\Lambda}$ é a matriz de taxas de chegada de Poisson. Se S for constituído por r estados, diz-se que (X, J) é um dMMPP de ordem r (r -dMMPP). Quando, em particular, $S = \{1, 2, \dots, r\}$ para algum $r \in \mathbb{N}$, então

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1r} \\ p_{21} & p_{22} & \dots & p_{2r} \\ \dots & \dots & \dots & \dots \\ p_{r1} & p_{r2} & \dots & p_{rr} \end{bmatrix} \text{ e } \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_r \end{bmatrix} \quad (3.3)$$

e escreve-se simplesmente que $(X, J) \sim \text{dMMPP}_r(\mathbf{P}, \mathbf{\Lambda})$.

A distribuição estacionária de J é designada por $\pi = [\pi_1, \pi_2, \dots, \pi_r]$, e verificam-se a seguintes condições:

$$\mathbf{e}(\mathbf{P} - \mathbf{I}) = 0 \quad (3.4)$$

$$\sum_i \pi_i = 1 \quad (3.5)$$

onde \mathbf{I} designa a matriz identidade de dimensão $r \times r$, e \mathbf{e} é um vector coluna de comprimento r com todos os elementos iguais a 1.

Consideremos a sobreposição de N MMPPs, $(X^{(i)}, J^{(i)}) \sim \text{dMMPP}_{r_i}(\mathbf{P}^{(i)}, \mathbf{\Lambda}^{(i)})$, $i = 1, 2, \dots, N$. O resultado desta sobreposição é ainda um processo dMMPP, isto é,

$$(X, J) = \left(\sum_{i=1}^N X^{(i)}, (J^{(1)}, J^{(2)}, \dots, J^{(N)}) \right) \sim \text{dMMPP}_S(\mathbf{P}, \mathbf{\Lambda}) \quad (3.6)$$

onde

$$S = \{1, \dots, r_1\} \times \{1, \dots, r_2\} \times \dots \times \{1, \dots, r_N\} \quad (3.7)$$

$$\mathbf{P} = \mathbf{P}^{(1)} \otimes \mathbf{P}^{(2)} \otimes \dots \otimes \mathbf{P}^{(N)} \quad (3.8)$$

$$\mathbf{\Lambda} = \mathbf{\Lambda}^{(1)} \oplus \mathbf{\Lambda}^{(2)} \oplus \dots \oplus \mathbf{\Lambda}^{(N)} \quad (3.9)$$

em que \oplus e \otimes designam respectivamente a soma e o produto de Kronecker.

Na Figura 3.1 está ilustrado o exemplo da sobreposição de dois 2-dMMPPs. Os 2-dMMPPs usados no exemplo têm as matrizes de taxas de Poisson

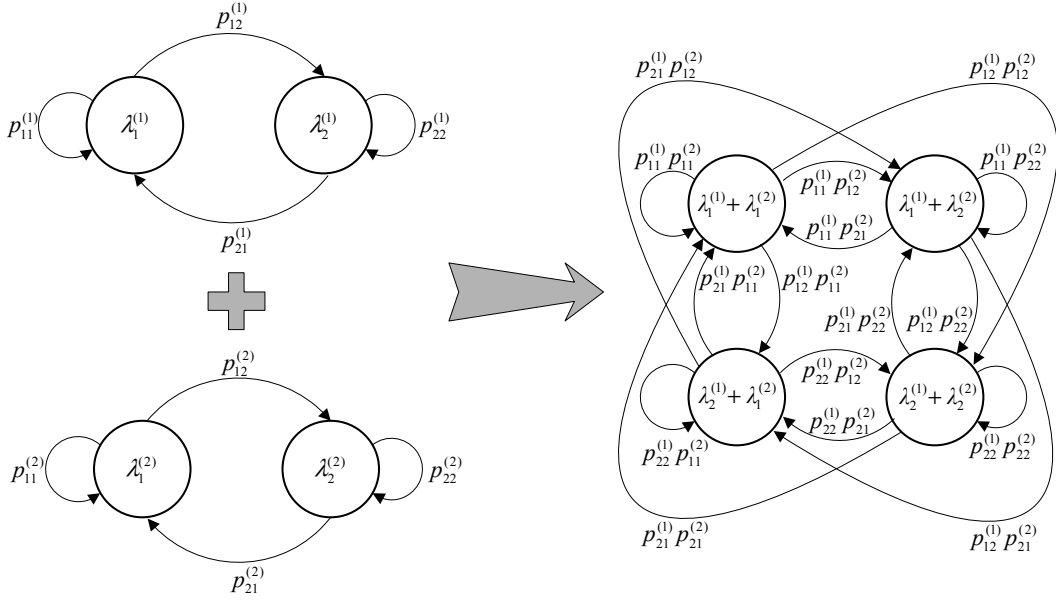


Figura 3.1: Exemplo da sobreposição de 2 2-dMMPPs.

$$\mathbf{\Lambda}^{(1)} = \begin{bmatrix} \lambda_1^{(1)} & 0 \\ 0 & \lambda_2^{(1)} \end{bmatrix} \quad \text{e} \quad \mathbf{\Lambda}^{(2)} = \begin{bmatrix} \lambda_1^{(2)} & 0 \\ 0 & \lambda_2^{(2)} \end{bmatrix} \quad (3.10)$$

e as matrizes de transição de estados

$$\mathbf{P}^{(1)} = \begin{bmatrix} p_{11}^{(1)} & p_{12}^{(1)} \\ p_{21}^{(1)} & p_{22}^{(1)} \end{bmatrix} \quad \text{e} \quad \mathbf{P}^{(2)} = \begin{bmatrix} p_{11}^{(2)} & p_{12}^{(2)} \\ p_{21}^{(2)} & p_{22}^{(2)} \end{bmatrix} \quad (3.11)$$

Desta sobreposição resulta um 4-dMMPP, cujas taxas de Poisson resultam da soma das taxas dos estados dos 2-dMMPPs e as probabilidades de transição resultam da multiplicação das probabilidades de transição dos 2-dMMPPs. As matrizes que descrevem as taxas de Poisson e as probabilidades de transição de estados do 4-dMMPP resultante podem ser obtidas pela soma e produto de Kronecker, i.e,

$$\mathbf{\Lambda} = \mathbf{\Lambda}^{(1)} \oplus \mathbf{\Lambda}^{(2)} = \begin{bmatrix} \lambda_1^{(1)} + \lambda_1^{(2)} & 0 & 0 & 0 \\ 0 & \lambda_1^{(1)} + \lambda_2^{(2)} & 0 & 0 \\ 0 & 0 & \lambda_2^{(1)} + \lambda_1^{(2)} & 0 \\ 0 & 0 & 0 & \lambda_2^{(1)} + \lambda_2^{(2)} \end{bmatrix} \quad (3.12)$$

e

$$\mathbf{P} = \mathbf{P}^{(1)} \otimes \mathbf{P}^{(2)} = \begin{bmatrix} p_{11}^{(1)} p_{11}^{(2)} & p_{11}^{(1)} p_{12}^{(2)} & p_{12}^{(1)} p_{11}^{(2)} & p_{12}^{(1)} p_{12}^{(2)} \\ p_{11}^{(1)} p_{21}^{(2)} & p_{11}^{(1)} p_{22}^{(2)} & p_{12}^{(1)} p_{21}^{(2)} & p_{12}^{(1)} p_{22}^{(2)} \\ p_{21}^{(1)} p_{11}^{(2)} & p_{21}^{(1)} p_{12}^{(2)} & p_{22}^{(1)} p_{11}^{(2)} & p_{22}^{(1)} p_{12}^{(2)} \\ p_{21}^{(1)} p_{21}^{(2)} & p_{21}^{(1)} p_{22}^{(2)} & p_{22}^{(1)} p_{21}^{(2)} & p_{22}^{(1)} p_{22}^{(2)} \end{bmatrix} \quad (3.13)$$

3.3 Modelo $M2^L$ -dMMPP

3.3.1 Descrição

O modelo proposto nesta Tese terá como base a sobreposição de L 2-dMMPPs independentes

$$(X^{(l)}, J^{(l)}) \sim \text{dMMPP}_2(\mathbf{P}^{(l)}, \mathbf{\Lambda}^{(l)}), l = 1, 2, \dots, L \quad (3.14)$$

e um M -dMMPP

$$(X^{(L+1)}, J^{(L+1)}) \sim \text{dMMPP}_M(\mathbf{P}^{(L+1)}, \mathbf{\Lambda}^{(L+1)}). \quad (3.15)$$

De notar que, em particular, para $l = 1, 2, \dots, L$,

$$\mathbf{P}^{(l)} = \begin{bmatrix} p_{11}^{(l)} & p_{12}^{(l)} \\ p_{21}^{(l)} & p_{22}^{(l)} \end{bmatrix}, \quad \mathbf{\Lambda}^{(l)} = \begin{bmatrix} \lambda_1^{(l)} & 0 \\ 0 & \lambda_2^{(l)} \end{bmatrix} \quad (3.16)$$

e é assumido que $p_{12}^{(l)} + p_{21}^{(l)} < 1$. Considera-se também que $J^{(1)}, J^{(2)}, \dots, J^{(L+1)}$ são cadeias ergódicas em regime estacionário. Para $l = 1, 2, \dots, L$ a distribuição estacionária de $J^{(l)}$ é designada por $\pi^{(l)} = [\pi_1^{(l)} \pi_2^{(l)}]$. De modo semelhante a distribuição estacionária de $J^{(L+1)}$ é designada por $\pi^{(L+1)} = [\pi_1^{(L+1)} \pi_2^{(L+1)} \dots \pi_M^{(L+1)}]$.

O resultado da sobreposição é o processo

$$(X, J) = \left(\sum_{l=1}^{L+1} X^{(l)}, (J^{(1)}, J^{(2)}, \dots, J^{(L+1)}) \right) \sim \text{dMMPP}_S(\mathbf{P}, \mathbf{\Lambda}) \quad (3.17)$$

onde

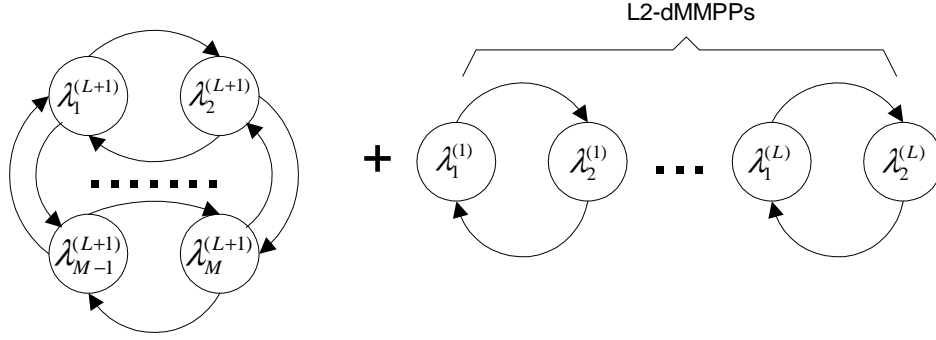
$$S = \{1, 2\}^L \times \{1, 2, \dots, M\} \quad (3.18)$$

$$\mathbf{P} = \mathbf{P}^{(1)} \otimes \mathbf{P}^{(2)} \otimes \dots \otimes \mathbf{P}^{(L+1)} \quad (3.19)$$

$$\mathbf{\Lambda} = \mathbf{\Lambda}^{(1)} \oplus \mathbf{\Lambda}^{(2)} \oplus \dots \oplus \mathbf{\Lambda}^{(L+1)} \quad (3.20)$$

De notar que a cadeia de Markov J está também em regime estacionário. (X, J) será referido como sendo o processo $M2^L$ -dMMPP. A sobreposição de um M -dMMPP e L 2-dMMPPs está ilustrada na Figura 3.2.

No procedimento de inferência proposto os valores de L e M não são fixados *a priori*, antes sendo determinados como parte do próprio procedimento. No entanto, na descrição que se segue serão considerados como fixos, sem perda de generalidade. Neste procedimento de inferência pretende-se que, relativamente aos incrementos do processo de chegadas (X) , os L 2-dMMPPs descrevam a função de autocovariância e


 Figura 3.2: Sobreposição de um modelo M -dMMPP e L modelos 2-dMMPP.

o M -dMMPP descreva a distribuição, tendo em conta a contribuição devida às distribuições dos L 2-dMMPP. De modo a facilitar a descrição do procedimento de inferência é conveniente definir o processo de incrementos associado a $X^{(1)}, X^{(2)}, \dots, X^{(L+1)}$, e a X , que será designado por $Y^{(1)}, Y^{(2)}, \dots, Y^{(L+1)}$, e Y , respectivamente. Assim,

$$Y_k^{(l)} = X_{k+1}^{(l)} - X_k^{(l)}, \quad l = 1, 2, \dots, L + 1 \quad (3.21)$$

e

$$Y_k = X_{k+1} - X_k \quad (3.22)$$

para $k = 0, 1, \dots$. De notar que Y_k é o número (total) de chegadas no intervalo de amostragem k e $Y_k^{(l)}$ é o número de chegadas devidas ao l -ésimo processo de chegadas. Assim,

$$Y_k = \sum_{l=1}^{L+1} Y_k^{(l)}, \quad k = 0, 1, 2, \dots \quad (3.23)$$

Além disso, $Y^{(1)}, Y^{(2)}, \dots, Y^{(L+1)}$, e Y , são seqüências estacionárias.

Designaremos as funções de probabilidade (marginal) dos processos $Y^{(1)}, Y^{(2)}, \dots, Y^{(L+1)}$, e Y , respectivamente por $\{f_l(k), k = 0, 1, 2, \dots\}$ e $\{f(k), k = 0, 1, 2, \dots\}$, e designaremos também a função de probabilidade de uma variável aleatória de Poisson com média μ por $\{g_\mu(k), k = 0, 1, 2, \dots\}$, para $\mu \in [0, +\infty[$

$$g_\mu(k) = e^{-\mu} \frac{\mu^k}{k!}, \quad k = 0, 1, 2, \dots \quad (3.24)$$

Para $l = 1, 2, \dots, L$, a distribuição marginal de $Y^{(l)}$ (isto é a distribuição de $Y_k^{(l)}$, para $k = 0, 1, \dots$) é a soma pesada de duas distribuições de Poisson com médias $\lambda_1^{(l)}$ e $\lambda_2^{(l)}$, e pesos $\pi_1^{(l)}$ e $\pi_2^{(l)}$, respectivamente. As funções de probabilidade de $Y^{(l)}$, $l = 1, 2, \dots, L$, serão então dadas por

$$f_l(k) = \pi_1^{(l)} g_{\lambda_1^{(l)}}(k) + \pi_2^{(l)} g_{\lambda_2^{(l)}}(k), \quad k = 0, 1, 2, \dots \quad (3.25)$$

e as funções de autocovariância por

$$\gamma_k^{(l)} = \text{Cov}(Y_0^{(l)}, Y_k^{(l)}) = \pi_1^{(l)} \pi_2^{(l)} \left(\lambda_2^{(l)} - \lambda_1^{(l)} \right)^2 e^{kc_l}, \quad k = 1, 2, \dots \quad (3.26)$$

onde $c_l = \ln(1 - p_{12}^{(l)} - p_{21}^{(l)})$. De notar que, em particular, as funções de autocovariância de $Y^{(1)}, Y^{(2)}, \dots, Y^{(L)}$ exibem um decaimento exponencial para zero.

Como se deseja que o processo M -dMMPP aproxime a distribuição dos incrementos do processo de chegada sem contribuir para a função de autocovariância dos incrementos do processo $M2^L$ -dMMPP, a cadeia de Markov $J^{(L+1)}$ não deve possuir memória. Isto é conseguido fazendo

$$\mathbf{P}^{(L+1)} = \begin{bmatrix} \pi_1^{(L+1)} & \pi_2^{(L+1)} & \dots & \pi_M^{(L+1)} \\ \pi_1^{(L+1)} & \pi_2^{(L+1)} & \dots & \pi_M^{(L+1)} \\ \dots & \dots & \dots & \dots \\ \pi_1^{(L+1)} & \pi_2^{(L+1)} & \dots & \pi_M^{(L+1)} \end{bmatrix}. \quad (3.27)$$

Esta solução implica que $Y^{(L+1)}$ seja uma sequência de variáveis aleatórias independentes e identicamente distribuídas cuja distribuição é uma mistura de M variáveis aleatórias de Poisson com médias $\lambda_i^{(L+1)}$ e pesos $\pi_i^{(L+1)}$, para $i = 1, 2, \dots, M$. Em consequência, a função de probabilidade de $Y^{(L+1)}$ é dada por

$$f_{L+1}(k) = \sum_{j=1}^M \pi_j^{(L+1)} g_{\lambda_j^{(L+1)}}(k), \quad k = 0, 1, 2, \dots \quad (3.28)$$

e a função de autocovariância é nula para todos os desfasamentos temporais positivos; isto é,

$$\gamma_k^{(L+1)} = \text{Cov}(Y_0^{(L+1)}, Y_k^{(L+1)}) = 0, \quad k \geq 1. \quad (3.29)$$

De acordo com (3.23), a função de probabilidade de Y é dada por:

$$f(k) = (f_1 \oplus f_2 \oplus \dots \oplus f_{L+1})(k) \quad (3.30)$$

onde \oplus designa a convolução de funções de probabilidade. Y é uma sequência de variáveis aleatórias com uma distribuição dada por uma soma pesada de variáveis aleatórias de Poisson (note-se que a soma de misturas independentes de variáveis de Poisson é ainda uma mistura de variáveis aleatórias de Poisson), e a função de probabilidade de Y é dada por

$$f(k) = \sum_{j_1=1}^2 \sum_{j_2=1}^2 \dots \sum_{j_L=1}^2 \sum_{j_{L+1}=1}^M \left(\prod_{l=1}^{L+1} \pi_{j_l}^{(l)} \right) g_{\sum_{l=1}^{L+1} \lambda_{j_l}^{(l)}}(k). \quad (3.31)$$

Considerando ainda (3.23) e de acordo com (3.26) e (3.29), conclui-se que a função de autocovariância de Y é dada por

$$\begin{aligned}\gamma_k &= \text{Cov}(Y_0, Y_k) = \sum_{l=1}^{L+1} \text{Cov}(Y_0^{(l)}, Y_k^{(l)}) \\ &= \sum_{l=1}^L \pi_1^{(l)} \pi_2^{(l)} \left(\lambda_2^{(l)} - \lambda_1^{(l)}\right)^2 e^{kc_l}\end{aligned}\tag{3.32}$$

para $k = 1, 2, \dots$

3.3.2 Procedimento de inferência dos parâmetros

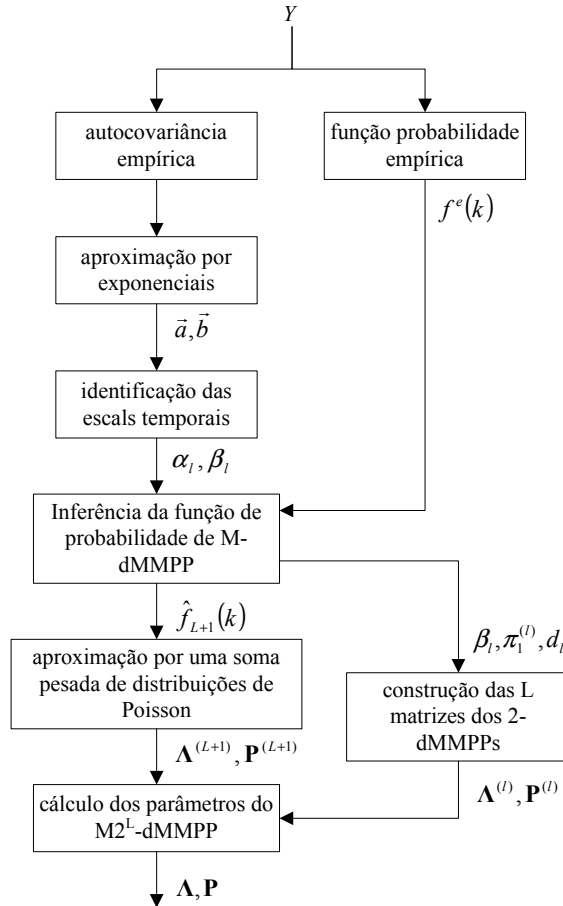
No que se segue, os processos $Y^{(1)}, Y^{(2)}, \dots, Y^{(L)}$ serão referidos como os processos 2-dMMPP, e os processos $Y^{(L+1)}$ e Y serão denominados, respectivamente, por processo M -dMMPP e processo $M2^L$ -dMMPP.

O procedimento de inferência dos parâmetros pode ser dividido em quatro passos: (i) aproximação da autocovariância empírica por uma soma pesada de exponenciais e identificação das escalas temporais de interesse, (ii) inferência da função de probabilidade do processo M -dMMPP e cálculo dos parâmetros dos 2-dMMPPs, (iii) inferência das taxas de chegada de Poisson e probabilidades de transição entre estados do M -dMMPP e (iv) cálculo dos parâmetros finais do modelo $M2^L$ -dMMPP. O diagrama de fluxos que ilustra o procedimento de inferência do modelo está representado na Figura 3.3. Nas sub-seções seguintes estes passos serão descritos em detalhe.

3.3.2.1 Aproximação da autocovariância e identificação das escalas temporais de interesse

A abordagem usada foi a de aproximar a autocovariância por um número elevado de exponenciais (com parâmetros reais) e depois agregar aquelas que tivessem um decaimento similar, ou seja, que possuísem constantes de tempo dentro da mesma escala temporal. Uma abordagem semelhante foi seguida em [YKT01, AN98, FW97].

Como primeiro passo, a autocovariância empírica é aproximada por uma soma de K exponenciais com pesos reais positivos e decaimentos reais negativos, usando para o efeito o algoritmo modificado de Prony [OS95]. O valor de K pode ser determinado recorrendo a um método iterativo, onde progressivamente se aumenta o número de exponenciais usadas na aproximação. K será o número máximo de exponenciais usadas na aproximação para o qual o algoritmo modificado de Prony resulta em amplitudes e constantes de tempo puramente reais. A aplicação do algoritmo de Prony tem como resultado dois vectores,


 Figura 3.3: Diagrama de fluxo do procedimento de inferência do modelo $M2^L$ -dMMPP.

$$\vec{a} = \begin{bmatrix} a_1 & \dots & a_K \end{bmatrix} \quad \vec{b} = \begin{bmatrix} b_1 & \dots & b_K \end{bmatrix} \quad (3.33)$$

os quais correspondem à função de aproximação

$$\text{Cov}_K(\vec{a}, \vec{b}) = \sum_{i=1}^K a_i e^{-b_i k}, \quad k = 1, 2, 3, \dots \quad (3.34)$$

Neste ponto é necessário identificar as componentes da autocovariância que caracterizam as diferentes escalas temporais. Definem-se L escalas temporais diferentes, nas quais os decaimentos (constantes de tempo das exponenciais) da autocovariância, $b_i, i = 1, \dots, K$, estão na mesma escala logarítmica. Para explicar este procedimento é útil assumir que $b_j \leq b_{j+1}, 1 \leq j \leq K - 1$. O valor de L é calculado usando o seguinte procedimento iterativo. Começando com $l = 1$ e $i_l = 1$ calcula-se i_{l+1} através da seguinte expressão

$$i_{l+1} = \min \{ K + 1, \inf \{ j : i_l < j \leq K \wedge [\log_{10}(b_j)] > [\log_{10}(b_{j-1})] \} \}, \quad (3.35)$$

em que $\lceil \dots \rceil$ representa o arredondamento para o inteiro imediatamente superior.

Se $i_{l+1} > K$ então faz-se $L = l$ e o procedimento termina; caso contrário incrementa-se l de uma unidade e repete-se o procedimento. De notar que, em particular

$$\lceil \log_{10}(b_{i_l}) \rceil = \lceil \log_{10}(b_{i_{l+1}}) \rceil = \dots = \lceil \log_{10}(b_{i_{l+1}-1}) \rceil \quad (3.36)$$

mas, se $i_{l+1} \leq K$,

$$\lceil \log_{10}(b_{i_l}) \rceil < \lceil \log_{10}(b_{i_{l+1}}) \rceil. \quad (3.37)$$

Para $j = 1, 2, \dots, L$, considera-se que os decaimentos desde b_{i_l} até $b_{i_{l+1}-1}$ caracterizam a mesma escala temporal. Então as $n_l = |i_{l+1} - i_l|$ componentes são agregadas num único componente com os seguintes parâmetros:

$$\alpha_l = \sum_{k=i_l}^{i_{l+1}-1} a_k \quad \text{e} \quad \beta_l = -\frac{\sum_{k=i_l}^{i_{l+1}-1} a_k b_k}{\alpha_l}. \quad (3.38)$$

Estes parâmetros são usados para ajustar a função de autocovariância do processo 2-dMMPP, $Y^{(l)}$, tendo em conta (3.26), (3.34) e (3.38). Destas relações resulta que

$$\alpha_l = d_l^2 \pi_1^{(l)} \pi_2^{(l)} \quad \text{e} \quad \beta_l = c_l \quad (3.39)$$

onde $d_l = \lambda_2^{(l)} - \lambda_1^{(l)}$, isto é, a função de autocovariância ajustada de $Y_1 + Y_2 + \dots + Y_L$ é

$$\sum_{l=1}^L \alpha_l e^{k\beta_l}, \quad k = 1, 2, \dots \quad (3.40)$$

A Figura 3.4 mostra o resultado da identificação das escalas temporais da função de autocovariância empírica da captura Bellcore (descrita na secção A.3 do Apêndice A), e à qual foi aplicado o modelo descrito neste capítulo (ver secção 3.4). Foram identificadas 3 escalas temporais caracterizadas pelos decaimentos $\beta_1 = 3.14 \times 10^{-4}$, $\beta_2 = 1.33 \times 10^{-2}$ e $\beta_3 = 1.44 \times 10^{-1}$.

3.3.2.2 Inferência da função de probabilidade do M -dMMPP e cálculo dos parâmetros dos L 2-dMMPPs

O próximo passo é a inferência da função massa de probabilidade (FMP) do M -dMMPP a partir da FMP empírica dos dados originais. A relação entre as funções de probabilidade dos processos 2-dMMPPs, M -dMMPP e $M2^L$ -dMMPP é definida por (3.30).

De modo a simplificar a desconvolução de $f_{L+1}(k)$ e $f_l(k)$, $l = 1, \dots, L$, considera-se que a taxa de chegada de Poisson no primeiro estado de cada um dos L 2-dMMPPs é

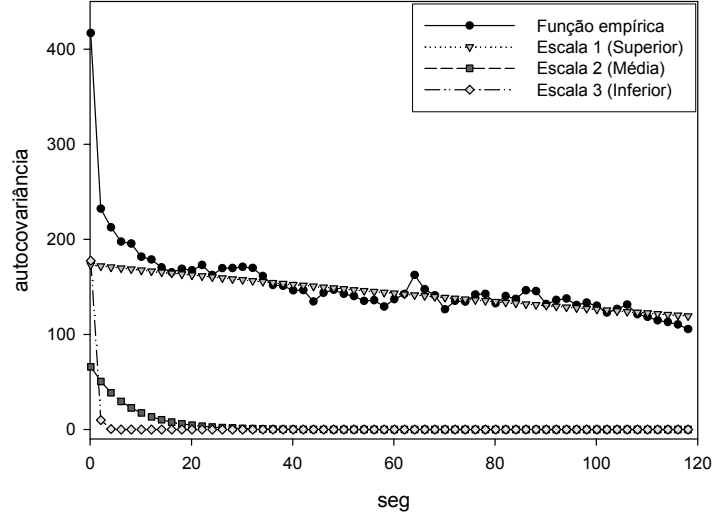


Figura 3.4: Identificação das escalas temporais de interesse.

zero; ou seja, $\lambda_1^{(l)} = 0$ e $\lambda_2^{(l)} = d_l$, para $l = 1, \dots, L$. De (3.39) resulta,

$$d_l = \sqrt{\frac{\alpha_l}{\pi_1^{(l)} \pi_2^{(l)}}}, \quad l = 1, 2, \dots, L. \quad (3.41)$$

A função de probabilidade do M -dMMPP, f_{L+1} , é inferida a partir da função de probabilidade empírica dos dados f^e , e das L funções de probabilidade dos 2-dMMPPs, designadas por \hat{f}_l , $l = 1, 2, \dots, L$. f_{L+1} está assim estreitamente relacionada, através da equação (3.30), com os parâmetros já inferidos na sub-secção 3.3.2.1 e com as probabilidades $\pi_1^{(l)}$, $l = 1, 2, \dots, L$. Mais precisamente, f_{L+1} tem de ser ajustada conjuntamente com os parâmetros $\pi_1^{(l)}$, $l = 1, \dots, L$, utilizando para o efeito um processo de minimização com restrições:

$$\min_{\{\pi_1^{(l)}, l=1, \dots, L\}, \{f_{L+1}(k), k=0, 1, \dots\}} \sum_k |o^e(k)| \quad (3.42)$$

onde

$$o^e(k) = f^e(k) - \left(\hat{f}_1 \oplus \dots \oplus \hat{f}_L \oplus f_{L+1} \right) (k) \quad (3.43)$$

sujeito às restrições definidas por (3.39) e a

$$\begin{aligned} 0 < \pi_1^{(l)} < 1, \quad l = 1, 2, \dots, L, \\ f_{L+1}(k) > 0, \quad k = 0, 1, \dots, \\ \text{e} \quad \sum_{k=0}^{+\infty} f_{L+1}(k) = 1. \end{aligned} \quad (3.44)$$

A função de probabilidade ajustada do M -dMMPP é designada por \hat{f}_{L+1} . Não se permite que $\pi_1^{(l)}$ seja 0 ou 1, pois se for 0 o l -ésimo 2-dMMPP degenera num processo

de Poisson simples e se for 1 torna-se um processo nulo. A função de probabilidade empírica, $f^e(k)$, é inferida para a gama de valores definida por

$$k = 0, 1, \dots, \max \left\{ 1.1 \sum_{l=1}^L d_l, \max\{Y\} \right\}. \quad (3.45)$$

As funções de probabilidade $\hat{f}_l(k), l = 1, \dots, L$ são calculadas para a mesma gama de valores. No entanto, para reduzir o número de pontos usados nas convoluções apenas se considera a sub-gama de valores definida por

$$k = \min_j \left\{ \hat{f}_l(j) > \xi_l \right\}, \min_j \left\{ \hat{f}_l(j) > \xi_l \right\} + 1, \dots, \max_j \left\{ \hat{f}_l(j) > \xi_l \right\} \quad (3.46)$$

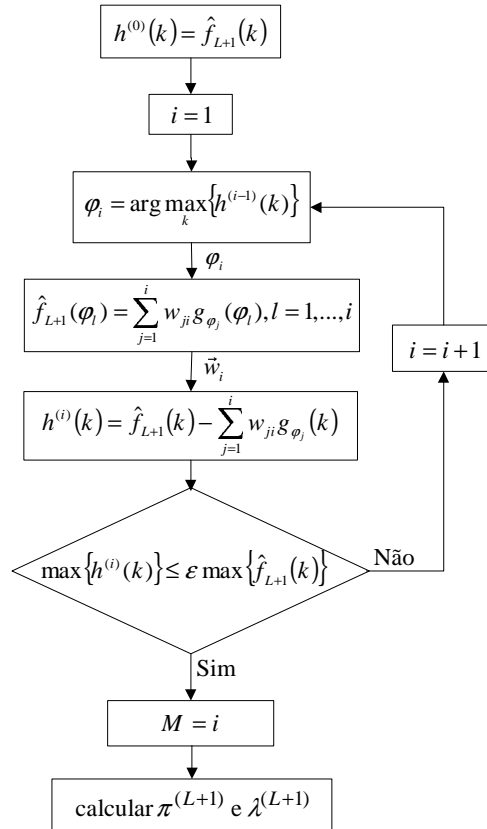
com $\xi_l = 10^{-4} \max\{\hat{f}_l(k)\}$. O processo de minimização com restrições dado por (3.42)–(3.44) é um problema de programação não-linear. Este tipo problemas é, em geral, de resolução difícil. Assim, foram consideradas duas aproximações: (i) fez-se $\pi_1^{(l)} = \pi_1^{(l+1)}$, $l = 1, \dots, L - 1$ e (ii) as possíveis soluções de $\pi_1^{(l)}$ foram restritas a uma gama discreta de valores, tais que $\pi_1^{(l)} = 0.001k$, $k = 1, \dots, 999$. Estas aproximações tiveram um impacto pouco significativo nos resultados obtidos com este procedimento de inferência de parâmetros, em particular naqueles apresentados na secção 3.4 deste capítulo. É utilizado um procedimento de procura para determinar o valor mínimo da função objectivo. Para cada valor possível de $\pi_1^{(l)}$ são efectuados os seguintes passos: (i) cálculo da FMP dos L 2-dMMPPs a partir de $\hat{f}_1 \oplus \hat{f}_1 \oplus \dots \oplus \hat{f}_L$ e (3.25), (ii) cálculo da estimativa da FMP do M -dMMPP, \hat{f}_{L+1} , desconvolvendo da FMP empírica a contribuição dos L 2-dMMPPs, e considerando todas as restrições impostas ao processo e (iii) cálculo do erro de aproximação entre a FMP empírica e a FMP conjunta dos L 2-dMMPPs e do M -dMMPP. Após o teste de todos os valores considerados de $\pi_1^{(l)}$, f_{L+1} será igual à estimativa \hat{f}_{L+1} para a qual o erro de aproximação das FMPs foi menor. Os correspondentes valores de $\pi_1^{(l)}$ serão as probabilidades em regime estacionário dos L 2-dMMPPs.

Neste ponto todos os parâmetros dos L 2-dMMPPs, $Y^{(1)}, Y^{(2)}, \dots, Y^{(L)}$, foram já determinados. As matrizes correspondentes

$$(\mathbf{P}^{(l)}, \mathbf{\Lambda}^{(l)}), l = 1, 2, \dots, L \quad (3.47)$$

podem ser construídas do seguinte modo:

$$\mathbf{P}^{(l)} = \begin{bmatrix} 1 - \pi_2^{(l)}(1 - e^{\beta_l}) & \pi_2^{(l)}(1 - e^{\beta_l}) \\ \pi_1^{(l)}(1 - e^{\beta_l}) & 1 - \pi_1^{(l)}(1 - e^{\beta_l}) \end{bmatrix} \quad (3.48)$$


 Figura 3.5: Algoritmo para o cálculo do número de estados e taxas de chegada do M -dMMPP.

e

$$\mathbf{\Lambda}^{(l)} = \begin{bmatrix} 0 & 0 \\ 0 & d_l \end{bmatrix}. \quad (3.49)$$

3.3.2.3 Inferência dos parâmetros do M -dMMPP

O próximo passo é a inferência do número de estados e das taxas de chegada de Poisson do M -dMMPP a partir de \hat{f}_{L+1} . Para efectuar esta operação \hat{f}_{L+1} é aproximada por uma soma pesada de funções de probabilidade de Poisson, isto é, como sendo a função de probabilidade de uma mistura de variáveis de Poisson com um número indeterminado (à partida) de componentes.

O ajuste é efectuado usando um algoritmo que progressivamente subtrai de \hat{f}_{L+1} uma função de probabilidade (de Poisson). Este procedimento está descrito no diagrama de fluxos da Figura 3.5. Seja $g_{\varphi_i}(k)$ a i -ésima função de probabilidade de Poisson, com média φ_i . Define-se $h^{(i)}(k)$ como sendo a diferença entre $\hat{f}_{L+1}(k)$ e a soma pesada de funções de probabilidade de Poisson após a i -ésima iteração. Inicialmente, faz-se $h^{(0)}(k) = \hat{f}_{L+1}(k)$. Em cada iteração, detecta-se o máximo de $h^{(i-1)}(k)$ e

o valor de k correspondente, $\varphi_i = \arg \max_k \{h^{(i-1)}(k)\}$, será considerado a i -ésima taxa de chegada de Poisson do M -dMMPP. São então calculados os pesos de cada uma das funções de probabilidade de Poisson já inferidas, $\vec{w}_i = [w_{1i}, w_{2i}, \dots, w_{ii}]$, usando para o efeito o seguinte conjunto de equações lineares:

$$\hat{f}_{L+1}(\varphi_l) = \sum_{j=1}^i w_{ji} g_{\varphi_j}(\varphi_l), \quad l = 1, \dots, i. \quad (3.50)$$

Deste modo é assegurado que o ajuste entre $\hat{f}_{L+1}(k)$ e a soma pesada de funções de probabilidade de Poisson é exacto nos pontos φ_l , para $l = 1, 2, \dots, i$. O passo final de cada iteração é o cálculo da nova função diferença

$$h^{(i)}(k) = \hat{f}_{L+1}(k) - \sum_{j=1}^i w_{ji} g_{\varphi_j}(k). \quad (3.51)$$

O algoritmo termina quando o máximo de $h^{(i)}(k)$ é inferior a uma percentagem pré-definida (ϵ) do máximo de $\hat{f}_{L+1}(k)$ e M será então igualado a i .

Depois de M ter sido determinado, os parâmetros do M -dMMPP, $\{(\pi_j^{(L+1)}, \lambda_j^{(L+1)}), j = 1, 2, \dots, M\}$, são então igualados a

$$\pi_j^{(L+1)} = \frac{w_{jM}}{\sum_{m=1}^M w_{mM}} \quad \text{e} \quad \lambda_j^{(L+1)} = \varphi_j. \quad (3.52)$$

Na Figura 3.6 é possível observar um exemplo da aplicação iterativa deste algoritmo a uma função de probabilidade. Após a primeira iteração foi detectado um máximo em 30 e conseqüentemente foi inferida uma função de probabilidade de Poisson com média 30 e peso 0.6235. Subtraindo à função de probabilidade inicial a função inferida obtém-se a função $h^{(1)}$, a qual tem um máximo em 45. Assim, a segunda função de probabilidade de Poisson inferida tem média 45. Recalculando os pesos das funções de probabilidade já inferidas obtemos um peso de 0.6 para a primeira e de 0.4 para a segunda. Após a segunda iteração o algoritmo termina, pois a função $h^{(2)}$ tem apenas valores residuais.

3.3.2.4 Construção do modelo $M2^L$ -dMMPP

Finalmente, o processo $M2^L$ -dMMPP pode ser construído recorrendo às equações (3.19) e (3.20), usando as matrizes $\mathbf{P}^{(i)}$ e $\mathbf{\Lambda}^{(i)}$, $i = 1, \dots, L$ calculadas de acordo com a expressão (3.16), a matriz $\mathbf{P}^{(L+1)}$ construída de acordo com as expressões (3.27) e $\mathbf{\Lambda}^{(L+1)}$ construídas de acordo com as expressões (3.3) e com os parâmetros $\lambda_j^{(L+1)}$, $j = 1, \dots, M$ obtidos por (3.52).

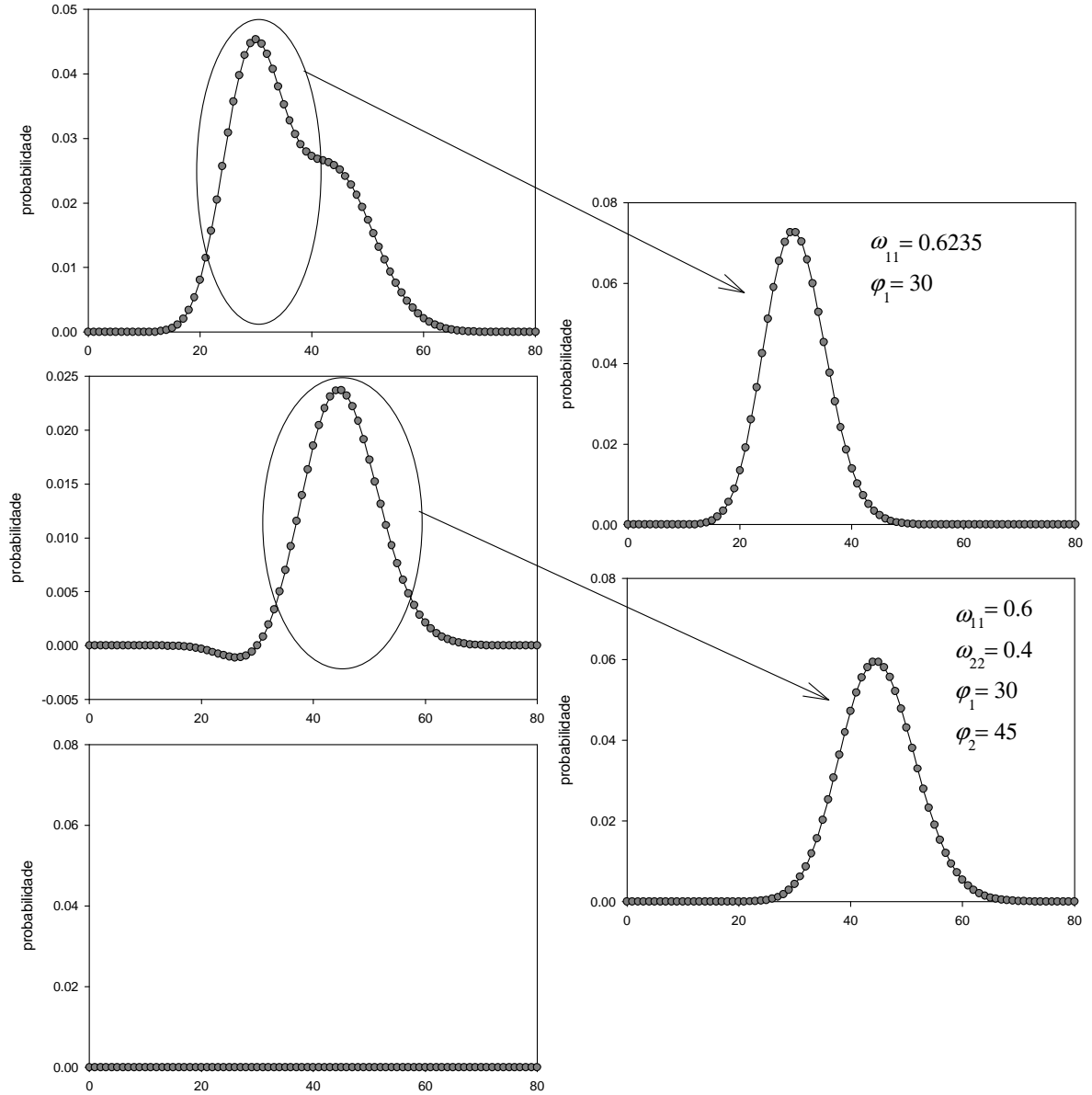


Figura 3.6: Exemplo do algoritmo de ajuste de uma função a uma soma pesada de funções de probabilidade de Poisson. (À esquerda, de cima para baixo) função que se pretende ajustar e funções diferença após uma, $h^{(1)}$, e duas, $h^{(2)}$, iterações. (À direita, de cima para baixo) Funções de probabilidade de Poisson inferidas na primeira e segunda iteração, g_{φ_1} e g_{φ_2} .

3.4 Resultados

Nesta secção serão apresentados e avaliados os resultados do modelo $M2^L$ -dMMPP e do respectivo procedimento de inferência de parâmetros introduzido nas secções anteriores. Na avaliação dos resultados do modelo $M2^L$ -dMMPP foram também incluídos os resultados de uma versão simplificada do modelo $M2^L$ -dMMPP, que será designado por $M2$ -dMMPP. O modelo $M2$ -dMMPP faz apenas o ajuste da cauda da função de autocovariância, utilizando a exponencial de menor decaimento das L exponenciais

usadas no ajuste da autocovariância do $M2^L$ -dMMPP. Pretende-se deste modo verificar se o ajuste da cauda da autocovariância é suficiente para efeitos de modelação de tráfego. Ambos os modelos foram aplicados a duas capturas de tráfego: (i) a captura Bellcore de Outubro e (ii) uma captura feita na Universidade de Aveiro (UA). Ambos as capturas estão descritos em pormenor no Apêndice A. O intervalo de amostragem usado foi 0.1 segundos em ambas as capturas. O parâmetro ϵ foi escolhido de modo a obter um compromisso entre o número de parâmetros do modelo e o ajuste eficaz das características do tráfego.

3.4.1 $M2^L$ -dMMPP

A captura Bellcore foi ajustada por um dMMPP com 40 estados, considerando $\epsilon = 0.2$. O procedimento de inferência devolveu os seguintes parâmetros: $\pi_1^{(l)} = 0.4, l = 1, \dots, L$, $L = 3$, $\alpha_1 = 1.73 \times 10^2$, $\alpha_2 = 6.69 \times 10^1$, $\alpha_3 = 2.05 \times 10^2$, $\beta_1 = 3.14 \times 10^{-4}$, $\beta_2 = 1.33 \times 10^{-2}$ e $\beta_3 = 1.44 \times 10^{-1}$. No que respeita à captura UA, esta foi ajustado a um dMMPP com 12 estados, considerando $\epsilon = 0.01$. No final da inferência de parâmetros obtiveram-se os seguintes resultados: $\pi_1^{(l)} = 0.4, l = 1, \dots, L$, $L = 2$, $\alpha_1 = 4.74 \times 10^1$, $\alpha_2 = 5.17 \times 10^1$, $\beta_1 = 4.81 \times 10^{-4}$ e $\beta_2 = 4.05 \times 10^{-2}$. Em ambos os casos, a função de autocovariância foi ajustada em 60 pontos, para deslocamentos temporais desde 1 até 1200 intervalos de amostragem, em intervalos de 20.

3.4.2 $M2$ -dMMPP

A captura Bellcore foi ajustada por um dMMPP com 22 estados, considerando $\epsilon = 0.05$. Um valor de $\pi_1^{(1)} = 0.3$ foi obtido com o procedimento de inferência de parâmetros. A exponencial usada no ajuste da cauda da função de autocovariância tinha como parâmetros $\alpha = 1.73 \times 10^2$ e $\beta = 3.14 \times 10^{-4}$, ou seja a exponencial com decaimento mais lento das usadas na sub-secção anterior. No que respeita à captura UA, esta foi ajustado a um dMMPP com 8 estados, considerando $\epsilon = 0.05$. No final da inferência de parâmetros obtiveram-se os seguintes resultados: $\pi_1^{(1)} = 0.3$, $\alpha_1 = 4.74 \times 10^1$ e $\beta_1 = 4.81 \times 10^{-4}$.

3.4.3 Comparação e avaliação

De modo a avaliar o comportamento dos métodos de inferência aqui apresentados efectuaram-se os seguintes testes: (i) comparação das estatísticas de primeira e segunda

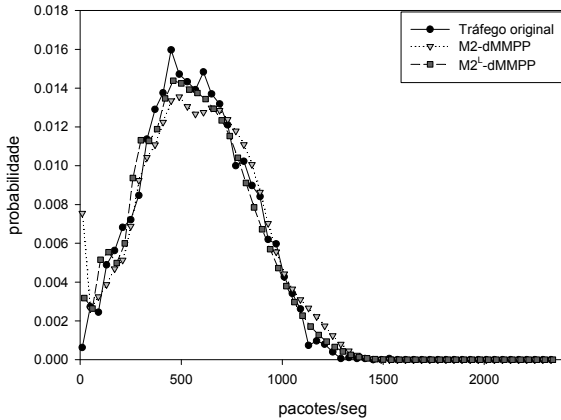


Figura 3.7: Ajuste da função massa de probabilidade, Bellcore.

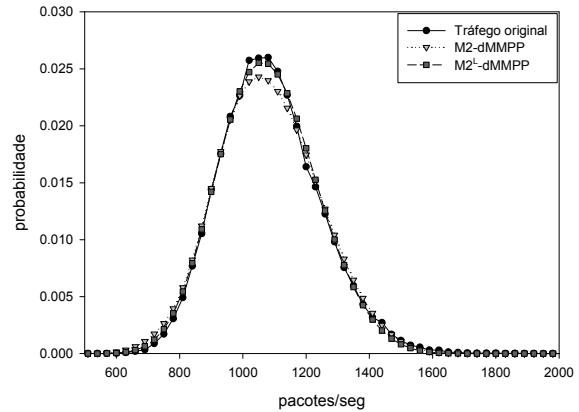


Figura 3.8: Ajuste da função massa de probabilidade, UA.

ordem e (ii) análise do comportamento do tráfego numa fila de espera, em termos de rácio de pacotes perdidos e atraso médio na fila de espera. Foram usados três tipos de tráfego: (i) o tráfego original considerando que todos os pacotes tinham tamanho fixo igual à média, (ii) tráfego gerado de acordo com o modelo $M2^L$ -dMMPP inferido, (iii) tráfego gerado de acordo com o modelo $M2$ -dMMPP inferido. Na análise do comportamento no tráfego numa fila, utilizou-se ainda o tráfego original mas considerando os tamanhos reais dos pacotes. Nas simulações efectuadas com base nos modelos inferidos foram calculados intervalos de confiança a 95% com base em 10 réplicas. Uma vez que o comprimento dos intervalos foi sempre muito reduzido optou-se por não os incluir nas figuras.

No caso das estatísticas de primeira ordem, pode-se observar a função massa de probabilidade (FMP) nas Figuras 3.7 e 3.8. Verifica-se que ambos os modelos conseguiram um ajuste muito bom da FMP. No que se refere às estatísticas de segunda ordem, mais concretamente à função de autocovariância (FAC) (Figuras 3.9 e 3.10), verifica-se que ambas as capturas de tráfego exibem características de memória longa. O modelo $M2$ -dMMPP obteve um bom ajuste da FAC apenas para valores de deslocamento temporal superiores (cauda da FAC). No caso do tráfego da captura Bellcore conseguiu-se um bom ajuste para deslocamentos temporais a partir dos 5 segundos. No caso da captura da UA só se conseguiu um ajuste aceitável para valores de deslocamentos superiores a 30 segundos. No que se refere ao modelo $M2^L$ -dMMPP, conseguiu-se um ajuste muito bom em toda a gama de deslocamentos temporais considerada.

De modo a avaliar o comportamento da captura Bellcore em fila de espera, fez-se variar a capacidade da fila de espera de 10 Kbytes até 8 Mbytes para uma taxa de serviço de 518 Kbytes/s (que corresponde a um factor de utilização de 0.7). Fez-

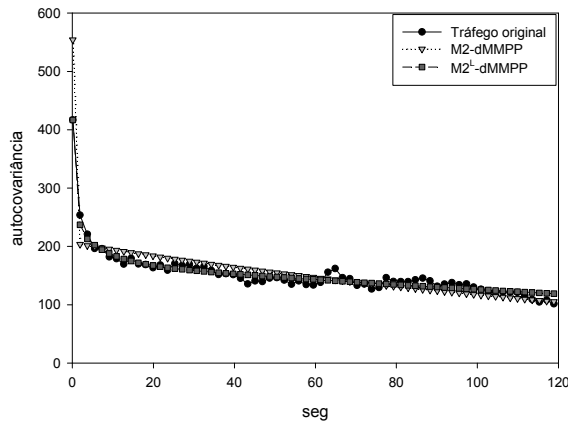


Figura 3.9: Ajuste da função de autocovariância, Bellcore.

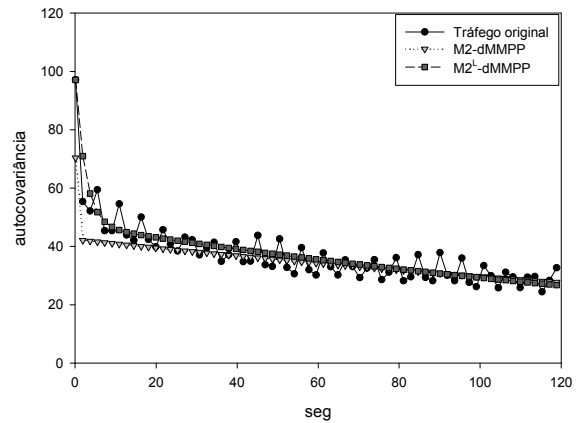


Figura 3.10: Ajuste da função de autocovariância, UA.

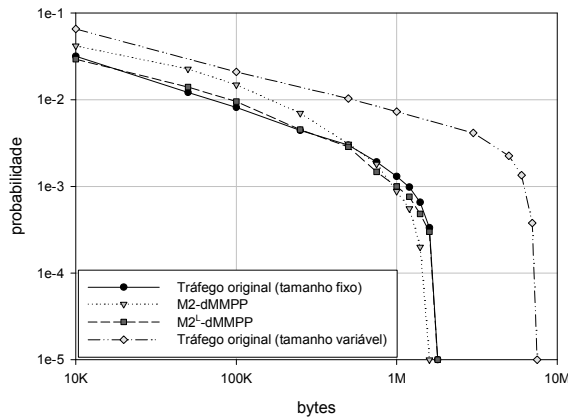


Figura 3.11: Rácio de perda de pacotes com uma utilização de 70%, Bellcore.

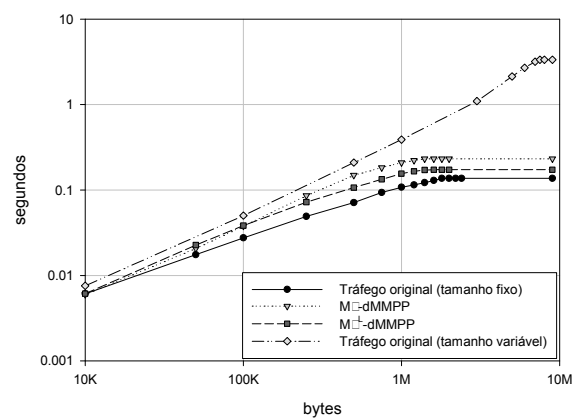


Figura 3.12: Atraso médio na fila de espera com uma utilização de 70%, Bellcore.

se ainda variar a unidade de armazenamento da fila de espera de 100 Kbytes até 70 Mbytes para uma taxa de serviço de 403 Kbytes/s (que corresponde a um factor de utilização de 0.9). Para a captura UA utilizaram-se factores de utilização de 0.9 e 0.98, os quais correspondem respectivamente a taxas de serviço de 726 Kbytes/s e 666 Kbytes/s. Em ambos os casos fez-se variar a unidade de armazenamento da fila de espera de 10 Kbytes até 40 Mbytes.

Nas Figuras 3.11 e 3.13 respeitantes aos resultados de rácio de perda de pacotes da captura Bellcore, pode-se observar que o modelo $M2^L$ -dMMPP obtém óptimos resultados para ambas as taxas de serviço. Com o modelo $M2$ -dMMPP obtém-se resultados ligeiramente piores, nomeadamente para a taxa de serviço de 70%. Uma explicação para este resultado é a falta de ajuste da FAC para deslocamentos temporais mais pequenos, também chamada de memória curta do processo. A modelação da memória curta parece assim ter uma importância maior quando a taxa de serviço é menor. Nas Figuras 3.12 e 3.14 são apresentados os resultados do atraso médio na fila de espera

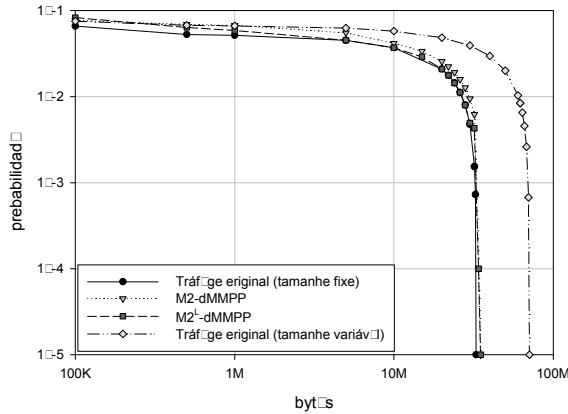


Figura 3.13: Rácio de perda de pacotes com uma utilização de 90%, Bellcore.

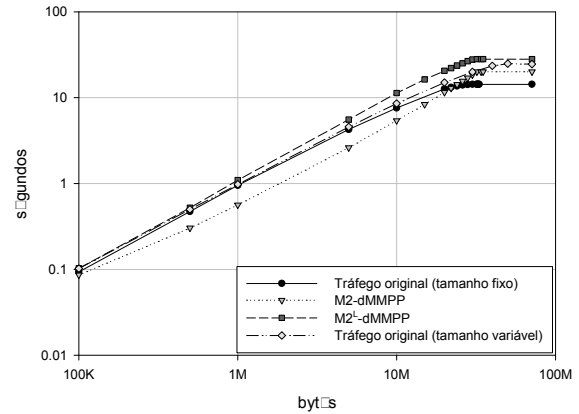


Figura 3.14: Atraso médio na fila de espera com uma utilização de 90%, Bellcore.

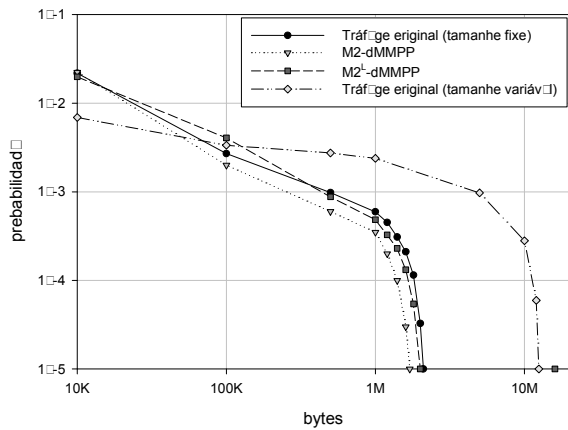


Figura 3.15: Rácio de perda de pacotes com uma utilização de 90%, UA.

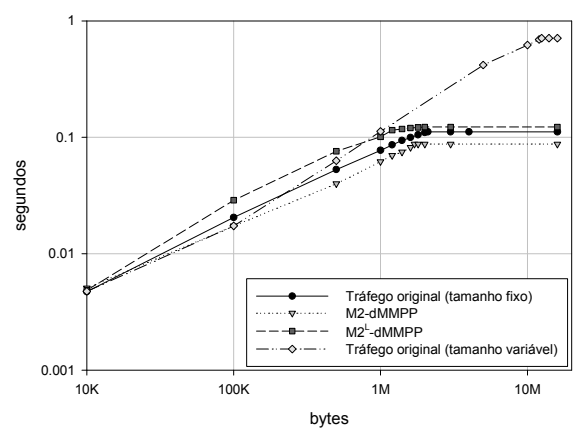


Figura 3.16: Atraso médio na fila de espera com uma utilização de 90%, UA.

obtidos nas simulações com a captura Bellcore. Observa-se que também aqui o modelo $M2^L$ -dMMPP obteve melhores resultados que o modelo $M2$ -dMMPP. Apesar de ambos os modelos apresentarem muito bons resultados o modelo $M2^L$ -dMMPP conseguiu capturar com um elevado grau de eficácia o comportamento do tráfego original, principalmente para unidades de armazenamento menores que 10 Mbytes.

Nas Figuras 3.15, 3.16, 3.17 e 3.18 são apresentados os resultados da simulação referentes à captura UA. As conclusões retiradas da observação dos mesmos são similares às da captura Bellcore.

Verifica-se que apesar dos resultados de simulação obtidos com o modelo $M2$ -dMMPP terem sido piores que os obtidos com o modelo $M2^L$ -dMMPP, estes não foram muito inferiores. O ajuste da memória curta do processo (ajuste da FAC para deslocamentos menores) tem alguma influência nos resultados que, contudo, parece não ser significativa. Pode-se concluir que em alguns casos pode ser vantajoso o uso do modelo mais simples, pois tem um menor número de parâmetros.

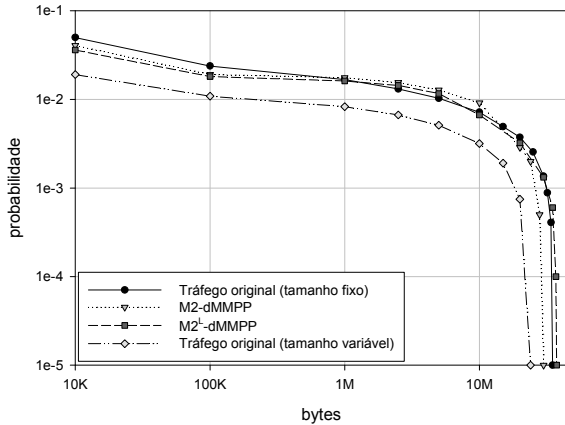


Figura 3.17: Rácio de perda de pacotes com uma utilização de 98%, UA.

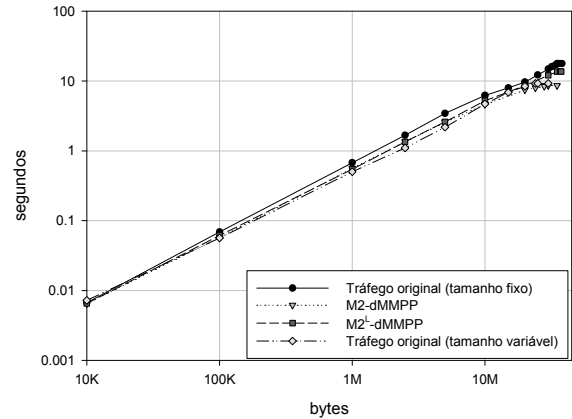


Figura 3.18: Atraso médio na fila de espera com uma utilização de 98%, UA.

Em todos os casos referidos verifica-se que a aproximação usada, de considerar os pacotes como sendo de igual tamanho, não permite a modelação eficaz do tráfego com tamanho de pacotes variável, quando se considera o tamanho real dos pacotes. É assim evidente a necessidade de usar modelos que também modelam o tamanho dos pacotes. No capítulo 4 será apresentado um modelo Markoviano capaz de modelar também o tamanho dos pacotes.

3.5 Conclusões

Neste capítulo foi apresentado um novo modelo de tráfego, e o respectivo procedimento de inferência de parâmetros. O modelo proposto é obtido por sobreposição de L 2-dMMPPs e de um M -dMMPP sem memória. Os 2-dMMPP modelam a função de autocovariância e o M -dMMPP modela a função massa de probabilidade, tomando em conta as restrições impostas pela modelação da autocovariância. Uma característica importante deste modelo é o facto de o número de estados do dMMPP não ser fixado à priori, permitindo uma adaptação ao tipo de tráfego que se está a caracterizar.

O modelo foi aplicado a tráfego real com características de memória longa e compararam-se tanto as estatísticas de primeira e segunda ordem, como o comportamento na passagem por uma fila de espera, do tráfego real e do tráfego gerado segundo o modelo inferido. Verificou-se que o modelo proposto conseguiu um ajuste muito bom das características do processo de chegada de pacotes. Os resultados mostram ainda, que apesar da relativa complexidade do tráfego, o procedimento de inferência foi capaz de ajustar o tráfego por dMMPPs com um número de estados relativamente reduzido. No entanto, quando o tráfego é formado por pacotes de tamanho variável (como é o

caso do tráfego IP) o modelo não captura eficazmente as suas características. Neste caso, é necessário incluir no modelo uma componente que permita ajustar as principais características estatísticas do processo de tamanhos dos pacotes. Este aspecto será considerado no capítulo seguinte.

Capítulo 4

Modelos Markovianos para os processos de chegadas e de tamanhos de pacotes

4.1 Introdução

No capítulo anterior foi ilustrada a importância da modelação conjunta do processo de chegada de pacotes e do processo de tamanhos de pacotes. Neste capítulo é apresentado um novo modelo de tráfego, e respectivo procedimento de inferência, que inclui a possibilidade de modelar os tamanhos de pacotes. O modelo proposto é um processo de Markov com chegadas em rajada em tempo discreto (dBMAP).

O processo Markoviano com chegadas em rajada (em tempo contínuo) BMAP foi apresentado por Lucantoni [Luc91] como a generalização de um processo Markoviano de chegadas [LMHN90]. O BMAP é um processo de chegadas genérico que tem as mesmas propriedades aditivas dos processos Markovianos, isto é, a sobreposição de BMAPs independentes é ainda um BMAP. Em [KLL03] foi proposto um modelo BMAP para modelar tráfego IP onde o procedimento de inferência de parâmetros tem por base o algoritmo EM (*expectation-maximization*). Em [Luc93] é feita uma resenha histórica dos BMAPs e das suas aplicações.

A versão em tempo discreto de um BMAP (a qual é designada por dBMAP ao longo deste capítulo, mas também é usualmente designada por D-BMAP) foi proposta por Blondia e Casals [BC92] e recebeu bastante atenção [Blo93, AC94, Ran95, GB96, Rid98], no entanto não tanta como a versão em tempo contínuo. O dBMAP pertence

à classe de processos aditivos de Markov uni-variados e como tal goza das propriedades desta classe de processos, algumas das quais serão apresentadas na secção 4.2. Entre outras aplicações, os dBMAs são usados para modelar fontes com chegadas em rajada e em reconhecimento de voz. Foi demonstrado que o resultado da sobreposição de processos dBMAs pode ser um processo com memória longa [GB96]. Um caso particular importante dos dBMAs é o processo de Poisson modulado à Markov em tempo discreto (dMMPP), descrito no capítulo 3.

Para a modelação de tráfego com tamanho de pacotes variável propõe-se neste capítulo da Tese o uso de um dBMAs em que as chegadas (de pacotes) ocorrem de acordo com um dMMPP e cada chegada é caracterizada por uma rajada (de bytes) cujo tamanho tem uma distribuição geral que depende da fase do dMMPP que descreve o processo de chegada de pacotes.

Este capítulo está organizado do seguinte modo. Na secção 4.2 é apresentado em pormenor o modelo e o respectivo procedimento de inferência de parâmetros. Na secção 4.3 são apresentados e discutidos os resultados da aplicação do modelo proposto a tráfego real. Finalmente na secção 4.4 são apresentadas algumas conclusões.

4.2 Modelo dBMAs

4.2.1 Descrição

O dBMAs pode ser visto como uma sequência aleatória de Markov cuja componente aditiva toma valores inteiros não-negativos. Assim, uma cadeia de Markov $(Y, J) = \{(Y_k, J_k), k \in \mathbb{N}_0\}$ com espaço de estados $\mathbb{N}_0 \times S$ é um dBMAs se

$$P(Y_{k+1} = m, J_{k+1} = j | Y_k = n, J_k = i) = \begin{cases} 0 & m < n \\ p_{ij} q_{ij}(m - n) & m \geq n \end{cases} \quad (4.1)$$

onde $\mathbf{P} = (p_{ij})_{i,j \in S}$ é uma matriz estocástica e, para cada par ordenado $(i, j) \in S^2$, $q_{ij} = \{q_{ij}(n), n \in \mathbb{N}_0\}$ é uma função de probabilidade em \mathbb{N}_0 , e $\mathbf{Q}(n) = (q_{ij}(n))_{i,j \in S}$. Isto implica, em particular que J é uma cadeia de Markov, chamada de *componente de Markov* ou *fase* de (Y, J) e S é um conjunto de estados moduladores ou o conjunto de fases. Quando o dBMAs (Y, J) é usado para modelar um processo de chegadas, Y_k pode ser interpretado como sendo o número total de chegadas (em bytes) até ao instante k e J será a fase do processo que influencia o número de pacotes que chegam num determinado intervalo de amostragem e a distribuição do tamanho destes.

Um caso particular importante do dBMAP é o dMMPP. Diz-se que o processo (X, J) que existe no espaço de estados $\mathbb{N}_0 \times S$ é um dMMPP com parâmetros $(\mathbf{P}, \mathbf{\Lambda})$, onde $\mathbf{P} = (p_{ij})_{i,j \in S}$ é uma matriz estocástica e $\mathbf{\Lambda} = (\lambda_{ij})_{i,j \in S} = (\lambda_i \delta_{ij})_{i,j \in S}$ é uma matriz diagonal com elementos não negativos (isto é, $\lambda_i \geq 0, i \in S$), se for um dBMAP com parametrização $(\mathbf{P}, \{\mathbf{Q}(n), n \in \mathbb{N}_0\})$, onde

$$q_{ij}(n) = e^{-\lambda_i} \frac{\lambda_i^n}{n!} \quad (4.2)$$

para $i, j \in S$ e $n \in \mathbb{N}$; isto é, $q_{ij} = \{q_{ij}(n), n \in \mathbb{N}_0\}$ é a função de probabilidade de uma variável aleatória de Poisson com média λ_i . Assim, um dMMPP é um dBMAP para o qual o número de chegadas num dado instante depende apenas da fase do dBMAP e quando o processo está na fase i o número e chegadas num dado instante tem uma distribuição de Poisson com média λ_i . O parâmetro λ_i pode ser nulo e, neste caso, não existe qualquer chegada quando o processo está na fase i .

No modelo aqui proposto será usado um dMMPP para modelar o processo de chegada de pacotes. Considera-se também que os tamanhos dos pacotes são independentes, tendo o tamanho dos pacotes que chegam na fase i uma função de probabilidade $q_i = \{q_i(n), n \in \mathbb{N}_0\}$. Deste modo, se (X, J) designar um dMMPP com espaço de estados $\mathbb{N}_0 \times S$ e com parametrização $(\mathbf{P}, \mathbf{\Lambda})$ que modela o processo de chegada de pacotes, o processo de chegadas (Y, J) é um dBMAP com um espaço de estados $\mathbb{N}_0 \times S$, satisfazendo (4.1) com

$$q_{ij}(n) = \sum_{l=0}^{+\infty} e^{-\lambda_i} \frac{\lambda_i^l}{l!} q_i^{(l)}(n) \quad (4.3)$$

para $i, j \in S$ e $n \in \mathbb{N}_0$, onde $q_i^{(l)}$ designa a convolução de ordem l de q_i . $q_{ij}(n)$ representa a probabilidade de após a transição da fase i para j existirem n chegadas (em bytes). Assim a inferência desta probabilidade leva em conta o número de pacotes que chegaram e a distribuição do tamanho destes. A probabilidade de haver l chegadas é dada por $e^{-\lambda_i} \lambda_i^l / l!$ e a probabilidade de l pacotes terem um tamanho total de n bytes é dada por $q_i^{(l)}(n)$; a probabilidade de l chegadas de pacotes corresponderem a n chegadas de bytes é dada por $e^{-\lambda_i} \lambda_i^l q_i^{(l)}(n) / l!$; a probabilidade $q_{ij}(n)$ resulta da soma desta componente para todos os casos possíveis do número de chegadas de pacotes no intervalo de tempo em causa, isto é, desde não haver chegadas ($l = 0$) até um número infinito de chegadas ($l = \infty$). Um dBMAP de 4 estados está ilustrado na Figura 4.1. Assim, (Y, J) é um dBMAP com um espaço de estados $\mathbb{N}_0 \times S$, de tal modo que, para

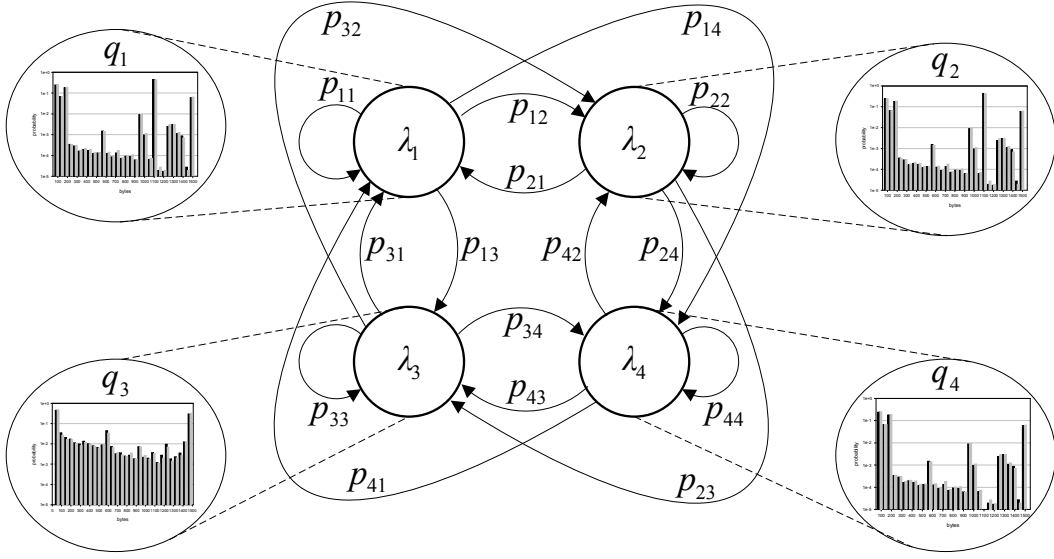


Figura 4.1: Exemplo de um dBMAP com 4 estados.

$n, m \in \mathbb{N}_0$,

$$P(Y_{k+1} = m + n, J_{k+1} = j | Y_k = m, J_k = i) = p_{ij} \sum_{l=0}^{+\infty} e^{-\lambda_i} \frac{\lambda_i^l}{l!} q_i^{(l)}(n). \quad (4.4)$$

Esta parametrização do dBMAP, designada por parametrização de tipo II, é a mais adequada para a subclasse de dBMAPs em questão, porque contém (i) os parâmetros $(\mathbf{P}, \mathbf{\Lambda})$ do dMMPP associado que modela o processo de chegada de pacotes e (ii) a distribuição do tamanho dos pacotes na fase i , q_i , para $i \in S$. Dado um dBMAP (Y, J) com um conjunto de fases S e uma parametrização de tipo II $(\mathbf{P}, \mathbf{\Lambda}, \{q_i, i \in S\})$, escreve-se

$$(Y, J) \sim \text{dBMAP}_S(\mathbf{P}, \mathbf{\Lambda}, \{q_i\}). \quad (4.5)$$

De modo semelhante se (X, J) é um dMMPP com um conjunto de fases S e com um parametrização $(\mathbf{P}, \mathbf{\Lambda})$, escreve-se

$$(X, J) \sim \text{dMMPP}_S(\mathbf{P}, \mathbf{\Lambda}). \quad (4.6)$$

Se S for constituído por r fases, diz-se que (Y, J) ((X, J)) é um dBMAP (dMMPP) de ordem r , r -dBMAP (r -dMMPP). Quando em particular, $S = \{1, 2, \dots, r\}$ para qualquer $r \in \mathbb{N}$, então

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1r} \\ p_{21} & p_{22} & \dots & p_{2r} \\ \dots & \dots & \dots & \dots \\ p_{r1} & p_{r2} & \dots & p_{rr} \end{bmatrix} \quad \text{and} \quad \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_r \end{bmatrix} \quad (4.7)$$

e escreve-se simplesmente $(Y, J) \sim \text{dBMAP}_r(\mathbf{P}, \mathbf{\Lambda}, \{q_i\})$ e $(X, J) \sim \text{dMMPP}_r(\mathbf{P}, \mathbf{\Lambda})$.

4.2.2 Procedimento de inferência de parâmetros

O procedimento de inferência de parâmetros do dBMAP está dividido em dois passos distintos: (i) inferência dos parâmetros do dMMPP subjacente e (ii) ajuste das distribuições do tamanho dos pacotes.

A inferência do dMMPP subjacente ao dBMAP, ou seja a modelação do processo de chegada de pacotes, é feita de acordo com o procedimento descrito na secção 3.3 do capítulo 3. Deste modo, o dMMPP subjacente ao dBMAP será um $M2^L$ -dMMPP.

Quanto à caracterização do tamanho dos pacotes, esta é feita de forma independente para cada um dos estados do $M2^L$ -dMMPP associado. Assim, o procedimento começa por associar a cada intervalo de amostragem (e respectivos pacotes) um dos estados do $M2^L$ -dMMPP, efectuando de seguida a inferência da distribuição do tamanho dos pacotes para cada um desses estados.

De modo a associar cada intervalo de amostragem a um dos estados do $M2^L$ -dMMPP, são varridos todos os intervalos de amostragem dos dados. Um intervalo de amostragem onde chegaram n pacotes é associado aleatoriamente a um estado de acordo com uma função de probabilidade $\theta(n) = \{\theta_1(n), \dots, \theta_{M2^L}(n)\}$, onde $\theta_i(n)$ representa a probabilidade de n chegadas de pacotes observadas terem sido originadas no estado i . Estes valores são dados por

$$\theta_i(n) = \frac{\pi_i g_{\lambda_i}(n)}{\sum_{j=1}^{M2^L} \pi_j g_{\lambda_j}(n)} \quad (4.8)$$

onde λ_j representa a taxa de chegadas de Poisson no j -ésimo estado do $M2^L$ -dMMPP, π_j corresponde à probabilidade em regime estacionário e $g_{\lambda}(n)$ representa a função de probabilidade de uma variável de Poisson com média λ .

Para inferir a distribuição dos tamanhos de pacotes de cada um dos estados, define-se E_i como sendo uma variável aleatória discreta que representa o tamanho dos pacotes do i -ésimo estado do $M2^L$ -dMMPP. Definem-se também $M2^L$ funções massa de probabilidade (FMPs), uma para cada um dos estados do $M2^L$ -dMMPP, dadas por $q_i(n) = P(E_i = n)$, com $n \in \Upsilon^i = \{v_1, v_2, \dots, v_{G_i}\}$ e $i = 1, 2, \dots, M2^L$; Υ^i é o conjunto (aqui também designado por alfabeto) dos G_i tamanhos de pacotes mais frequentes associados com o i -ésimo estado do $M2^L$ -dMMPP.

A inferência da distribuição do tamanho dos pacotes começa por considerar a série temporal de tamanhos de pacotes (individuais) $\{C_k, k = 1, 2, \dots, K\}$, onde C_k é o tamanho do k -ésimo pacote e K o número total de pacotes. É efectuada de seguida

uma partição desta série temporal em $M2^L$ subconjuntos definidos por

$$D^i = \{C_k : S(k) = i, k = 1, 2, \dots, K\}, i = 1, 2, \dots, M2^L, \quad (4.9)$$

onde $S(k)$ representa o estado que foi associado ao k -ésimo pacote, de acordo com o critério atrás descrito que recorre a $\theta(n)$. Assim, o subconjunto D^i agrupa os tamanhos de todos os pacotes que foram associados ao i -ésimo estado do dMMPP associado. Estes subconjuntos serão usados para inferir os conjuntos dos G_i tamanhos de pacotes mais frequentes do estado i (por exemplo usando histogramas), que serão agrupados nos alfabetos $\Upsilon^i = \{v_1, v_2, \dots, v_{G_i}\}, i = 1, 2, \dots, M2^L$. Após isso, cada elemento do subconjunto D^i é arredondado para o elemento mais próximo de Υ^i , resultando assim os novos subconjuntos

$$\tilde{D}^i = \{\Phi_{\Upsilon^i}(x), x \in D^i\}, i = 1, 2, \dots, M2^L \quad (4.10)$$

onde $\Phi_{\Omega}(x)$ representa a função que arredonda x para o elemento mais próximo do vector Ω . Finalmente, a partir dos subconjuntos \tilde{D}^i são inferidas as correspondentes FMP $q_i(n) = P\{E_i = n\}, n \in \Upsilon^i$.

É ainda possível considerar o alfabeto dos pacotes mais frequentes independentemente da fase do processo dBMAP. Este alfabeto será designado por Υ e, na sua construção, são considerados os G tamanhos de pacotes mais frequentes da série $\{C_k, k = 1, 2, \dots, K\}$.

4.3 Resultados

O modelo dBMAP e respectivo procedimento de inferência foram aplicados a duas capturas de tráfego: (i) a captura Bellcore de Outubro (pOct.TL) e (ii) uma captura feita na Universidade de Aveiro (UA); ambas as capturas estão descritas em pormenor no Apêndice A. O intervalo de amostragem usado foi 0.1 segundos em ambas as capturas.

A captura Bellcore foi ajustada a um dBMAP de 40 estados e a captura UA a um dBMAP com 12 estados. Os dMMPPs subjacentes são os descritos na secção 3.4.1.

Tanto no caso da captura Bellcore como no caso da captura UA verificou-se que os alfabetos inferidos eram, em cada estado, praticamente idênticos. Assim optou-se por usar alfabetos independentes da fase do dBMAP (com 20 elementos), tanto mais que isso conduzia a um número de parâmetros significativamente menor. No caso da

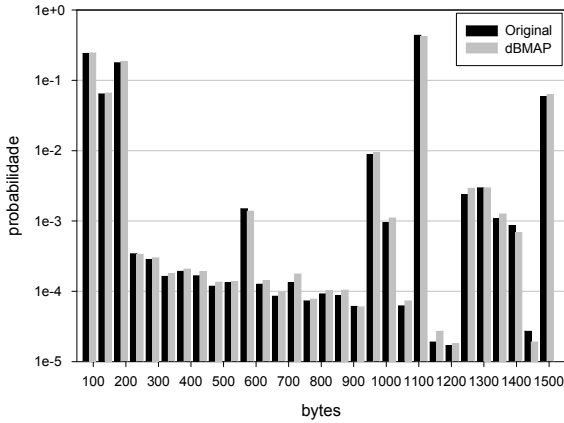


Figura 4.2: Ajuste da função massa de probabilidade, processo TP, Bellcore.

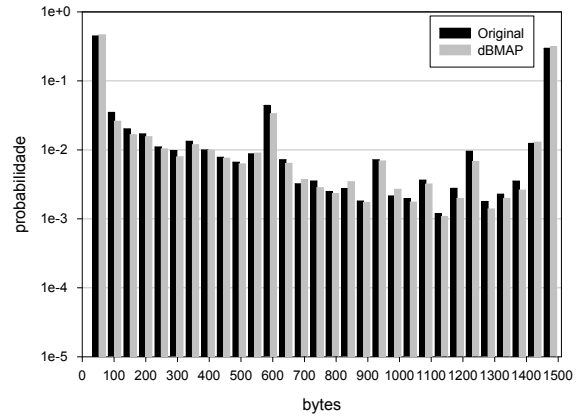


Figura 4.3: Ajuste da função massa de probabilidade, processo TP, UA.

captura Bellcore o alfabeto inferido foi $\Upsilon = \{64\ 74\ 94\ 102\ 110\ 126\ 130\ 142\ 150\ 162\ 174\ 190\ 570\ 938\ 986\ 1082\ 1090\ 1242\ 1282\ 1518\}$ e no caso da captura UA foi $\Upsilon = \{40\ 48\ 52\ 60\ 77\ 114\ 213\ 351\ 386\ 552\ 576\ 608\ 628\ 837\ 932\ 1216\ 1400\ 1420\ 1462\ 1500\}$.

O modelo de tráfego e a precisão dos procedimentos de inferência a ele associados foram avaliados segundo vários critérios. Compararam-se as funções massa de probabilidade (FMP) do processo de chegada em bytes por segundo (BPS) e do processo de tamanhos de pacotes (TP), obtidos com os dados originais e com o modelo dBMAP. Foram também comparadas as funções de autocovariância (FAC) do processo de BPS. Também foi analisado o comportamento do tráfego quando submetido a uma fila de espera, comparando-se o rácio de pacotes perdidos (RPP) e o atraso médio na fila de espera (AMFE) obtidos por simulação, com dois tipos de tráfego: (i) o tráfego original e (ii) tráfego gerado de acordo com o modelo dBMAP inferido. A modelação do processo de chegada de pacotes através de um dMMPP foi avaliada na secção 3.4 do capítulo 3. Conforme foi aí verificado, o modelo $M2^L$ -dMMPP conseguiu resultados muito bons tanto no ajuste das estatísticas de primeira e segunda ordem como na captura do comportamento do tráfego na passagem por uma fila de espera. Nas simulações efectuadas com base no modelo inferido foram calculados intervalos de confiança a 95% com base em 10 réplicas. Uma vez que o comprimento dos intervalos foi sempre muito reduzido optou-se por não os incluir nas figuras.

Nas Figuras 4.2 e 4.3, podem ser observados os resultados do ajuste do processo do tamanho dos pacotes. Para ambas as capturas consideradas, verifica-se que o modelo dBMAP foi capaz de capturar as estatísticas de primeira ordem do processo dos tamanhos de pacotes com um elevado grau de precisão.

No que respeita ao ajuste das estatísticas do processo de chegada de BPS (o qual

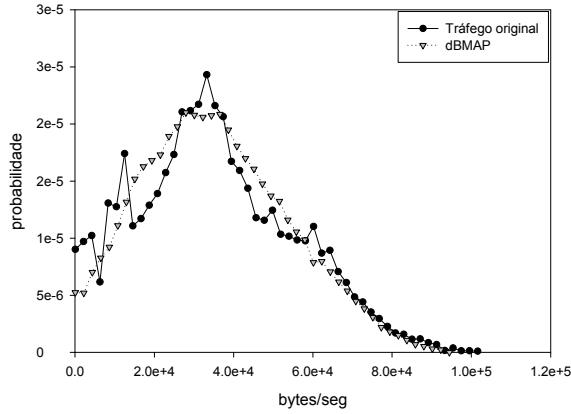


Figura 4.4: Ajuste da função massa de probabilidade, processo BPS, Bellcore.

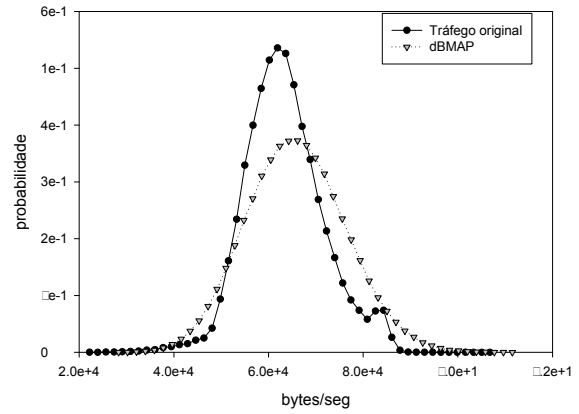


Figura 4.5: Ajuste da função massa de probabilidade, processo BPS, UA.

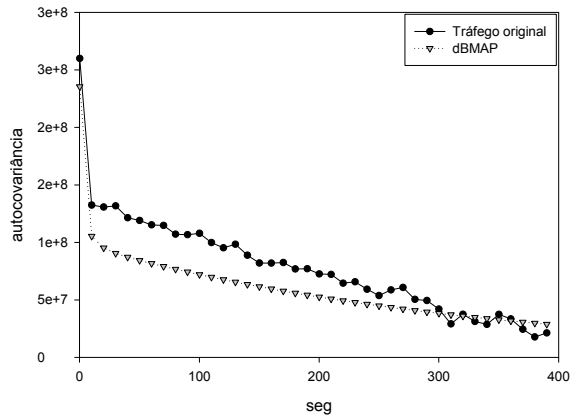


Figura 4.6: Ajuste da função de autocovariância, processo BPS, Bellcore.

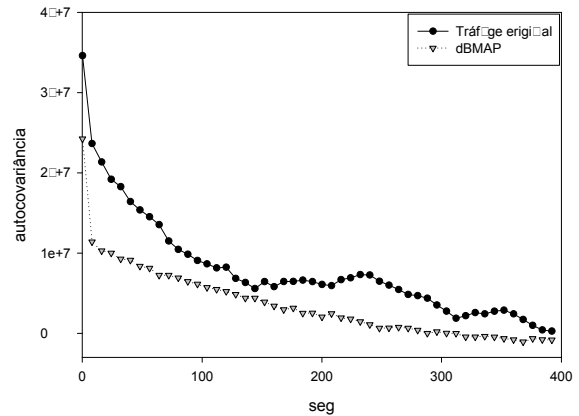


Figura 4.7: Ajuste da função de autocovariância, processo BPS, UA.

resulta da conjugação dos processos de chegada e tamanhos dos pacotes) verifica-se que o procedimento não conseguiu um bom ajuste das mesmas (Figuras 4.4, 4.5, 4.6 e 4.7). Uma vez que o processo de chegada de pacotes foi muito bem ajustado pelo dMMPP (Figuras 3.7, 3.8, 3.9 e 3.10), e que a autocovariância do processo de BPS do tráfego gerado a partir dos modelos dBMAP inferidos é menor que a do tráfego original (Figuras 4.6 e 4.7), esta diferença pode explicar-se pelo facto de o modelo dBMAP não conseguir descrever a autocorrelação do processo do tamanho dos pacotes.

Apesar de um ajuste menos bom das estatísticas de primeira e segunda ordem pode-se observar nas Figuras 4.8, 4.9, 4.10 e 4.11, que o modelo conseguiu capturar o comportamento da captura Bellcore quando submetida a uma fila de espera. No que se refere à captura UA (Figuras 4.12, 4.13, 4.14 e 4.15), o modelo conseguiu igualmente muito bons resultados.

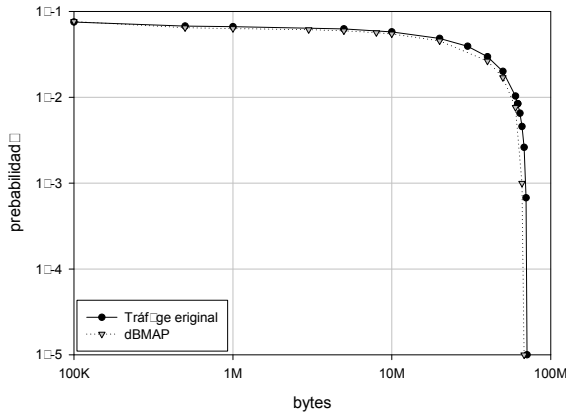


Figura 4.8: Rácio de perda de pacotes com uma utilização de 90%, Bellcore.

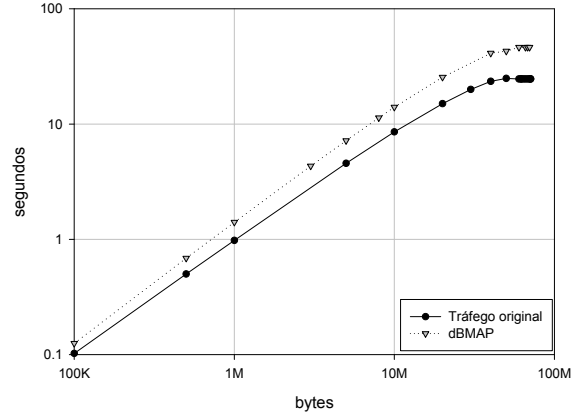


Figura 4.9: Atraso médio na fila de espera com uma utilização de 90%, Bellcore.

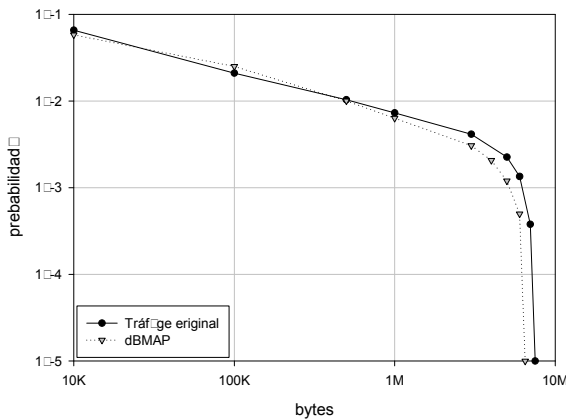


Figura 4.10: Rácio de perda de pacotes com uma utilização de 70%, Bellcore.

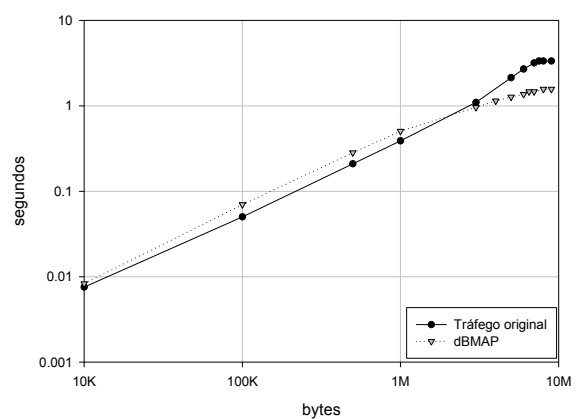


Figura 4.11: Atraso médio na fila de espera com uma utilização de 70%, Bellcore.

4.4 Conclusões

Neste capítulo foi proposto um modelo de tráfego baseado num processo discreto de Markov com chegadas em rajada (dBMAP). No modelo aqui proposto, as chegadas de pacotes ocorrem de acordo com um processo discreto de Poisson modulado à Markov (dMMPP) e cada pacote tem um tamanho definido de acordo com uma distribuição geral que depende da fase do dMMPP subjacente. Foi desenvolvido um procedimento de inferência de parâmetros para este dBMAP que permite o ajuste simultâneo (i) da autocovariância e distribuição marginal do processo de chegada de pacotes e (ii) da distribuição dos tamanhos de pacotes.

O procedimento de inferência de parâmetros foi aplicado a tráfego real com características de memória longa. Como já se havia verificado no capítulo anterior o procedimento ajusta com um elevado grau de precisão as funções de autocovariância e massa de probabilidade do processo de chegada de pacotes. Quanto ao processo do

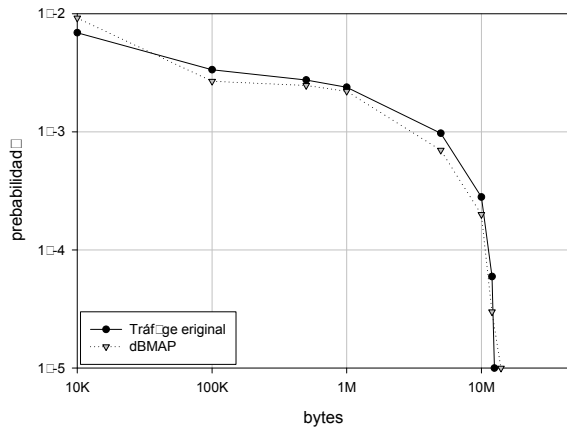


Figura 4.12: Rácio de perda de pacotes com uma utilização de 90%, UA.

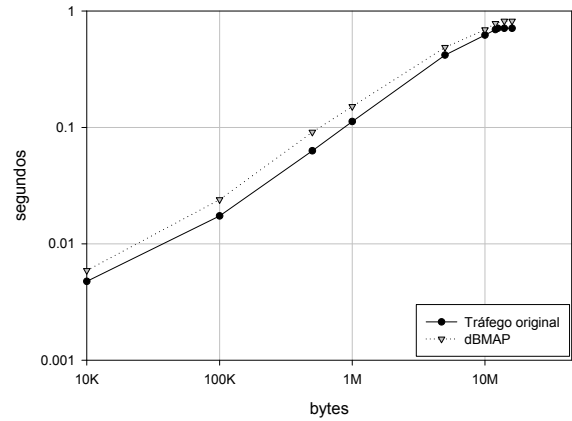


Figura 4.13: Atraso médio na fila de espera com uma utilização de 90%, UA.

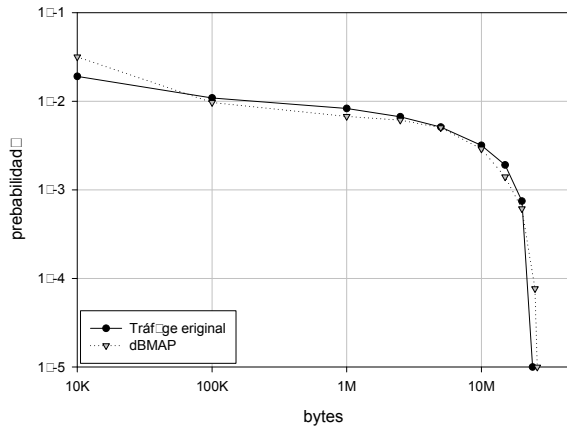


Figura 4.14: Rácio de perda de pacotes com uma utilização de 98%, UA.

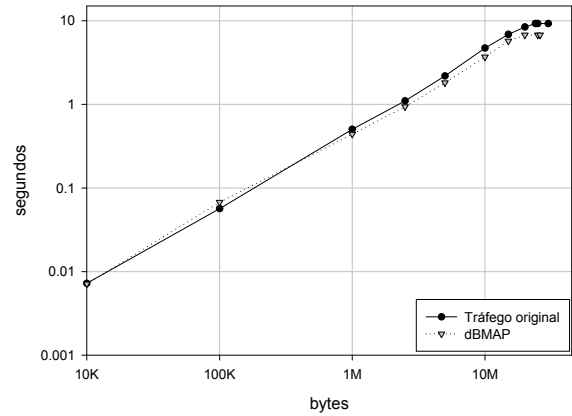


Figura 4.15: Atraso médio na fila de espera com uma utilização de 98%, UA.

tamanho dos pacotes, o procedimento faz um bom ajuste da função de probabilidade. No entanto, o ajuste das estatísticas do processo de bytes por segundo (o qual resulta da conjugação dos processos de chegada e tamanhos dos pacotes), não é muito bom, o que pode ser explicado pelo facto do modelo não capturar a autocorrelação do tamanho dos pacotes. Apesar disto, o modelo dBMAP (e o procedimento de inferência de parâmetros associado) revelou-se um modelo eficaz na captura do comportamento do tráfego em fila de espera.

Capítulo 5

Modelos com base em sistemas de Lindenmayer para o processo de chegadas de pacotes

5.1 Introdução

Neste capítulo será apresentado um modelo de tráfego baseado nos chamados sistemas de Lindenmayer e o procedimento de inferência de parâmetros respectivo. Os sistemas de Lindenmayer foram introduzidos em 1968 pelo biólogo A. Lindenmayer como um método para modelar o crescimento de plantas. A capacidade dos sistemas de Lindenmayer para modelar fenómenos fractais motivou o desenvolvimento de um modelo de tráfego com base nos mesmos. Em particular, usaram-se sistemas de Lindenmayer estocásticos, pois permitem a introdução de componentes aleatórias no tráfego.

Este capítulo está organizado do seguinte modo. Na secção 5.2 faz-se uma introdução aos sistemas de Lindenmayer (sistemas-L). Na secção 5.3 é descrito em pormenor o modelo proposto e o respectivo procedimento de inferência de parâmetros. Na secção 5.4 é feita uma comparação detalhada entre os sistemas-L e as cascatas aleatórias, sendo também apresentado o fundamento físico que suporta os modelos de tráfego baseados em sistemas-L. Na secção 5.5 são apresentados e discutidos os resultados da aplicação do modelo proposto a tráfego real. Finalmente na secção 5.6 são apresentadas algumas conclusões.

5.2 Sistemas de Lindenmayer

Um sistema de Lindenmayer (daqui em diante designado por sistema-L) é uma máquina que opera sobre sequências de símbolos. Estes sistemas foram apresentados em 1968 pelo biólogo A. Lindenmayer como um método para modelar o crescimento de plantas [Lin68]. São caracterizados por um alfabeto, um axioma e um conjunto de regras de produção (ou crescimento). O alfabeto é um conjunto de símbolos e as regras de produção definem transformações de determinado símbolo numa sequência de símbolos. A partir de uma sequência ou símbolo inicial (axioma), um sistema-L constrói iterativamente sequências de símbolos substituindo cada símbolo pela respectiva sequência definida nas regras de produção. Os sistemas-L são um método recursivo de construção de sequências com características fractais. Os sistemas-L dizem-se estocásticos quando nas regras de produção é introduzida uma componente aleatória. Uma descrição detalhada dos sistemas-L pode ser encontrada em [PJS92].

A curva de Koch é um fractal que pode ser construído através de um sistema-L descrito do seguinte modo:

$$\begin{aligned} \text{Alfabeto:} & \quad \{F, +, -\} \\ \text{Axioma:} & \quad F \\ \text{Regra:} & \quad F \rightarrow F+F-F+F \end{aligned}$$

onde 'F' pode ser interpretado como a indicação para desenhar um segmento de recta na direcção corrente, '+' representa uma rotação no sentido anti-horário e '-' uma rotação no sentido horário. Este sistema-L resulta em:

$$\begin{array}{ll} \text{Início} & F \\ 1^{\text{a}} \text{ iteração} & F + F - -F + F \\ 2^{\text{a}} \text{ iteração} & F + F - -F + F + F + F - -F + F - -F + F - -F + F + F + F - -F + F \\ 3^{\text{a}} \text{ iteração} & F + F - -F + F + F + F - -F + F - -F + F - -F + F + F + F - -F + F \\ & + F + F - -F + F + F + F - -F + F - -F + F - -F + F + F + F - -F + \\ & F - -F + F - -F + F + F + F - -F + F - -F + F - -F + F + F + F - \\ & -F + F + F + F - -F + F + F + F - -F + F - -F + F - -F + F + F + F \\ & - -F + F \\ \dots & \dots \end{array}$$

O resultado das primeiras quatro iterações do processo de construção da curva de Koch, para rotações de 60°, está ilustrado na Figura 5.1. Na mesma figura, é possível observar que partes da curva de Koch numa determinada iteração tem similaridades

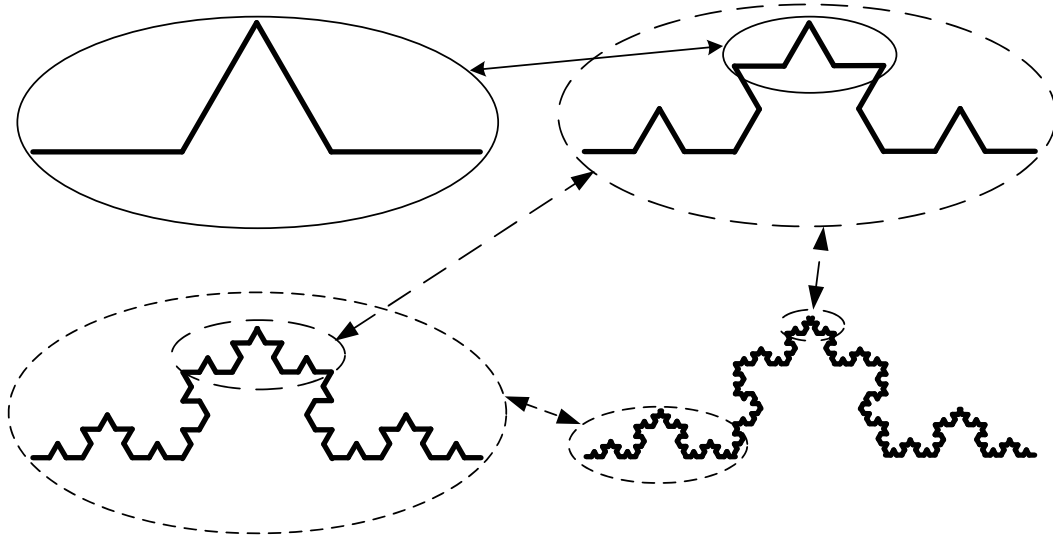


Figura 5.1: Exemplo de um sistema-L e das similaridades entre escalas, curva de Koch.

com as curvas de Koch de iterações anteriores. Note-se que a ordem dos símbolos é importante num sistema-L.

Um outro exemplo de aplicação de um sistemas-L é o crescimento de um organismo por divisão celular. Neste exemplo, existem dois tipos de células representadas pelas letras A e B. A divisão celular é modelada substituindo estas letras por sequência de letras: uma célula A divide-se em duas células (uma do tipo A e outra do tipo B) representadas pela sequência AB; uma célula B divide-se em duas células do tipo A, representadas pela sequência AA. O organismo modelado por este sistema-L cresce efectuando sucessivas divisões celulares. Na altura do nascimento, este organismo é constituído por uma única célula do tipo A. Após uma divisão o organismo tem duas células representadas pela sequência AB. Após a segunda divisão o organismo já possui 4 células (sequência ABAA), e após três divisões o organismo tem oito células (sequência ABAAABAB). Este processo de crescimento pode ser descrito da seguinte forma:

Alfabeto: $\{A,B\}$
 Axioma: A
 Regras: $A \rightarrow AB$
 $B \rightarrow AA$

que resulta em:

Início	A
1ª iteração	AB
2ª iteração	ABAA
3ª iteração	ABAAABAB
4ª iteração	ABAAABABABAAABAA
...	...

Quando as regras de produção são estocásticas o sistema-L diz-se estocástico. Num sistema-L estocástico podem existir várias regras para um único símbolo, sendo uma regra específica aplicada com uma determinada probabilidade. Usando o exemplo anterior, uma regra de produção pode converter A em AB com probabilidade 0.4 ou em BB com probabilidade 0.6 (em vez de converter sempre A em AB). Neste caso, após 3 iterações são várias as sequências possíveis: ABAAABAB, BBAAABAB, BBBB BBBB, etc.

Nas Figuras 5.2 e 5.3 são apresentados alguns resultados da aplicação dos sistemas-L à modelação de plantas. O sistema-L da Figura 5.2, resultou da aplicação da regra $F \rightarrow F[+F]F[-F]F$ ao axioma F , com um ângulo de rotação de 23.5° . Os símbolos '[' e ']' funcionam em conjunto: '[' memoriza o actual estado do gráfico e ']' repõe o estado do gráfico anteriormente memorizado, permitindo o crescimento independente de partes do fractal. Quanto ao sistema-L da Figura 5.3, resultou da aplicação da regra $F \rightarrow FF - [-F + F + F] + [+F - F - F]$ ao axioma F , com um ângulo de rotação de 15° .

Na Figura 5.4 podemos observar a aplicação dos sistemas-L a um objecto mais abstracto, demonstrando assim a versatilidade dos mesmos. Este sistema usa um alfa-

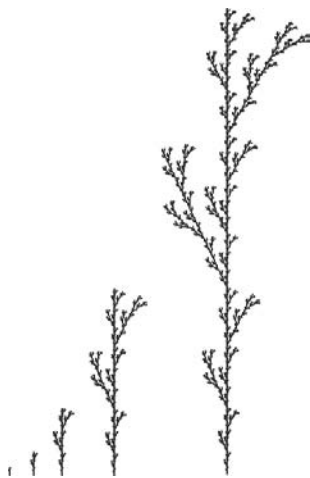


Figura 5.2: Exemplo de sistemas-L aplicado ao crescimento de plantas, erva.

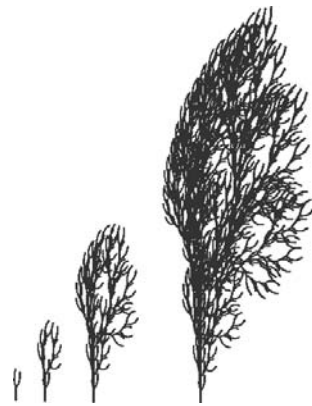


Figura 5.3: Exemplo de sistema-L aplicado ao crescimento de plantas, arbusto.

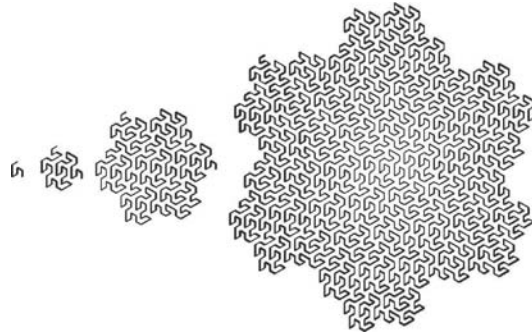


Figura 5.4: Exemplo de sistemas-L, labirinto.



Figura 5.5: Exemplos de sistemas-L estocásticos aplicados ao crescimento de plantas, arbustos.

betto constituído pelos símbolos $\{X, Y, +, -\}$, onde 'X' e 'Y' podem ser interpretados como a indicação para desenhar um segmento de recta na direcção corrente, e tal como nos exemplos anteriores, '+' representa uma rotação no sentido anti-horário e '-' uma rotação no sentido horário. Como 'X' e 'Y' representam ramos do fractal com potencial de evolução diferente, este sistema-L possui duas regras de produção distintas:

$$\begin{aligned} \text{Alfabeto:} & \quad \{X, Y, +, -\} \\ \text{Axioma:} & \quad X \\ \text{Regra 1:} & \quad X \rightarrow X + Y + +Y - X - -XX - Y + \\ \text{Regra 2:} & \quad Y \rightarrow -X + YY + +Y + X - -X - Y \end{aligned}$$

Na Figura 5.5 ilustra-se a utilização de sistemas-L estocásticos. A introdução de uma componente aleatória às regras de produção do sistema-L da Figura 5.3 dá origem, ao fim do mesmo número de iterações, a arbustos distintos mas onde as características fundamentais estão presentes. Este sistema-L é definido por:

$$\begin{aligned} \text{Alfabeto:} & \quad \{F, +, -, [,]\} \\ \text{Axioma:} & \quad F \\ \text{Regra:} & \quad (\text{com probabilidade } 0.5) F \rightarrow FF - [-F + F + F] + [+F - F - F] \\ & \quad (\text{com probabilidade } 0.5) F \rightarrow FF + [-F + F + F] - [+F - F - F] \end{aligned}$$

Outro exemplo de um sistema-L estocástico é o que dá origem a uma sequência de rectângulos com diferentes áreas. Designando um rectângulo de área i por X_i , este sistema pode ser definido por:

Alfabeto:	$\{X_1, X_2, X_3, X_4, X_5\}$
Axioma:	X_3
Regra 1:	$X_1 \rightarrow X_1X_1$
Regra 2:	(com probabilidade 1/3) $X_2 \rightarrow X_1X_3$ (com probabilidade 1/3) $X_2 \rightarrow X_3X_1$ (com probabilidade 1/3) $X_2 \rightarrow X_2X_2$
Regra 3:	(com probabilidade 1/5) $X_3 \rightarrow X_3X_3$ (com probabilidade 1/5) $X_3 \rightarrow X_2X_4$ (com probabilidade 1/5) $X_3 \rightarrow X_4X_2$ (com probabilidade 1/5) $X_3 \rightarrow X_1X_5$ (com probabilidade 1/5) $X_3 \rightarrow X_5X_1$
Regra 4:	(com probabilidade 1/3) $X_4 \rightarrow X_3X_5$ (com probabilidade 1/3) $X_4 \rightarrow X_5X_3$ (com probabilidade 1/3) $X_4 \rightarrow X_4X_4$
Regra 5:	$X_5 \rightarrow X_5X_5$

As regras de produção foram definidas de modo que a média da área dos rectângulos produzidos seja igual à área do rectângulo que lhes deu origem. Se a noção de massa for aplicada à área média dos rectângulos, então estamos perante um sistema-L que se diz ter a propriedade da conservação de massa. Algumas das realizações possíveis deste sistema-L podem ser observadas na Figura 5.6. Uma extrapolação possível deste sistema-L é a associação da área de cada rectângulo a uma quantidade de tráfego observado num intervalo de tempo (por exemplo número de pacotes ou número de bytes). Estamos assim perante um sistema-L estocástico capaz de modelar uma sequência de tráfego.

5.3 Modelo de tráfego para o ritmo de chegada de pacotes

Nesta secção é proposto um modelo de tráfego baseado em sistemas-L estocásticos, para a caracterização do processo do ritmo de chegada de pacotes. Este modelo será

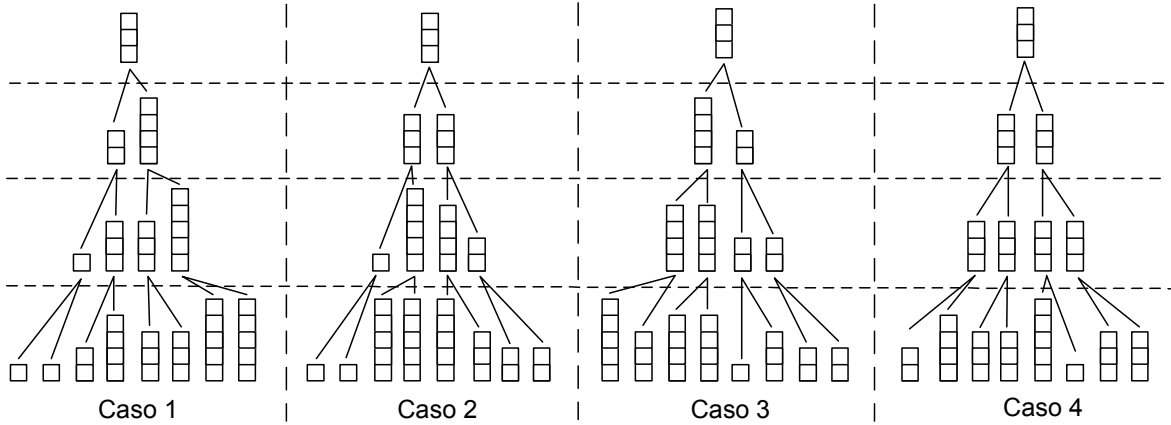


Figura 5.6: Exemplos de construção de sistemas-L estocásticos, retângulos.

designado por "modelo de sistema-L único" por forma a diferenciá-lo de outros modelos baseados em sistema-L que serão introduzidos no capítulo seguinte. O sistema-L deste modelo tem um alfabeto de ritmos de chegada, λ_i , definido por

$$\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_L\}, \lambda_i \in \mathbb{R}_0^+, i = 1, \dots, L. \quad (5.1)$$

e tem regras de produção que aleatoriamente geram dois ritmos de chegada a partir de outro. Sem perda de generalidade é assumido que $\lambda_1 < \lambda_2 < \dots < \lambda_L$.

O processo de tráfego é construído progressivamente, regido por um sistema-L, onde em cada iteração é produzida uma nova escala temporal. Partindo da escala temporal superior, onde o tráfego é caracterizado por um único ritmo de chegada num único intervalo de tempo, cada iteração gera uma escala mais fina (i) dividindo cada intervalo de tempo (pai) em dois subintervalos de igual duração (filhos) e (ii) associando a cada um destes novos subintervalos um ritmo de chegada de acordo com as regras de produção do sistema-L subjacente. As escalas temporais são agrupadas em gamas de escalas temporais, permitindo-se a definição de regras de produção diferentes para cada uma destas gamas de escalas. A introdução desta característica visa possibilitar a modelação de características de similaridade escalar distintas. Haverá similaridade escalar do tráfego se a representação em escala logarítmica das energias de ordem q (normalmente a energia é de segunda ordem, $q = 2$) em função da escala tiver um comportamento linear (ver secção 2.3.4.2 do capítulo 2); se a representação gráfica mostrar um comportamento globalmente não-linear, ainda assim podem ser detectadas gamas temporais com um comportamento linear. A construção do processo de tráfego está ilustrada na Figura 5.7.

Para caracterizar o processo de tráfego define-se $X_{(j,r)}^{(i)} \in \Lambda$ como sendo o ritmo de

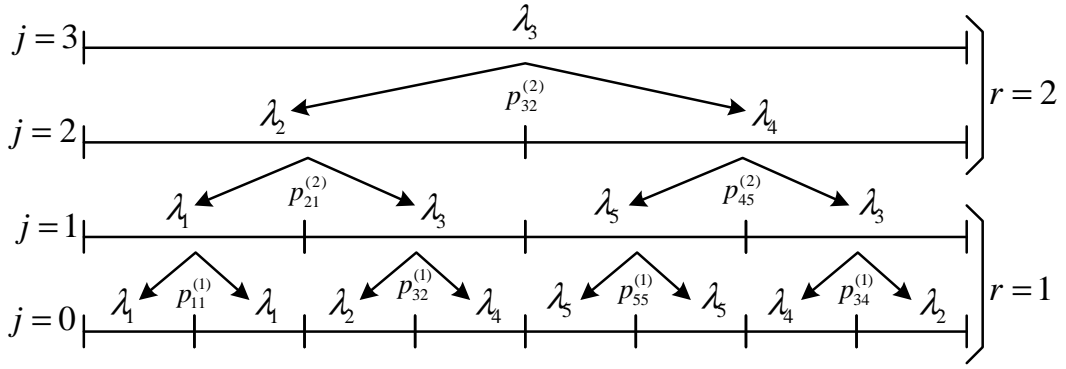


Figura 5.7: Representação da evolução do modelo de sistema-L único.

chegadas no intervalo temporal i da escala temporal j e gama temporal r . O número de escalas temporais é definido por S e o número de gamas temporais por R . Por conveniência, faz-se j decrescer de $j = S - 1$ (na escala temporal superior) até $j = 0$ (escala temporal inferior). Também, r decresce de $r = R$ (a gama temporal das escalas temporais superiores) até $r = 1$ (a gama temporal das escalas temporais inferiores). Então, o número de intervalos temporais na escala temporal j , que será designado por N_j , é 2^{S-j-1} . Assumindo uma duração unitária para os intervalos na escala temporal inferior, $j = 0$, a duração de um intervalo na escala j será 2^j . De maneira a relacionar as escalas temporais com as gamas de escalas temporais, define-se j_r como sendo a escala superior j na gama r . Então para o caso ilustrado na Figura 5.7, teremos $S = 4$, $R = 2$, $j_2 = 3$ e $j_1 = 1$.

De maneira a garantir que o ritmo de chegadas médio é o mesmo em todas as escalas temporais, e para manter o significado físico, é imposta a seguinte condição às regras de produção:

$$X_{(j,r)}^{(i)} = \frac{1}{2}X_{(j-1,r')}^{(2i-1)} + \frac{1}{2}X_{(j-1,r'')}^{(2i)} \quad (5.2)$$

isto é, a atribuição dos ritmos de chegada é tal que a média dos ritmos de chegada nos subintervalos filhos, da direita e da esquerda, será igual ao ritmo de chegada do intervalo pai. Com esta condição, o processo de tráfego pode ser descrito pelo axioma $X_{(S-1,R)}^{(1)}$, o ritmo de chegadas na escala temporal superior, e pela regras de produção definidas por

$$X_{(j,r)}^{(i)} = \lambda_l \xrightarrow{p_{lq}^{(r)}} \begin{cases} X_{(j-1,r')}^{(2i-1)} = \lambda_q \\ X_{(j-1,r'')}^{(2i)} = 2\lambda_l - \lambda_q \end{cases} \quad (5.3)$$

onde $\sum_{q=1}^L p_{lq}^{(r)} = 1, \forall l$. Assim, um ritmo de chegada λ_l no intervalo i , escala j e gama

r produz, com probabilidade $p_{lq}^{(r)}$, um ritmo de chegada λ_q no subintervalo filho da esquerda $2i - 1$ e ritmo de chegada $2\lambda_l - \lambda_q$ no subintervalo filho da direita $2i$, da próxima escala temporal $j - 1$ e gama r' . As regras de produção podem ser totalmente descritas por R matrizes com dimensão $L \times L$,

$$\mathbf{P}^{(r)} = \left(p_{lq}^{(r)} \right), \quad l, q = 1, \dots, L, \quad r = 1, \dots, R. \quad (5.4)$$

De maneira a garantir que o alfabeto é fechado no que respeita às regras de produção, são impostas as seguintes condições: (i) $\lambda_i - \lambda_{i-1} = \frac{\lambda_L - \lambda_1}{L-1}$, $i = 2, 3, \dots, L$, isto é, os valores λ_i serão equidistantes; (ii) $p_{lq}^{(r)} = 0$ se $q > l + \min(L - l, l - 1)$ ou $q < l - \min(L - l, l - 1)$.

Finalmente, o modelo de sistema-L único constrói, na escala j e gama r , a sequência de ritmos de chegada

$$Z_{(j,r)} = \{X_{(j,r)}^{(i)}, i = 1, \dots, N_j\}. \quad (5.5)$$

A partir da sequência $Z_{(0,1)}$ é possível gerar os instantes específicos de chegada dos pacotes, por exemplo, distribuindo uniformemente os $X_{(0,1)}^{(i)}$ pacotes ao longo do i -ésimo intervalo de tempo.

5.3.1 Procedimento de inferência do modelo de sistema-L único

O procedimento de inferência determina os parâmetros do sistema-L a partir de dados reais. Começa-se por definir o intervalo de amostragem do tráfego Δ e considerar a sequência temporal do número de chegada de pacotes em cada um dos intervalos de amostragem. Esta sequência temporal empírica será designada por $\{A_k, k = 1, 2, \dots, K\}$, onde A_k representa o número de chegadas no intervalo de amostragem k . Por conveniência, o comprimento da sequência temporal K deve ser uma potência de 2. O procedimento de inferência pode ser dividido em três passos: (i) determinação do alfabeto e axioma do sistema-L, (ii) identificação das gamas de escalas temporais e (iii) inferência da regras de produção do sistema-L. O diagrama de fluxo deste procedimento está representado na Figura 5.8.

O alfabeto do sistema-L é formado por L valores equidistantes de ritmos de chegada. Existem várias possibilidades para determinar os elementos do alfabeto. Uma hipótese é forçar a inclusão do valor mínimo e do valor máximo dos ritmos de chegada observados. Outra é forçar a inclusão do valor médio e garantir que os valores mínimo

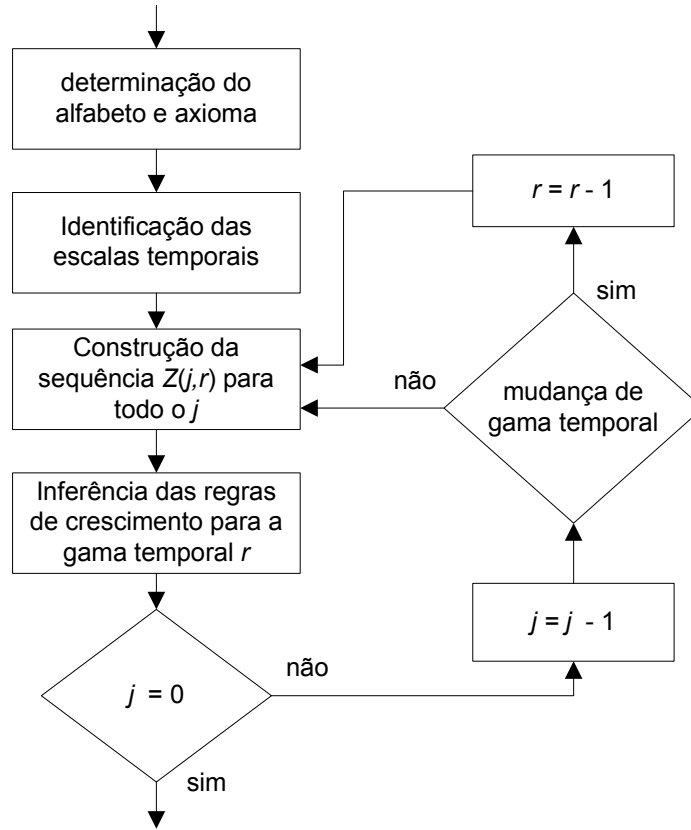


Figura 5.8: Diagrama de fluxo do procedimento de inferência do modelo de sistem-L único.

e máximo estão contidos na gama de valores cobertos pelo alfabeto. Nos resultados apresentados neste capítulo foi considerada a primeira opção. Note-se, que para valores de L relativamente elevados, os alfabetos e consequentemente os resultados obtidos com os dois métodos de construção são semelhantes.

Por forma a garantir que o ritmo médio de chegadas do modelo inferido é igual ao ritmo médio dos dados, impõe-se que o axioma seja igual ao ritmo médio de chegadas de $\{A_k\}$ arredondado para o elemento mais próximo do alfabeto, isto é,

$$X_{(S-1,R)}^{(1)} = \Phi_{\Lambda} \left((1/K\Delta) \sum_{k=1}^K A_k \right) \quad (5.6)$$

onde $\Phi_{\Lambda}(x)$ representa a função que arredonda x para o elemento de Λ mais próximo. A identificação das gamas temporais baseia-se no método de análise escalar por ondulas descrito em [AFTV00] e na secção 2.3.4.2 do capítulo 2 deste Tese, que recorre a diagramas de energias em escala logarítmica. As gamas temporais correspondem ao conjunto de escalas temporais para as quais, dentro dos limites dos intervalos de confiança, os valores de y_j estão sobre uma recta.

O passo final é a inferência das regras de produção do sistema-L, as quais são total-

mente caracterizadas pelas matrizes $\mathbf{P}^{(r)}$. Primeiro as seqüências de ritmos de chegada são arredondados para os elementos mais próximo do alfabeto, definindo-se assim as seqüências $Y_{(j,r)}$ para cada uma das escalas temporais. Isto é feito determinando os ritmos de chegada $X_{(j,r)}^{(i)}$ a partir de $\{A_k\}$ da seguinte forma

$$X_{(j,r)}^{(i)} = \Phi_{\Lambda} \left((N_j / K \Delta) \sum_{k=K(i-1)/N_j+1}^{Ki/N_j} A_k \right) \quad (5.7)$$

com $i = 1, \dots, N_j$, para cada j . Designando por $c_{lq}^{(r)}$ o número de vezes que, na escala temporal j , o pai $X_{(j,r)}^{(i)} = \lambda_l$ gerou um filho (no subintervalo da esquerda) $X_{(j-1,r')}^{(2i-1)} = \lambda_q$, as regras de produção são inferidas através de

$$p_{lq}^{(r)} = c_{lq}^{(r)} / \sum_{q=1}^L c_{lq}^{(r)}, \quad l = 1, \dots, L, \quad r = 1, \dots, R. \quad (5.8)$$

5.4 Relação com as cascatas conservativas

Os modelos com base em cascatas aleatórias conservativas, descritos na secção 2.4.8 do capítulo 2, são modelos multifractais que têm um processo de construção semelhante ao dos sistemas-L.

Nos sistemas-L para o processos de chegada de pacotes, a massa no intervalo i da escala j pode ser interpretada com o sendo $2^j X_{(j,r_a)}^{(i)}$. Para além da propriedade de preservação de massa, os sistemas-L possuem outra característica que não existe nas cascatas conservativas (nem nas cascatas aleatórias) e que permite capturar uma importante característica das redes de dados. Nos sistemas-L a massa é redistribuída em subintervalos (esquerdo e direito), sendo esta redistribuição feita de um modo dependente da massa do intervalo pai. Na construção das cascatas conservativas esta dependência não existe. Consideremos o exemplo, apresentado em [FGW98], da dinâmica de uma sessão HTTP: os cliques de um utilizador geram pedidos HTTP, os pedidos HTTP geram fluxos e os fluxos são formados por pacotes individuais. Ora, o número de pedidos HTTP por unidade de tempo que um servidor pode processar é necessariamente limitado. Deste modo, a maneira como os cliques de um utilizador geram pedidos depende do próprio número de cliques e, portanto, um servidor HTTP que tenha mais pedidos para processar vai distribuir as respostas a esses pedidos mais espaçadamente no tempo. Conclui-se então que a massa (neste caso o número de cliques) influencia o modo como a distribuição temporal das respostas aos pedidos é feita. A um nível protocolar mais baixo, consideremos a maneira como os fluxos produzem pacotes. Se

a rede estiver congestionada, o mecanismo de controlo de congestionamento do TCP impõe que os pacotes sejam distribuídos no tempo de maneira mais esparsa. Mais uma vez, a massa (neste caso agora, o número de fluxos) influencia a maneira como os pacotes são distribuídos no tempo.

Para efectuar uma análise mais precisa deste aspecto, avaliou-se se a propriedade de independência do gerador subjacente às cascatas conservativas estaria ou não presente em dados reais, usando para tal a captura Bellcore. O gerador, tal como definido no contexto das cascatas conservativas (capítulo 2 secção 2.4.8), representa a fracção de massa distribuída ao intervalo filho da esquerda. No sistema-L único esta fracção é $2^{j-1}X_{(j-1,r')}^{(2i-1)}/2^jX_{(j,r)}^{(i)} = \lambda_l/2\lambda_q$. Definiu-se $W_{a_{jl}}$ como sendo o gerador da sequência de ritmos de chegada, na escala j , condicionado pelo ritmo de chegada λ_l . O gerador $W_{a_{jl}}$ é uma variável aleatória discreta que, para cada l , pode tomar valores $\lambda_l/2\lambda_q$ e pode ser facilmente inferido a partir dos dados. Claramente, se o pressuposto de independência for verdadeiro então $W_{a_{jl}}$ deve ter a mesma distribuição para todo o l . Nas Figuras 5.9 e 5.10 está representada a média e variância de $W_{a_{jl}}$, para cada λ_l e para $j = 1, 2, 3, 4$ e 5 . Os resultados mostram que, para uma mesma escala temporal, o gerador pode ter diferentes valores de média e variância. Por exemplo, para $j = 1$ o gerador $W_{a_{jl}}$ tem uma variância de 0.04 para $\lambda_l = 21$ pacotes/seg e 0.001 para $\lambda_l = 100$ pacotes/seg. Verifica-se assim que o gerador possui distribuições distintas o que mostra que o pressuposto de independência não é válido no caso da captura Bellcore. O mesmo comportamento foi observado em outras capturas de tráfego. Para concluir, o efeito imposto pela limitação dos recursos disponíveis é um factor importante que tem um forte impacto no processo de geração de tráfego. Este efeito pode ser capturado por uma descrição baseada em sistemas-L mas não através de cascatas conservativas.

5.5 Resultados numéricos

O modelo e respectivo procedimento de inferência foi aplicado a duas capturas de tráfego: (i) a captura Bellcore de Outubro e (ii) uma captura feita na Universidade de Aveiro (UA) a qual exhibe características de similaridade escalar multifractal não trivial. Ambos as capturas estão descritas em pormenor no Apêndice A. O intervalo de amostragem usado foi 0.1 segundos em ambas as capturas.

A captura Bellcore foi modelada por um sistema-L estocástico com um alfabeto de $L = 236$ ritmos de chegada entre um mínimo de 10 pacotes/seg e um máximo de

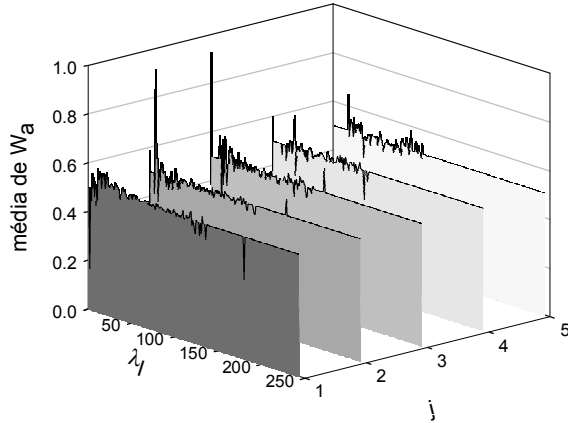


Figura 5.9: Média de $W_{a_{j|l}}$ em função de λ_l e j , ritmo de chegada de pacotes da captura Bellcore.

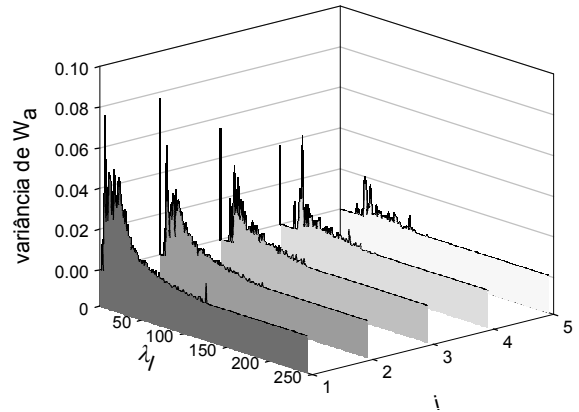


Figura 5.10: Variância de $W_{a_{j|l}}$ em função de λ_l e j , ritmo de chegada de pacotes da captura Bellcore.

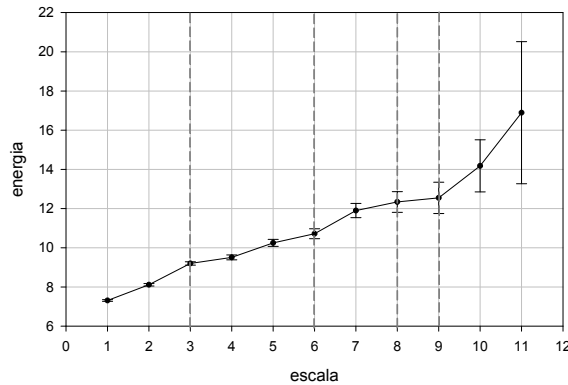


Figura 5.11: diagrama de energias em escala logarítmica, Bellcore.

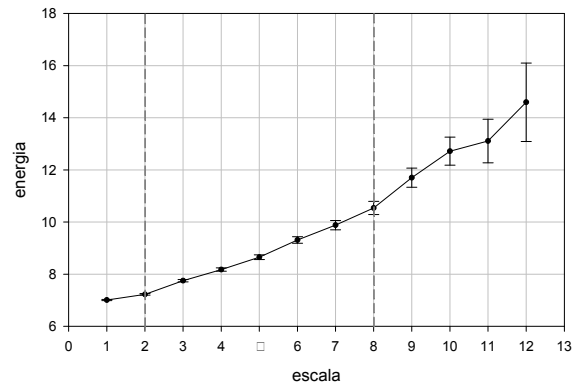


Figura 5.12: diagrama de energias em escala logarítmica, UA.

2360 pacotes/seg, em intervalos de 10 pacotes/seg. De referir que o valor do intervalo entre elementos do alfabeto foi escolhido de modo a que a diferença em número de chegadas por intervalo de amostragem seja um. Através do diagrama de energias em escala logarítmica identificaram-se 5 gamas de escalas temporais (de um total de 14 escalas temporais) definidas por $j_1 = 3, j_2 = 6, j_3 = 8, j_4 = 9$ e $j_5 = 14$ (Figura 5.11). A captura UA foi modelada por um sistema-L com um alfabeto de $L = 214$ ritmos de chegada, de 450 pacotes/seg a 2580 pacotes/seg em intervalos de 10 pacotes/seg. Na Figura 5.12 está representado o diagrama de energias em escala logarítmica da captura UA, sendo possível verificar que existem 3 gamas de escalas (para um total de 17 escalas temporais) definidas por $j_1 = 2, j_2 = 8$ e $j_3 = 17$. A estimação de parâmetros levou menos de 30 segundos, usando uma implementação em MATLAB a correr num PC equipado com um processador AMD Thunderbird 1.2 GHz e 1.5 GBytes de RAM. Isto mostra que o procedimento de inferência é computacionalmente eficiente.

Para efeitos de comparação, o tráfego foi ainda ajustado a uma cascata conservativa

utilizando dois métodos de inferência de parâmetros distintos. No primeiro método, proposto em [GWF99], o gerador é, em cada escala temporal, ajustado a uma distribuição gaussiana truncada, com média $1/2$, sendo a variância ajustada individualmente em cada escala. A cascata conservativa inferida segundo este método será designada por "cascata gaussiana". O segundo método recorre ao uso de um gerador com uma distribuição discreta. Também neste método o gerador é distinto em cada escala temporal, mas agora não existe qualquer restrição no que se refere à média. Este procedimento de ajuste possui um maior número de parâmetros ajustáveis (não se está restrito ao ajuste da variância) introduzindo assim um maior número de graus de liberdade na modelação do tráfego. Este modelo será designado por "cascata discreta".

A aplicabilidade do modelo de tráfego e a precisão do procedimento de inferência a ele associado foram avaliados segundo vários critérios. Compararam-se as funções massa de probabilidade (FMP) e de autocovariância do processo de contagem de pacotes obtido com os modelos inferidos e com os dados originais. Foi também analisado o comportamento do tráfego quando sujeito a uma fila de espera, em termos de rácio de pacotes perdidos e de atraso médio na fila de espera, estimados por simulação. Foram usados quatro tipos de tráfego: (i) o tráfego original considerando que todos os pacotes tinham tamanho fixo igual ao tamanho médio, (ii) tráfego gerado de acordo com o modelo de sistema-L único estocástico, (iii) tráfego gerado de acordo com uma cascata gaussiana e (iv) tráfego gerado de acordo com uma cascata discreta. Na análise do comportamento do tráfego em fila de espera foi ainda usado o tráfego original mas considerando os tamanhos reais dos pacotes.

A características multifractais do tráfego foram avaliadas usando diagramas lineares multiescalares, introduzidos na secção 2.3.5 do capítulo 2. A Figura 5.13 mostra que a captura UA, e o tráfego gerado de acordo com os correspondentes modelos de sistema-L único e de cascatas conservativas, têm todas características de similaridade escalar multifractal. Uma análise semelhante foi efectuada à captura Bellcore revelando que este tráfego não possui características de similaridade escalar multifractal (Figura 5.14). Em ambos os casos, verifica-se que o modelo baseado em sistemas-L forneceu uma boa aproximação das curvas de h_q em função dos momentos, mas o mesmo não aconteceu com as cascatas conservativas (note-se a semelhança das curvas relativas aos dados originais e ao modelo de sistema-L único).

No caso da FMP (Figura 5.15), tanto as cascatas conservativas com o modelo de sistema-L único obtiveram óptimos resultados para a captura UA. O desempenho dos

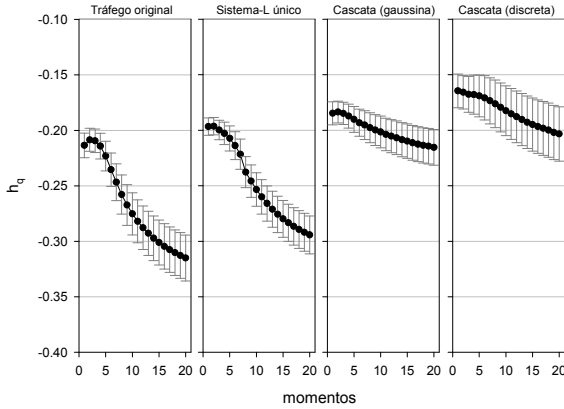


Figura 5.13: Diagramas lineares multiscaleares, UA.

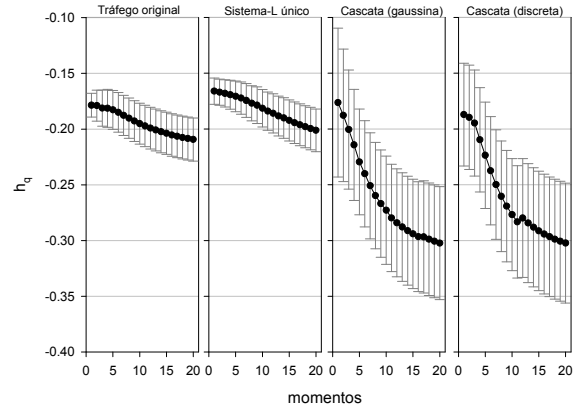


Figura 5.14: Diagramas lineares multiscaleares, Bellcore.

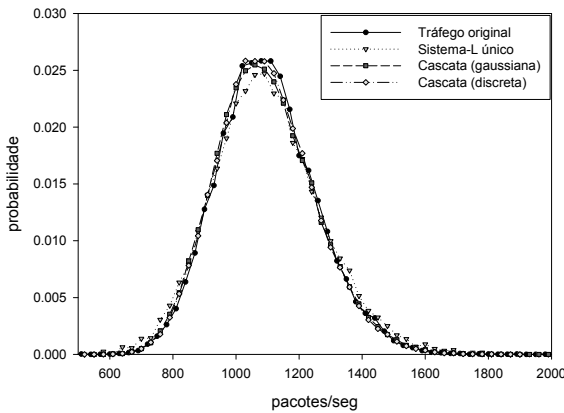


Figura 5.15: Ajuste da função massa de probabilidade, UA.

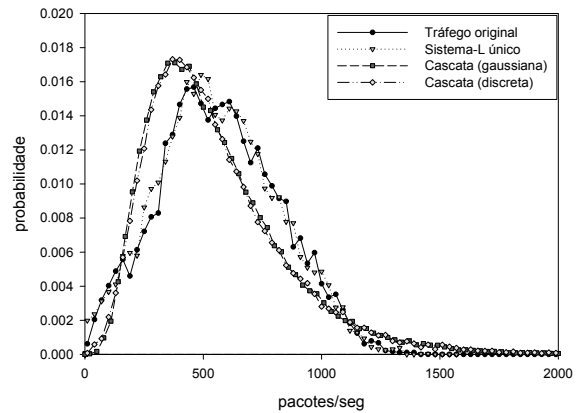


Figura 5.16: Ajuste da função massa de probabilidade, Bellcore.

modelos com base em cascatas conservativas não foi tão bom para a captura Bellcore (Figura 5.16). Este facto pode explicar-se pela ausência de dependência entre escalas do gerador W , o que dificulta o ajuste de distribuições de probabilidade assimétricas por parte das cascatas conservativas. A dependência existente na construção dos sistemas-L permite, por exemplo, fazer com que um pai com uma massa elevada dê sempre origem a uma massa elevada num dos seus filhos, permitindo assim que na escala temporal inferior apareçam mais valores de um dos extremos do alfabeto. É assim possível a obtenção de distribuições de probabilidade assimétricas. Para a função de autocovariância (Figura 5.17), o ajuste conseguido por ambos os modelos foi mais uma vez bastante bom no caso da captura UA, mas claramente inferior para as cascatas conservativas no caso da captura Bellcore (Figura 5.18).

De modo a avaliar o comportamento da captura Bellcore quando sujeita a uma fila de espera fez-se variar a unidade de armazenamento da fila de espera de 10 Kbytes até 8 Mbytes para uma taxa de serviço de 518 Kbytes/s (que corresponde a um factor de

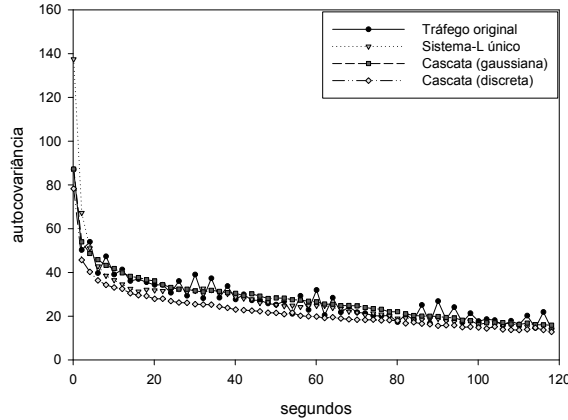


Figura 5.17: Ajuste da função de autocovariância, UA.

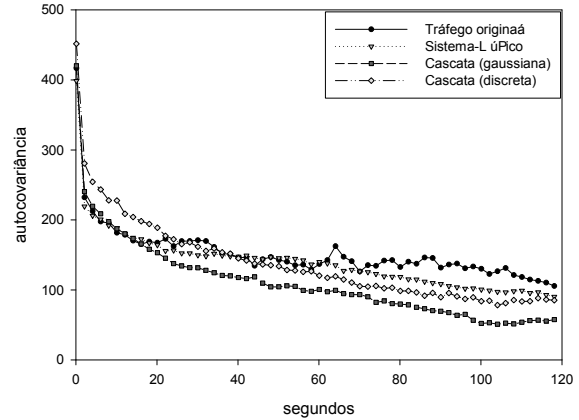


Figura 5.18: Ajuste da função de autocovariância, Bellcore.

utilização de 0.7). Fez-se ainda variar a unidade de armazenamento da fila de espera de 100 Kbytes até 70 Mbytes para uma taxa de serviço de 403 Kbytes/s (que corresponde a um factor de utilização de 0.9).

Para a captura UA utilizaram-se factores de utilização de 0.9 e 0.98, os quais correspondem respectivamente a taxas de serviço de 726 Kbytes/s e 666 Kbytes/s. Em ambos os casos fez-se variar a unidade de armazenamento da fila de espera de 10 Kbytes até 40 Mbytes.

As Figuras 5.19, 5.20, 5.21, 5.22, 5.23, 5.24, 5.25 e 5.26 mostram que, para ambos os tráfegos e factores de utilização, os resultados para o comportamento do tráfego na passagem por uma fila de espera foi muito bom no caso do modelo se sistema-L único, enquanto que para o modelo de cascatas conservativas existiram discrepâncias. É interessante verificar que, para o caso da captura UA, estas discrepâncias ocorreram apesar de um bom ajuste das estatísticas de primeira e segunda ordem. Este facto mostra que para caracterizar tráfego multifractal não é suficiente o ajuste das estatísticas de primeira e segunda ordem. Verifica-se ainda que um modelo de tráfego que considere todos os pacotes com tamanhos iguais não consegue modelar o comportamento de tráfego onde o tamanho dos pacotes é variável. O processo do tamanho dos pacotes tem uma influência fundamental no comportamento do tráfego numa rede, e não pode por isso ser negligenciado.

5.6 Conclusões

Neste capítulo foram apresentados os sistemas de Lindenmayer, um modelo de tráfego baseado nos mesmos e o procedimento de inferência de parâmetros respectivo. Os sis-

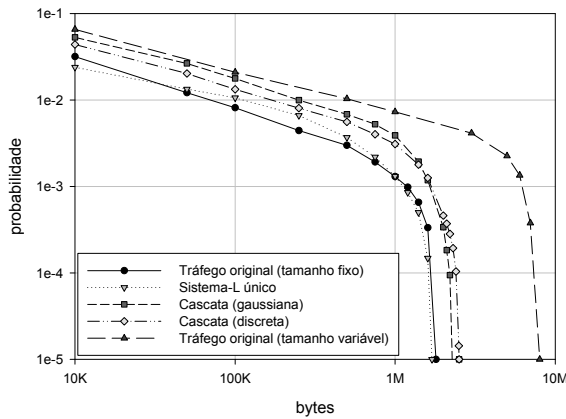


Figura 5.19: Rácio de perda de pacotes com uma utilização de 70%, Bellcore.

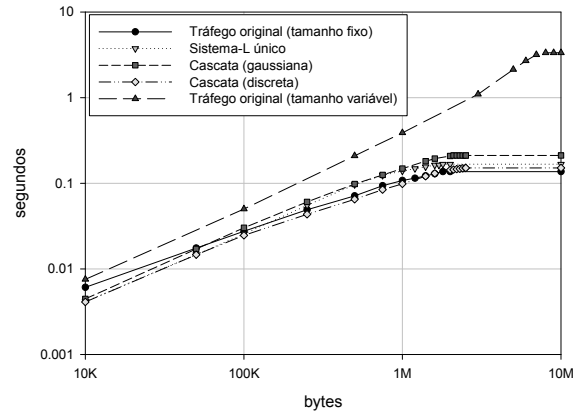


Figura 5.20: Atraso médio na fila de espera com uma utilização de 70%, Bellcore.

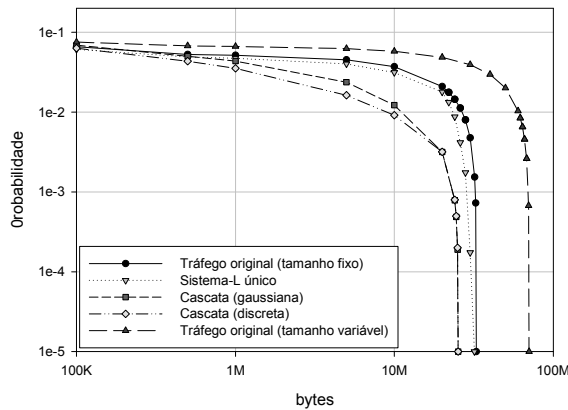


Figura 5.21: Rácio de perda de pacotes com uma utilização de 90%, Bellcore.

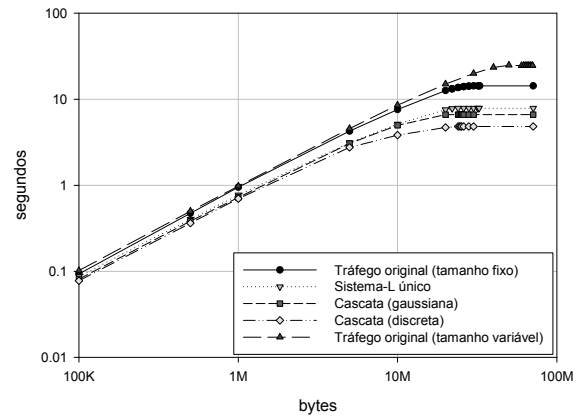


Figura 5.22: Atraso médio na fila de espera com uma utilização de 90%, Bellcore.

temas de Lindenmayer foram propostos como um método para modelar o crescimento de plantas, mas a capacidade destes para modelar fenómenos fractais motivou o desenvolvimento de um modelo de tráfego com base nos mesmos. Usaram-se sistemas de Lindenmayer estocásticos, os quais permitem a introdução de componentes aleatórias no sistema. Uma extensão destes mesmos sistemas em que as características do sistema variam ao longo das diferentes escalas temporais permite que o modelo proposto descreva as características multiescalares do tráfego.

O modelo, designado por modelo de sistema-L único, foi aplicado a tráfego real exibindo características de similaridade escalar multifractal não trivial. Compararam-se as estatísticas de primeira e segunda ordem, o comportamento do tráfego na passagem por uma fila de espera e as características multifractais do tráfego, dos processos de chegada de pacotes do tráfego real, do tráfego gerado segundo o modelo de sistema-L único e ainda de tráfego gerado segundo modelos baseados em cascatas conservativas. O modelo de sistema-L único obteve resultados muito bons em todas as componentes

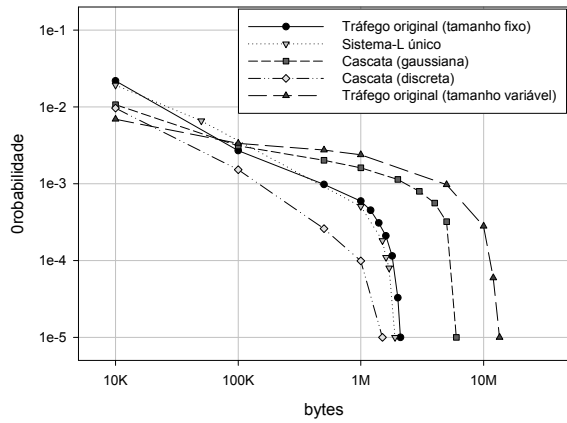


Figura 5.23: Rácio de perda de pacotes com uma utilização de 90%, UA.

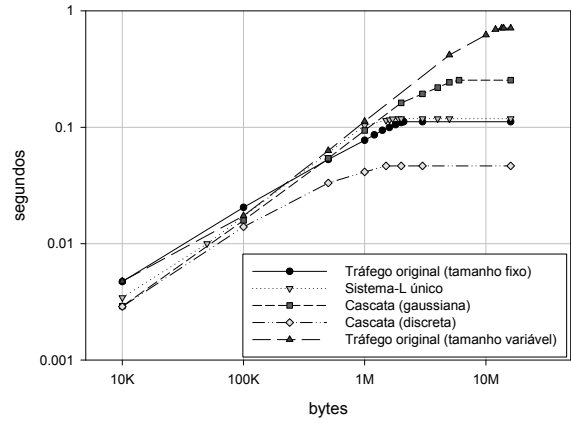


Figura 5.24: Atraso médio na fila de espera com uma utilização de 90%, UA.

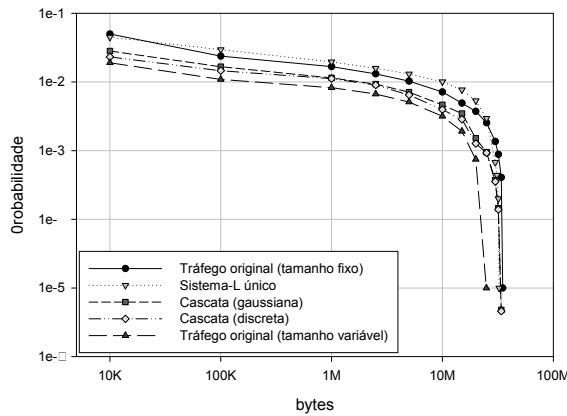


Figura 5.25: Rácio de perda de pacotes com uma utilização de 98%, UA.

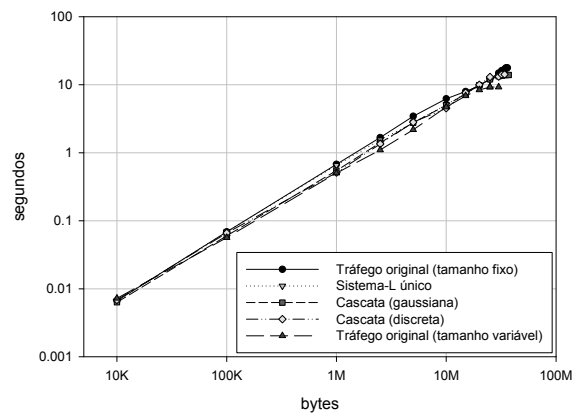


Figura 5.26: Atraso médio na fila de espera com uma utilização de 98%, UA.

avaliadas, sendo o desempenho deste modelo muito superior ao dos modelos de tráfego baseados em cascatas. No entanto quando se consideram os tamanhos reais dos pacotes, verifica-se que o modelo de sistema-L único não consegue descrever eficazmente o comportamento em fila de espera. Assim, pode-se concluir que para modelar eficazmente tráfego IP com modelos baseados em sistemas-L, à semelhança do que acontecia com os modelos Markovianos, é necessário incluir nos modelos uma componente de modelação do processo de tamanho dos pacotes. Este aspecto será considerado no capítulo seguinte.

Capítulo 6

Modelos com base em sistemas de Lindenmayer para os processos de chegadas e de tamanhos de pacotes

6.1 Introdução

Neste capítulo serão apresentados três novos modelos de tráfego com capacidade de modelar tanto o ritmo de chegada de pacotes como o tamanho destes. Estes modelos foram desenvolvidos com o objectivo de modelar todas as componentes do tráfego IP (ou qualquer outro tipo de tráfego com tamanho de pacotes variável). Os modelos aqui apresentados permitem a captura das características multiescalares do tráfego, ao nível dos processos de chegada de pacotes e de tamanho de pacotes.

O primeiro modelo consiste em dois sistemas-L independentes, em que um modela o ritmo de chegada de pacotes e o outro o tamanho dos pacotes, sendo por isso apelidado de sistema-L duplo (SLD). O segundo modelo é um sistema-L bi-dimensional onde o ritmo de chegada de pacotes e o tamanho dos mesmos são modelados conjuntamente, o que originou a escolha do nome sistema-L conjunto (SLC). O terceiro modelo conjuga um sistema-L único, que modela a chegada dos pacotes, com o tamanho dos pacotes caracterizado por um conjunto de distribuições dependentes da taxa de chegadas, e é chamado de sistema-L com funções massa de probabilidade (SLFMP).

Este capítulo está organizado do seguinte modo. Na secção 6.2 são apresentados em pormenor os modelos de tráfego. Na secção 6.3 são descritos em detalhe os procedimentos de inferência de parâmetros dos modelos. Na secção 6.4 são apresentadas

as cascatas aleatórias duplas e a relação destas com os sistemas-L com capacidade de modelar o tamanho dos pacotes. Na secção 6.5 são apresentados e discutidos os resultados da aplicação dos modelos propostos a tráfego real. Finalmente na secção 6.6 são apresentadas algumas conclusões.

6.2 Modelos de tráfego

6.2.1 Sistema-L duplo

Este modelo de tráfego tem por base dois sistemas-L estocásticos independentes, um para o ritmo de chegadas e outro para os tamanhos dos pacotes. Para os ritmos de chegadas considera-se um sistema-L em tudo idêntico ao apresentado no capítulo anterior (sistema-L único) o qual possui o seguinte alfabeto de ritmos de chegadas

$$\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_L\}, \lambda_i \in \mathbb{R}_0^+, i = 1, \dots, L. \quad (6.1)$$

Para os tamanhos dos pacotes considera-se outro sistema-L semelhante, mas onde o alfabeto é constituído por tamanhos médios de pacote, γ_i :

$$\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_G\}, \gamma_i \in \mathbb{R}_0^+, i = 1, \dots, G. \quad (6.2)$$

As regras de produção geram aleatoriamente dois tamanhos médios de pacote a partir de um anterior. Sem perda de generalidade foi assumido que $\gamma_1 < \gamma_2 < \dots < \gamma_G$.

A série temporal de ritmos de chegadas e tamanhos médios de pacote é, construída progressivamente, governada por dois sistemas-L independentes. O tráfego é caracterizado por um único ritmo de chegada e um único tamanho médio de pacote em cada intervalo de tempo. Em cada iteração é gerada uma escala temporal mais fina (i) dividindo cada intervalo de tempo (pai) em dois subintervalos de igual duração (filhos) e (ii) associando um ritmo de chegada e um tamanho médio de pacote a cada um dos novos subintervalos, de acordo com as regras de produção de cada um dos sistemas-L estocásticos. Para cada um dos sistemas-L as escalas temporais são agrupadas em gamas de escalas (não necessariamente as mesmas para os dois sistemas-L) permitindo-se a definição de diferentes regras de produção para cada uma dessas gamas de escalas. O procedimento de inferência das gamas de escalas é o descrito no capítulo anterior. A construção do processo de tráfego está ilustrada na Figura 6.1.

Para caracterizar o processo de tráfego foi definido o par ordenado $(X_{(j,r_a)}^{(i)}, Y_{(j,r_s)}^{(i)})$, onde $X_{(j,r_a)}^{(i)} \in \Lambda$ e $Y_{(j,r_s)}^{(i)} \in \Gamma$ são respectivamente o ritmo de chegada e o tamanho

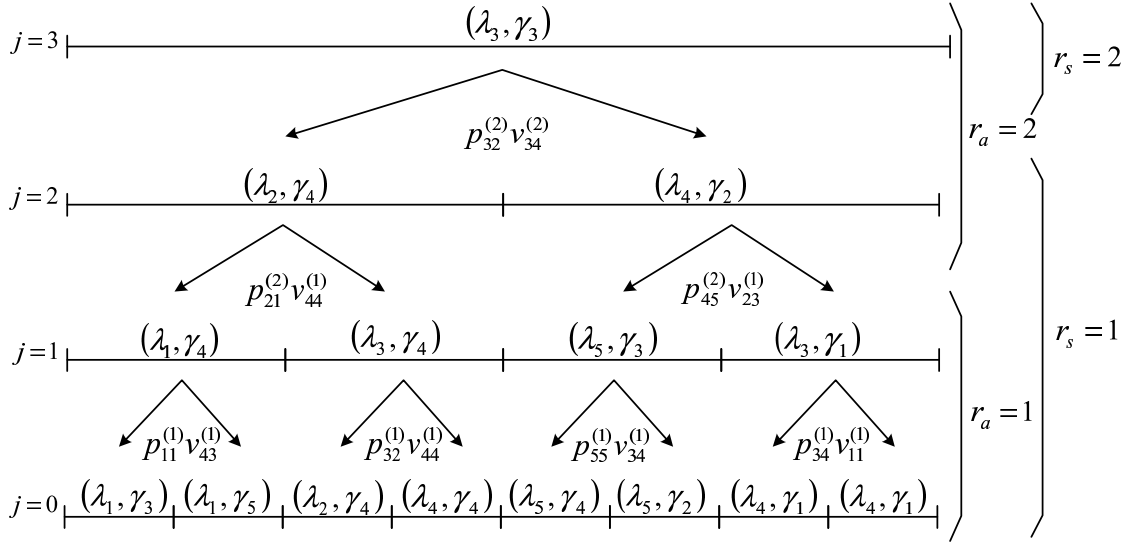


Figura 6.1: Representação da evolução do modelo SLD.

médio de pacote no intervalo de tempo i da escala temporal j e gamas de escalas temporais r_a para os ritmos de chegadas e r_s para os tamanhos médios. O número de escalas temporais é S , o número de gamas de escalas temporais é R_a para os ritmos de chegadas e R_s para os tamanhos médios. Também aqui, j decresce de $j = S - 1$ (na escala temporal mais larga) até $j = 0$ (escala temporal mais fina); e r_a (r_s) decresce de R_a (R_s) (a gama temporal das escalas temporais mais largas) até 1 (a gama temporal das escalas temporais mais finas). O número de intervalos de tempo na escala j será $N_j = 2^{S-j-1}$. Mais uma vez, assumindo uma duração unitária para os intervalos na escala temporal inferior, $j = 0$, a duração de um intervalo na escala j será 2^j . De modo a relacionar as escalas temporais com as gamas temporais, define-se j_{r_a} e j_{r_s} como a escala superior j nas gamas r_a e r_s respectivamente. Então para o caso ilustrado na Figura 6.1, teremos $S = 4$, $R_a = 2$, $R_s = 2$, $j_2 = 3$ e $j_1 = 1$ para os ritmos, e $j_2 = 3$ e $j_1 = 2$ para os tamanhos.

De modo a garantir que o ritmo de chegadas médio e o tamanho médio de pacote (calculado para todos os intervalos de tempo) são os mesmos em todas as escalas temporais, mantendo assim o significado físico, são impostas as seguintes condições às regras de produção:

$$X_{(j,r_a)}^{(i)} = \frac{1}{2}X_{(j-1,r'_a)}^{(2i-1)} + \frac{1}{2}X_{(j-1,r'_a)}^{(2i)} \quad (6.3)$$

$$Y_{(j,r_s)}^{(i)} = \frac{1}{2}Y_{(j-1,r'_s)}^{(2i-1)} + \frac{1}{2}Y_{(j-1,r'_s)}^{(2i)} \quad (6.4)$$

isto é, a geração de ritmos de chegadas (tamanhos médios) de pacotes é tal que a média dos ritmos de chegadas (tamanhos médios) nos subintervalos filhos da direita

e da esquerda será igual ao ritmo de chegada (tamanho médio) do intervalo pai. De notar que estas regras não garantem que a quantidade de bytes num determinado intervalo pai seja igual à soma dos bytes dos intervalos filhos respectivos, apesar de garantirem que a quantidade total de bytes é a mesma em todas as escalas temporais. Seria possível impor condições mais complexas para a relação dos ritmos de chegadas e dos tamanhos médios dos pacotes, mas isso tornaria impossível definir um alfabeto de tamanhos médios de pacotes finito.

O processo de tráfego pode ser descrito pelos axiomas $X_{(S-1, R_a)}^{(1)}$ e $Y_{(S-1, R_s)}^{(1)}$, respectivamente o ritmo de chegadas e tamanho médio de pacote na escala temporal superior, e pelas regras de produção independentes definidas por

$$X_{(j, r_a)}^{(i)} = \lambda_l \xrightarrow{p_{lq}^{(r_a)}} \begin{cases} X_{(j-1, r'_a)}^{(2i-1)} = \lambda_q \\ X_{(j-1, r'_a)}^{(2i)} = 2\lambda_l - \lambda_q \end{cases} \quad (6.5)$$

e

$$Y_{(j, r_s)}^{(i)} = \gamma_m \xrightarrow{v_{mn}^{(r_s)}} \begin{cases} Y_{(j-1, r'_s)}^{(2i-1)} = \gamma_n \\ Y_{(j-1, r'_s)}^{(2i)} = 2\gamma_m - \gamma_n \end{cases} \quad (6.6)$$

onde $\sum_{q=1}^L p_{lq}^{(r_a)} = 1, \forall l$, e $\sum_{n=1}^G v_{mn}^{(r_s)} = 1, \forall m$. Assim, um ritmo de chegadas λ_l (tamanho médio γ_m) no intervalo i , escala j e gama r_a (r_s) produz, com probabilidade $p_{lq}^{(r_a)}$ ($v_{mn}^{(r_s)}$), um ritmo de chegadas λ_q (tamanho médio γ_n) no subintervalo filho da esquerda $2i - 1$ e um ritmo de chegadas $2\lambda_l - \lambda_q$ (tamanho médio $2\gamma_m - \gamma_n$) no subintervalo filho da direita $2i$, da próxima escala temporal $j - 1$ e gama r'_a (r'_s). As regras de produção para os ritmos de chegadas podem ser totalmente descritas por R_a matrizes com dimensão $L \times L$

$$\mathbf{P}^{(r_a)} = \left(p_{lq}^{(r_a)} \right), l, q = 1, \dots, L, r_a = 1, \dots, R_a \quad (6.7)$$

e as regras de produção dos tamanhos médios de pacotes podem ser totalmente descritas por R_s matrizes com dimensão $G \times G$

$$\mathbf{V}^{(r_s)} = \left(v_{mn}^{(r_s)} \right), m, n = 1, \dots, G, r_s = 1, \dots, R_s \quad (6.8)$$

É ainda necessário garantir que o alfabeto é fechado no que respeita às regras de produção. São então impostas as seguintes condições para o sistema-L dos ritmos de chegadas: (i) $\lambda_i - \lambda_{i-1} = \frac{\lambda_L - \lambda_1}{L-1}, i = 2, 3, \dots, L$, isto é, os valores λ_i serão equidistantes; (ii) $p_{lq}^{(r_a)} = 0$ se $q > l + \min(L-l, l-1)$ ou $q < l - \min(L-l, l-1)$. Condições similares existem para o sistema-L dos tamanhos médios de pacote: (i) $\gamma_i - \gamma_{i-1} = \frac{\gamma_G - \gamma_1}{G-1}, i = 2, 3, \dots, G$; (ii) $v_{mn}^{(r_s)} = 0$ se $n > m + \min(G-m, m-1)$ ou $n < m - \min(G-m, m-1)$.

Finalmente, o processo de tráfego na escala j pode ser definido pela seguinte série temporal

$$Z_{(j,r_a,r_s)} = \{(X_{(j,r_a)}^{(i)}, Y_{(j,r_s)}^{(i)}), i = 1, \dots, N_j\} \quad (6.9)$$

Este modelo tem um total de $R_a L^2 + L + R_s G^2 + G$ parâmetros.

A geração do processo de tráfego completo a partir do modelo pode ser feita em três passos: (i) em primeiro lugar são geradas as sequências temporais $X_{(0,1)}^{(i)}$ e $Y_{(0,1)}^{(i)}$ com $i = 1, \dots, N_j$, (ii) em segundo lugar, para cada intervalo temporal, os instantes de chegada dos pacotes são definidos de acordo com uma distribuição uniforme (o número de chegadas é definido pelo ritmo de chegada de pacotes desse intervalo); (iii) finalmente é atribuído a todos os pacotes do i -ésimo intervalo um tamanho igual ao tamanho médio definido por $Y_{(0,1)}^{(i)}$.

6.2.2 Sistema-L conjunto

Nesta secção é apresentado o modelo de sistema-L conjunto, o qual descreve conjuntamente o ritmo de chegada e o tamanho médio dos pacotes. Este modelo difere do modelo SLD pois baseia-se num único sistema-L bi-dimensional em que o alfabeto é formado por pares ordenados de ritmos de chegadas e tamanhos médios de pacotes. Deste modo é introduzida no modelo a capacidade de correlacionar os ritmos de chegada de pacotes com o tamanho destes. A sequência temporal é construída de acordo com um sistema-L estocástico bi-dimensional, em que o alfabeto é definido por

$$\{(\lambda_l, \gamma_g) : \lambda_l \in \Lambda, \gamma_g \in \Gamma\} \quad (6.10)$$

onde

$$\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_L\}, \lambda_l \in \mathbb{R}_0^+, l = 1, \dots, L \quad (6.11)$$

$$\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_G\}, \gamma_g \in \mathbb{R}_0^+, g = 1, \dots, G \quad (6.12)$$

e onde as regras de produção geram aleatoriamente dois pares (de ritmos e tamanhos médios) a partir de outro.

O processo de tráfego é construído iterativamente de acordo com o sistema-L estocástico bi-dimensional, em que em cada iteração é produzida uma nova escala temporal. Começando pela escala temporal superior, em cada iteração são gerados dois pares ordenados de ritmo de chegada e tamanho médio de pacotes de acordo com regras de produção do sistema-L. Também neste modelo é permitido o agrupamento de escalas temporais em gamas de escalas temporais com regras de produção próprias. Mas agora

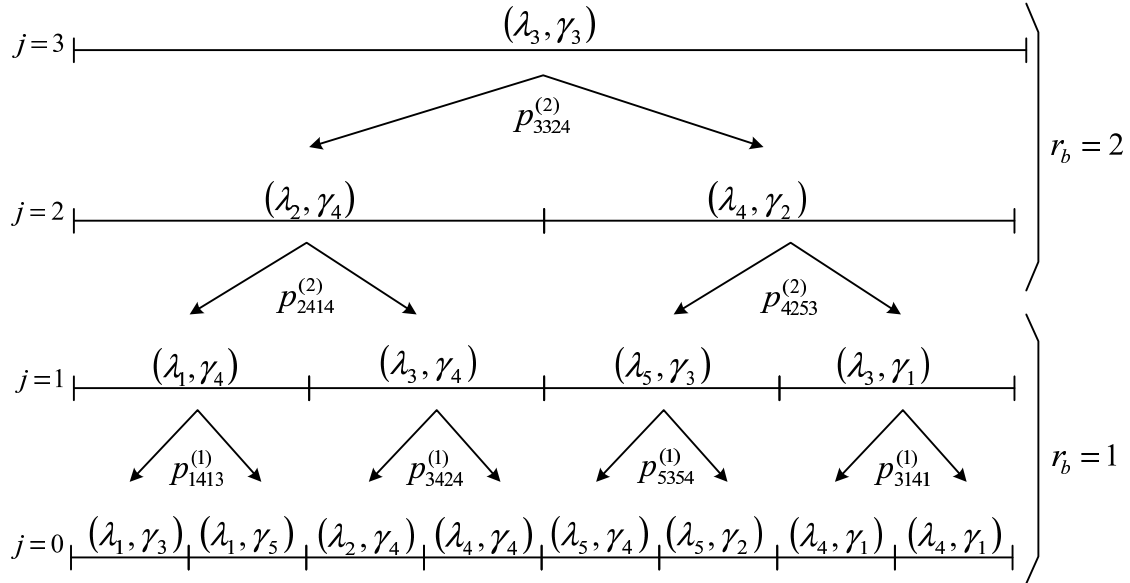


Figura 6.2: Representação da evolução do modelo SLC.

o agrupamento das escalas é feito analisando as características de similaridade escalar do processo de tráfego conjunto de ritmos de chegadas e tamanho médio dos pacotes, ou seja o processo de chegada em bytes por segundo. A construção do processo de tráfego está ilustrada na Figura 6.2.

O processo de tráfego continua a ser caracterizado por um par ordenado $(X_{(j,r_b)}^{(i)}, Y_{(j,r_b)}^{(i)})$, onde $X_{(j,r_b)}^{(i)} \in \Lambda$ e $Y_{(j,r_b)}^{(i)} \in \Gamma$ são respectivamente o ritmo de chegada e o tamanho médio dos pacotes no intervalo de tempo i da escala temporal j e gama de escalas temporais r_b . O número de gamas de escalas temporais é R_b . As definições de escala temporal, gama de escalas temporais e dos diferentes parâmetros associados às escalas são idênticas às da secção anterior.

Mais uma vez, é necessário garantir que o ritmo de chegadas médio e o tamanho médio dos pacotes (calculados sobre todos os intervalos de tempo) são os mesmos em todas as escalas temporais. São então impostas as seguintes condições às regras de produção:

$$X_{(j,r_b)}^{(i)} = \frac{1}{2}X_{(j-1,r'_b)}^{(2i-1)} + \frac{1}{2}X_{(j-1,r'_b)}^{(2i)} \quad (6.13)$$

$$Y_{(j,r_b)}^{(i)} = \frac{1}{2}Y_{(j-1,r'_b)}^{(2i-1)} + \frac{1}{2}Y_{(j-1,r'_b)}^{(2i)} \quad (6.14)$$

isto é, tal como na secção anterior, a média dos ritmos (tamanhos médios) dos pacotes nos subintervalos filhos da direita e da esquerda será igual ao ritmo (tamanho médio) do intervalo pai. De notar mais uma vez que a regra (6.14) não garante que a quantidade de bytes num determinado intervalo pai seja igual à soma dos bytes nos respectivos intervalos filhos.

O processo de tráfego pode ser descrito pelo axioma $(X_{(S-1, R_b)}^{(1)}, Y_{(S-1, R_b)}^{(1)})$, o par ordenado do ritmo de chegadas e tamanho médio dos pacotes na escala temporal superior, e pelas regras de produção definidas por

$$(X_{(j, r_b)}^{(i)}, Y_{(j, r_b)}^{(i)}) = (\lambda_l, \gamma_m) \xrightarrow{p_{lmqn}^{(r_b)}} \begin{cases} (X_{(j-1, r'_b)}^{(2i-1)}, Y_{(j-1, r'_b)}^{(2i-1)}) = (\lambda_q, \gamma_n) \\ (X_{(j-1, r'_b)}^{(2i)}, Y_{(j-1, r'_b)}^{(2i)}) = (2\lambda_l - \lambda_q, 2\gamma_m - \gamma_n) \end{cases} \quad (6.15)$$

onde $\sum_{q=1}^L \sum_{n=1}^G p_{lmqn}^{(r_b)} = 1, \forall l, \forall m$. Assim, um par com um ritmo de chegadas λ_l e tamanho de médio γ_m no intervalo i , escala j e gama r_b produz, com probabilidade $p_{lmqn}^{(r_b)}$, um par com um ritmo λ_q e tamanho médio γ_n no subintervalo filho da esquerda $2i - 1$ e um par com um ritmo de chegadas $2\lambda_l - \lambda_q$ e tamanho de médio $2\gamma_m - \gamma_n$ no subintervalo filho da direita $2i$, da próxima escala temporal $j - 1$ e gama r'_b . As regras de produção para os ritmos de chegadas podem ser totalmente descritas por R_b matrizes com dimensão $L \times G \times L \times G$

$$\mathbf{P}^{(r_b)} = \left(p_{lmqn}^{(r_b)} \right) \quad (6.16)$$

com $l, q = 1, \dots, L, m, n = 1, \dots, G$ e $r_b = 1, \dots, R_b$.

Também para este sistema-L é necessário garantir que o alfabeto é fechado no que respeita às regras de produção. São assim impostas as seguintes condições: (i) $\lambda_i - \lambda_{i-1} = \frac{\lambda_L - \lambda_1}{L-1}, i = 2, 3, \dots, L$, isto é, os valores λ_i serão equidistantes; (ii) $\gamma_i - \gamma_{i-1} = \frac{\gamma_G - \gamma_1}{G-1}, i = 2, 3, \dots, G$, isto é, os valores γ_i serão equidistantes; (iii) $p_{lmqn}^{(r_b)} = 0$ if $q > l + \min(L-l, l-1)$ ou $q < l - \min(L-l, l-1)$; (iv) $p_{lmqn}^{(r_b)} = 0$ se $n > m + \min(G-m, m-1)$ ou $n < m - \min(G-m, m-1)$.

Finalmente, o processo de tráfego na escala j pode ser definido pela seguinte sequência temporal

$$Z_{(j, r_b)} = \{(X_{(j, r_b)}^{(i)}, Y_{(j, r_b)}^{(i)}), i = 1, \dots, N_j\} \quad (6.17)$$

O modelo SLC tem um total de $R_b L^2 G^2 + L + G$ parâmetros.

A geração do processo de tráfego completo a partir do modelo começa por gerar a sequência temporal $Z_{(0,1)}$. A partir daí segue os mesmos passos do modelo anterior.

6.2.3 Sistema-L com FMP

O modelo apresentado nesta subsecção utiliza um sistema-L único (ver secção 5.3 do capítulo 5) para modelar o processo de chegada de pacotes e um conjunto de distribuições discretas para caracterizar o processo do tamanho dos pacotes. O tamanho dos

pacotes é apenas caracterizado na escala temporal mais fina. Além disso, contrariamente ao que acontecia nos dois modelos anteriores, o sistema-L com FMPs caracteriza o tamanho de cada pacote individualmente, em vez do tamanho médio de pacote no intervalo de amostragem. A cada ritmo de chegadas é associada uma função massa de probabilidade (FMP) que descreve o tamanho dos pacotes. Assim, haverá um total de L FMPs.

Para caracterizar o processo dos tamanhos dos pacotes, recorre-se à variável aleatória discreta E_l que representa o tamanho de um pacote associado ao ritmo de chegadas λ_l . Definem-se também L funções massa de probabilidade (FMPs), uma para cada um dos ritmos de chegadas, dadas por $h^l(e) = P(E_l = e)$, com $e \in \Upsilon^l = \{v_1, v_2, \dots, v_{Q_l}\}$ e $l = 1, 2, \dots, L$; Υ^l é o conjunto (aqui também designado por alfabeto) dos Q_l tamanhos de pacotes associados ao ritmo de chegada λ_l . De referir que todos os pacotes que cheguem no intervalo de tempo associado ao ritmo de chegadas λ_l terão um tamanho atribuído segundo a FMP h^l .

O modelo SLFMP tem um total de $R_a L^2 + L + 2 \sum_{l=1}^L Q_l$ parâmetros.

Neste caso, a geração do processo de tráfego completo a partir do modelo compreende os três passos seguintes: (i) em primeiro lugar é gerada a sequência de ritmos de chegadas de pacotes na escala temporal mais fina dada pelo sistema-L correspondente; (ii) em segundo lugar, para cada intervalo temporal, os instantes de chegada dos pacotes são definidos de acordo com uma distribuição uniforme (o número de chegadas é definido pelo ritmo de chegadas de pacotes desse intervalo); (iii) finalmente são atribuídos os tamanhos aos pacotes, de acordo com as FMPs associadas aos diferentes ritmos de chegadas.

6.3 Procedimentos de inferência de parâmetros

6.3.1 Sistema-L duplo

O procedimento de inferência determina os parâmetros dos dois sistemas-L a partir de dados reais. Começa-se por fixar um intervalo de amostragem Δ e construir a série temporal do número de chegadas de pacotes e do tamanho médio dos mesmos em cada um dos intervalos de amostragem, que será designada por $\{(A_k, B_k), k = 1, 2, \dots, K\}$, onde A_k e B_k representam respectivamente o número de chegadas de pacotes e tamanho médio dos pacotes no intervalo de amostragem k . Por conveniência, o tamanho da

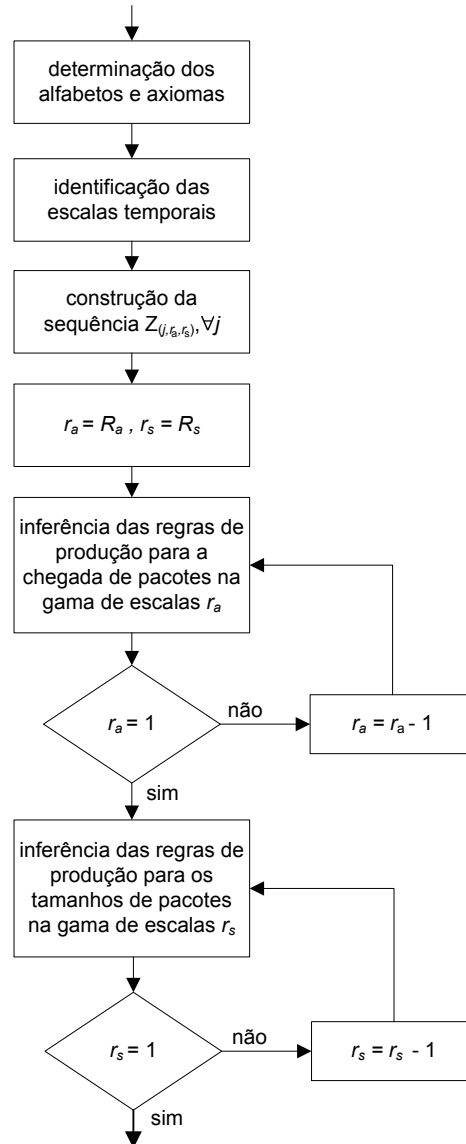


Figura 6.3: Diagrama de fluxo do procedimento de inferência do modelo SLD.

série temporal K deverá ser uma potência de 2. O procedimento de inferência pode ser dividido em três passos: (i) determinação do alfabeto e do axioma dos ritmos de chegadas e tamanhos médios de pacotes, (ii) identificação das gamas de escalas temporais para cada sistema-L e (iii) inferência das regras de produção dos ritmos de chegadas e tamanhos médios dos pacotes. O diagrama de fluxos de cada um dos sistemas-L que constituem o modelo está representado na figura 6.3.

Os alfabetos dos sistemas-L consistem em L e G valores equidistantes desde o valor mínimo ao máximo presente nos dados. Tanto o valor de L como o de G devem ser um compromisso entre a precisão e complexidade do modelo. À semelhança do modelo de sistema-L único, esta é apenas uma das possibilidades para a construção do alfabeto.

Devido à propriedade de preservação de massa dos sistemas-L definida em (6.3) e (6.4), impõe-se que o axioma do ritmo de chegadas de pacotes seja igual ao ritmo médio de chegada de pacotes de $\{A_k\}$ arredondado para o elemento do alfabeto mais próximo, isto é,

$$X_{(S-1, R_a)}^{(1)} = \Phi_\Lambda \left((1/K\Delta) \sum_{k=1}^K A_k \right) \quad (6.18)$$

e, de forma análoga, o axioma dos tamanhos médios dos pacotes é dado por

$$Y_{(S-1, R_s)}^{(1)} = \Phi_\Gamma \left((1/K) \sum_{k=1}^K B_k \right) \quad (6.19)$$

onde $\Phi_\Omega(x)$ representa a função que arredonda x para o elemento mais próximo de Ω .

Tal como na inferência do sistema-L único (capítulo 5), a identificação das gamas temporais baseia-se no método de análise escalar por onduletas descrito na secção 2.3.4.2 do capítulo 2 deste Tese. Este método é aplicado tanto ao ritmo de chegadas como ao tamanho médio dos pacotes.

O próximo passo é a inferência das regras de produção dos dois sistemas-L, as quais são caracterizados de forma completa pelas matrizes $\mathbf{P}^{(r_a)}$ e $\mathbf{V}^{(r_s)}$. Inicialmente, os dados são arredondados de modo a definir as sequências de ritmos de chegadas e tamanhos médios de pacotes em cada escala temporal. Os ritmos de chegadas de pacotes $X_{(j, r_a)}^{(i)}$ são obtidos a partir de $\{A_k\}$ através de

$$X_{(j, r_a)}^{(i)} = \Phi_\Lambda \left((N_j/K\Delta) \sum_{k=K(i-1)/N_j+1}^{K i/N_j} A_k \right) \quad (6.20)$$

e os tamanhos médios dos pacotes $Y_{(j, r_s)}^{(i)}$ através de

$$Y_{(j, r_s)}^{(i)} = \Phi_\Gamma \left((N_j/K) \sum_{k=K(i-1)/N_j+1}^{K i/N_j} B_k \right) \quad (6.21)$$

com $i = 1, \dots, N_j$, para cada j .

Definindo $c_{lq}^{(r_a)}$ como sendo o número de vezes que, para todos as escalas da gama r_a , o pai $X_{(j, r_a)}^{(i)} = \lambda_l$ produziu um filho no subintervalo da esquerda $X_{(j-1, r'_a)}^{(2i-1)} = \lambda_q$, as probabilidades das regras de produção dos ritmos de chegadas são dadas por

$$p_{lq}^{(r_a)} = c_{lq}^{(r_a)} / \sum_{q=1}^L c_{lq}^{(r_a)}, \quad l = 1, \dots, L, \quad r_a = 1, \dots, R_a \quad (6.22)$$

De mesmo modo define-se $d_{mn}^{(r_s)}$ como sendo o número de vezes que, para todos as escalas da gama r_s , o pai $Y_{(j, r_s)}^{(i)} = \gamma_m$ produziu um filho no subintervalo da esquerda

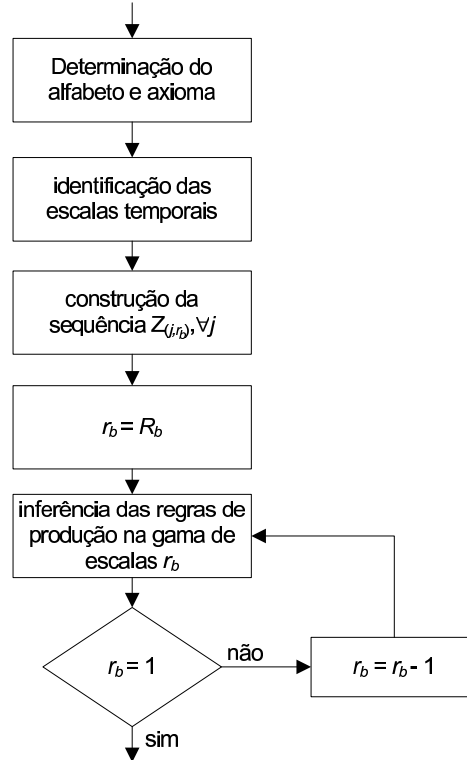


Figura 6.4: Diagrama de fluxo do procedimento de inferência do modelo SLC.

$Y_{(j-1,r'_s)}^{(2i-1)} = \gamma_n$. As probabilidades das regras de produção para os tamanhos médios são então dadas por

$$v_{mn}^{(r_s)} = d_{mn}^{(r_s)} / \sum_{n=1}^G d_{mn}^{(r_s)}, \quad m = 1, \dots, G, \quad r_s = 1, \dots, R_s \quad (6.23)$$

6.3.2 Sistema-L conjunto

Tal como no modelo anterior, a inferência de parâmetros começa pela construção da série temporal empírica $\{(A_k, B_k), k = 1, 2, \dots, K\}$. O procedimento de inferência está também dividido em 3 partes: (i) determinação do alfabeto e axioma, (ii) identificação das gamas de escalas temporais, (iii) inferência das regras de produção. O diagrama de fluxos do procedimento de inferência deste sistema-L está representado na Figura 6.4.

O alfabeto deste sistema-L bi-dimensional consistirá num conjunto com $L \times G$ pares ordenados (λ_l, γ_g) , onde tanto os elementos λ_l como os γ_g serão equidistantes, variando do máximo ao mínimo presente nos dados. O axioma será constituído pelo ritmo médio de chegadas de pacotes de $\{A_k\}$ arredondado para o elemento de Λ mais próximo e

pela média de $\{B_k\}$ arredondada para o elemento de Γ mais próximo, isto é,

$$X_{(S-1, R_b)}^{(1)} = \Phi_\Lambda \left((1/K\Delta) \sum_{k=1}^K A_k \right) \quad (6.24)$$

$$Y_{(S-1, R_b)}^{(1)} = \Phi_\Gamma \left((1/K) \sum_{k=1}^K B_k \right). \quad (6.25)$$

As regras de produção deste sistema-L bidimensional são totalmente caracterizadas pela matriz $\mathbf{P}^{(r_b)}$. A inferência das regras de produção à semelhança dos modelos anteriores começa pela construção das sequências de ritmos de chegadas e de tamanhos médios de pacotes para cada escala temporal. A inferência destas sequências é feita de acordo com (6.20) e (6.21).

A particularidade deste modelo está na inferência conjunta das probabilidades de geração de ritmos de chegadas e tamanhos médios de pacotes. Definido agora $c_{lmqn}^{(r_b)}$ como sendo o número de vezes que, para todas as escalas temporais da gama temporal r_b , o pai $(X_{(j, r_b)}^{(i)}, Y_{(j, r_b)}^{(i)}) = (\lambda_l, \gamma_m)$ produz um filho no intervalo da esquerda $(X_{(j-1, r_b')}^{(2i-1)}, Y_{(j-1, r_b')}^{(2i-1)}) = (\lambda_q, \gamma_n)$, as probabilidades das regras de produção podem ser obtidas através de

$$p_{lmqn}^{(r_b)} = c_{lmqn}^{(r_b)} / \sum_{q=1}^L \sum_{n=1}^G c_{lmqn}^{(r_b)} \quad (6.26)$$

com $l = 1, \dots, L, m = 1, \dots, G$ e $r_b = 1, \dots, R_b$.

6.3.3 Sistema-L com FMPs

O sistema-L que caracteriza o processo de chegada de pacotes é inferido exactamente como no caso do sistema-L único (subsecção 5.3.1 do capítulo 5). Note-se que neste primeiro passo é implicitamente associado a cada pacote individual um determinado ritmo de chegada (na escala mais fina). O ajuste do processo do tamanho dos pacotes começa por considerar a série temporal de tamanhos de pacotes (individuais) $\{C_n, n = 1, 2, \dots, N\}$, onde C_n é o tamanho do n -ésimo pacote de um total de N pacotes. É efectuada de seguida uma partição desta série temporal em L subconjuntos definidos por

$$D^l = \{C_n : L(n) = \lambda_l, n = 1, 2, \dots, N\}, l = 1, 2, \dots, L, \quad (6.27)$$

onde $L(n)$ representa o ritmo de chegadas que foi associado ao n -ésimo pacote. Assim, o subconjunto D^l agrupa os tamanhos de todos os pacotes que foram associados ao ritmo de chegadas λ_l . Estes subconjuntos serão usados para inferir os alfabetos dos

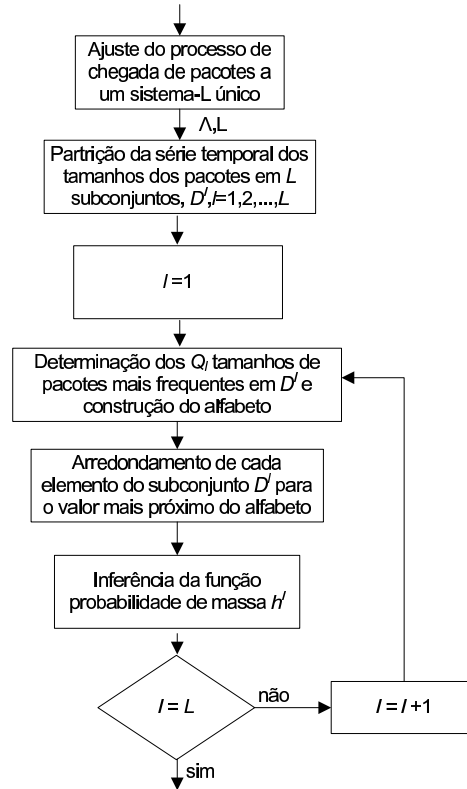


Figura 6.5: Diagrama de fluxo do procedimento de inferência do modelo de sistema-L com FMPs.

tamanhos dos pacotes (individuais) associados a cada ritmo de chegadas. De seguida são determinados os Q_l tamanhos de pacotes mais frequentes usando, por exemplo, histogramas dos subconjuntos D^l . Estes tamanhos são agrupados nos alfabetos $\Upsilon^l = \{v_1, v_2, \dots, v_{Q_l}\}$. Neste ponto, cada elemento do subconjunto D^l é arredondado para o elemento mais próximo de Υ^l , dando assim origem a novos subconjuntos

$$\tilde{D}^l = \{\Phi_{\Upsilon^l}(x), x \in D^l\}, l = 1, 2, \dots, L \quad (6.28)$$

Finalmente, a partir dos subconjuntos \tilde{D}^l são inferidas as correspondentes FMP $h^l(e) = P\{E_l = e\}$, $e \in \Upsilon^l$. O diagrama de fluxos do procedimento de inferência deste sistema-L está representado na Figura 6.5.

É possível ainda considerar que o conjunto dos pacotes mais frequentes é independente do ritmo de chegadas, isto é $\Upsilon = \Upsilon^l = \Upsilon^{l+1}$, $l = 1, \dots, L$. Neste caso, para a construção do conjunto Υ seriam considerados os Q tamanhos de pacotes mais frequentes do conjunto total de dados.

6.4 Relação com cascatas conservativas

Nesta secção será feita uma extensão da análise da secção 5.4, comparando os sistemas-L com capacidade de modelação dos tamanhos dos pacotes com as chamadas cascatas conservativas duplas. Este modelo é uma extensão do proposto em [FGW98], sendo constituído por duas cascatas conservativas independentes, uma que modela o processo de chegadas e outra que modela o tamanho médio dos pacotes. O procedimento de inferência dos parâmetros de cada cascata conservativa, é idêntico ao descrito na secção 5.5. Também neste caso são considerados dois tipos de modelos que se distinguem quanto ao tipo de gerador: cascatas cujos geradores têm distribuição Gaussiana truncada (designados por cascatas duplas Gaussianas) e cascatas cujos geradores têm distribuição discreta (designados por cascatas duplas discretas).

Para efectuar esta análise, definiu-se $W_{b_{j|m}}$ como sendo o gerador da sequência de tamanhos médios de pacotes, na escala j , condicionado pelo ritmo de chegadas γ_m . O gerador $W_{b_{j|m}}$ é uma variável aleatória discreta que, para cada m , pode tomar os valores $\gamma_m/2\gamma_n$ e pode ser facilmente inferido a partir dos dados. Se o pressuposto de independência for verdadeiro então $W_{b_{j|m}}$ deve ter a mesma distribuição para todo o m . Nas Figuras 6.6 e 6.7 está representada a média e variância de $W_{b_{j|m}}$, para cada γ_m e para $j = 1, 2, 3, 4$ e 5 . Os resultados mostram que, tal como acontecia para o processo de chegadas (secção 5.4), os geradores apresentam diferentes valores para a média e variância numa mesma escala temporal. Por exemplo, para $j = 1$ o gerador $W_{b_{j|m}}$ tem uma variância de 0.0516 para $\gamma_m = 341$ bytes e 3.51×10^{-3} para $\gamma_m = 701$ bytes. Verifica-se assim que o gerador possui distribuições dependentes do valor do tamanho médio dos pacotes da escala temporal superior. Portanto o pressuposto de independência do gerador não é igualmente válido para o processo do tamanho médio de pacotes, no caso da captura Bellcore. Este resultado reforça a ideia de que os modelos baseados em sistemas-L podem ser vantajosos relativamente aos baseados em cascatas conservativas. Este aspecto será analisado com maior detalhe na secção seguinte.

6.5 Resultados numéricos

Mais uma vez, os modelos e respectivos procedimento de inferência foram aplicados a duas capturas de tráfego: (i) a captura Bellcore de Outubro e (ii) uma captura feita na Universidade de Aveiro (UA). Ambas as capturas estão descritas em pormenor no Apêndice A. O intervalo de amostragem usado foi de 0.1 segundos em ambas as

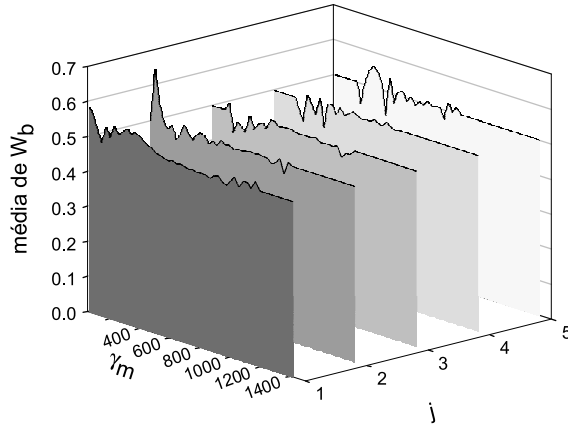


Figura 6.6: Média de $W_{b|j|_m}$ em função de γ_m e j , tamanho médio dos pacotes da captura Bellcore.

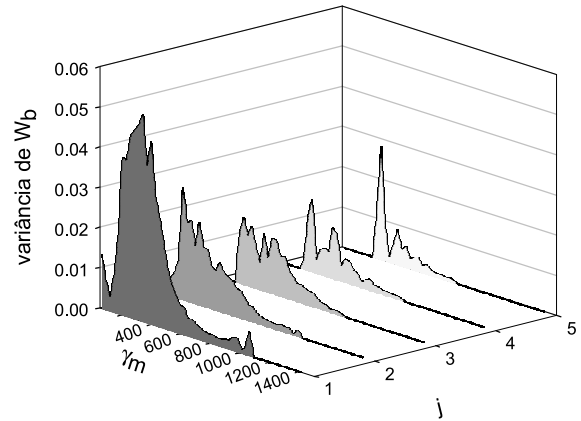


Figura 6.7: Variância de $W_{b|j|_m}$ em função de γ_m e j , tamanho médio dos pacotes da captura Bellcore.

capturas.

6.5.1 Modelos inferidos

6.5.1.1 Sistema-L duplo

A captura Bellcore foi em primeiro lugar ajustada por um SLD. O processo de chegada de pacotes foi modelado pelo sistema-L único já inferido na secção 5.5 do capítulo 5, onde alfabeto tinha um comprimento de $L = 236$ ritmos de chegadas e tinham sido identificadas 5 gamas de escalas temporais com características escalares similares. O processo de tamanho dos pacotes foi modelado por um sistema-L com um alfabeto de $G = 50$ tamanhos médios de pacote, variando de 64 bytes a 1421 bytes, em intervalos de 28 bytes. A estratégia usada na escolha do alfabeto dos tamanhos médios de pacote foi diferente da usada na escolha do alfabeto dos ritmos de chegadas. No caso do alfabeto dos ritmos de chegadas definiu-se à partida o intervalo entre elementos do alfabeto; para o alfabeto dos tamanhos médios de pacote definiu-se à partida qual ia ser o comprimento do alfabeto. Utilizando um diagrama de energias em escala logarítmica identificaram-se 3 gamas de escalas temporais (de um total de 14 escalas temporais) para o processo dos tamanhos dos pacotes, definidas por $j_1 = 3, j_2 = 7$ e $j_3 = 14$ (Figura 6.8).

A captura UA foi ajustada a um SLD onde mais uma vez o processo de chegada de pacotes foi modelado pelo sistema-L único, já inferido na secção 5.5 do capítulo 5, onde o alfabeto tinha um comprimento de $L = 214$ ritmos de chegadas e tinham sido identificadas 3 gamas de escalas temporais. O processo de tamanho dos pacotes foi

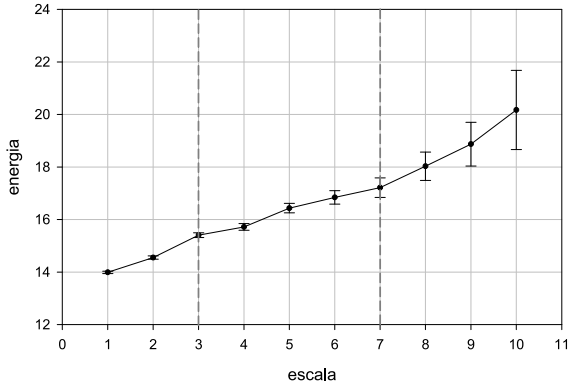


Figura 6.8: diagrama de energias em escala logarítmica, processo TMP, Bellcore.

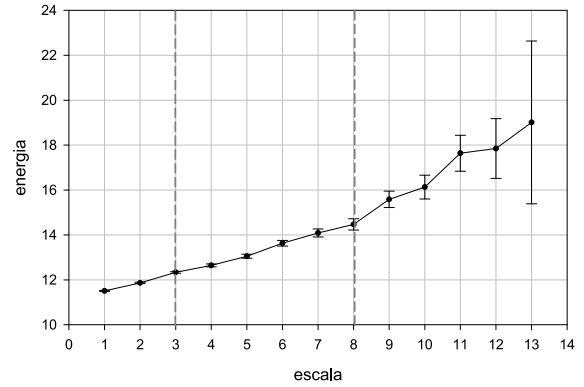


Figura 6.9: diagrama de energias em escala logarítmica, processo TMP, UA.

modelado por um sistema-L com um alfabeto de $G = 50$ tamanhos médios de pacote, variando de 240 bytes a 1012 bytes, em intervalos de 16 bytes. Através do diagrama de energias em escala logarítmica identificaram-se 3 gamas de escalas temporais (de um total de 17 escalas temporais) para o processo dos tamanhos dos pacotes, definidas por $j_1 = 3, j_2 = 8$ e $j_3 = 17$ (Figura 6.9).

6.5.1.2 Sistema-L conjunto

A captura Bellcore foi ajustada a um SLC com um alfabeto bi-dimensional de 7080 elementos, correspondendo a $L = 236$ ritmos de chegadas e $G = 30$ tamanhos médios de pacotes. Os critérios usados na escolha dos tamanhos dos alfabetos foram os já utilizados anteriormente; a única diferença foi a escolha de um menor número de tamanhos médios de pacote possíveis devido à maior complexidade deste modelo. No caso do ritmo de chegadas, o mínimo e o máximo observados foram 10 e 2360 pacotes/s respectivamente, e foi usado um intervalo entre elementos do alfabeto de 10 pacotes/s. No caso dos tamanhos médios dos pacotes, o máximo e mínimo observados foram 64 bytes e 1421 bytes respectivamente e foi usado um intervalo entre elementos de 47 bytes. Através da análise do processo de chegadas (de bytes por segundo) utilizando um diagrama de energias em escala logarítmica identificaram-se 4 gamas de escalas temporais (de um total de 14 escalas temporais) definidas por $j_1 = 5, j_2 = 8, j_3 = 9$ e $j_4 = 14$ (Figura 6.10).

A aplicação do procedimento de inferência à captura UA, resultou num SLC com um alfabeto bi-dimensional de 6420 elementos correspondendo a $L = 214$ ritmos de chegadas e $G = 30$ tamanhos médios de pacote, de 450 a 2580 pacotes/s com intervalos de 10 pacotes/s, e 240 a 1012 bytes com intervalos de 27 bytes. A análise do processo de chegadas (de bytes por segundo) revelou a existência de 3 gamas de escalas temporais

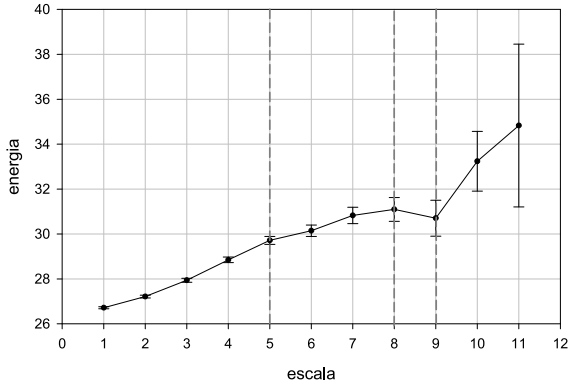


Figura 6.10: diagrama de energias em escala logarítmica, processo BPS, Bellcore.

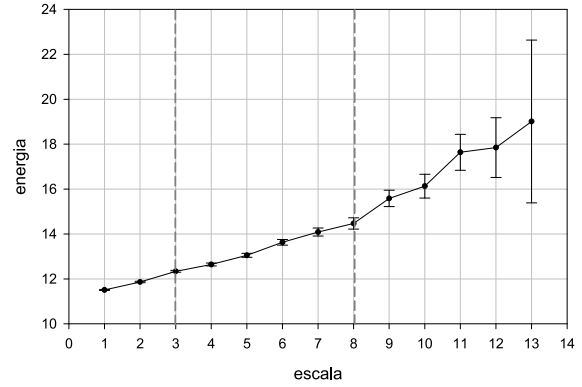


Figura 6.11: diagrama de energias em escala logarítmica, processo BPS, UA.

(de um total de 17 escalas temporais) definidas por $j_1 = 3, j_2 = 8$ e $j_3 = 17$ (Figura 6.11).

6.5.1.3 Sistema-L com FMP

A captura Bellcore foi ajustada ao sistema-L único da secção 5.5 do capítulo 5. Deste modo, o tamanho dos pacotes é caracterizado por 236 histogramas, cada um deles associado a um dos ritmos de chegadas do sistema-L. Após inferir os alfabetos $\Upsilon^l, l = 1, \dots, L$, verificou-se que estes eram muito próximos e, portanto, optou-se por considerar um único alfabeto independente do ritmo de chegadas dos pacotes. Assim, a partir de um histograma foram determinados os 20 tamanhos de pacotes mais frequentes nos dados (em bytes): $\Upsilon = \{64, 74, 94, 102, 110, 126, 130, 142, 150, 162, 174, 190, 570, 938, 986, 1082, 1090, 1242, 1282, 1518\}$.

Para a captura UA usou-se também o sistema-L único já inferido na secção 5.5 do capítulo 5. Também para esta captura se optou por considerar um alfabeto independente do ritmo de chegadas dos pacotes e com 20 elementos: $\Upsilon = \{40, 48, 52, 60, 77, 114, 213, 351, 386, 552, 576, 608, 628, 837, 932, 1216, 1400, 1420, 1462, 1500\}$.

6.5.2 Resultados comparativos

A aplicabilidade dos modelos de tráfego e a precisão dos procedimentos de inferência a eles associados foram avaliados segundo vários critérios. Compararam-se as funções massa de probabilidade (FMP) e de autocovariância (FAC) dos processos de chegada em bytes por segundo (BPS) e do processo de tamanhos médios de pacotes (TMP), calculadas para os dados originais e para o tráfego gerado segundo os modelos inferidos. Também foi analisado o comportamento do tráfego quando sujeito a uma fila de espera,

comparando o rácio de pacotes perdidos (RPP) e o atraso médio na fila de espera (AMFE), obtidos por simulação. Foram usados seis tipos de tráfego: (i) o tráfego original, (ii) tráfego gerado de acordo com o modelo SLD, (iii) tráfego gerado de acordo com o modelo SLC, (iv) tráfego gerado de acordo com o modelo SLFMP, (v) tráfego gerado de acordo com uma cascata dupla gaussiana e (vi) tráfego gerado de acordo com uma cascata dupla discreta. Os parâmetros das cascatas usadas para gerar o tráfego do tipo (v) e (vi) foram inferidos segundo os procedimentos descritos na secção 5.5 do capítulo 5, aplicados aos processos de chegadas e de tamanho médio de pacotes. De referir que os resultados estatísticos e de simulação do tráfego gerado a partir dos diferentes modelos inferidos foram calculados considerando um conjunto de 10 réplicas independentes. Foram calculados intervalos de confiança a 95% com base nas 10 réplicas, mas uma vez que o comprimento dos intervalos foi sempre muito reduzido optou-se por não os incluir nas figuras.

Foram analisadas as características multifractais do tráfego segundo o método descrito na secção 2.3.5 do capítulo 2 e já utilizado no capítulo 5. A Figura 6.12 mostra que os processos de BPS da captura UA e do tráfego gerado de acordo com os modelos de sistema-L possuem todas as características de similaridade escalar multifractal não trivial. Uma análise semelhante foi efectuada para o processo de TMP (Figura 6.14), tendo-se verificado também a existência de características de similaridade escalar multifractal não trivial. O modelo SLC conseguiu replicar com bastante precisão as componentes multiescalares (ou seja, a curva de h_q em função dos momentos) de ambos os processos de tráfego analisados. O modelo SLD conseguiu modelar também as componentes multiescalares de ambos os processos apesar de o fazer com um menor grau de aproximação. Com o SLFMP foi possível apenas modelar aproximadamente as componentes multiescalares do processo BPS, sendo que as características multiescalares do processo TMP não foram capturadas. Este resultado é explicado pelo facto do modelo SLFMP apenas modelar as estatísticas de primeira ordem do processo TMP. Na Figura 6.13 pode-se verificar que o processo BPS do tráfego gerado a partir dos modelos baseados em cascatas não possui características de similaridade escalar multifractal não trivial. No que respeita ao processo de TMP, pode-se observar na Figura 6.15 que o tráfego da cascata discreta não possui características de similaridade escalar multifractal não trivial. O tráfego da cascata gaussiana possui características de similaridade escalar multifractal não trivial mas estas são relativamente diferentes das do tráfego original.

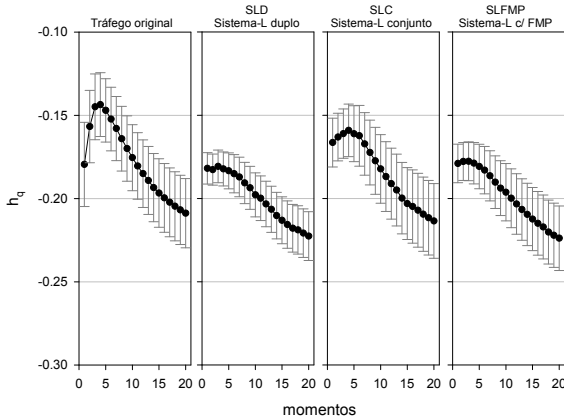


Figura 6.12: Diagramas lineares multiescalares com sistemas-L, processo BPS, UA.

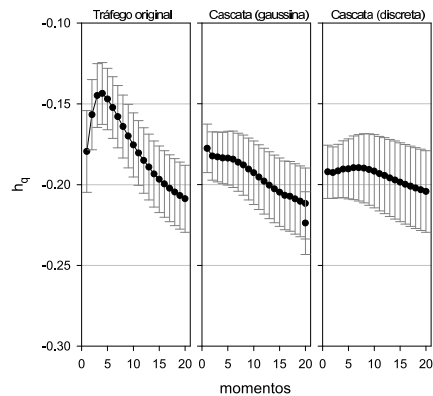


Figura 6.13: Diagramas lineares multiescalares com cascatas conservativas, processo BPS, UA.

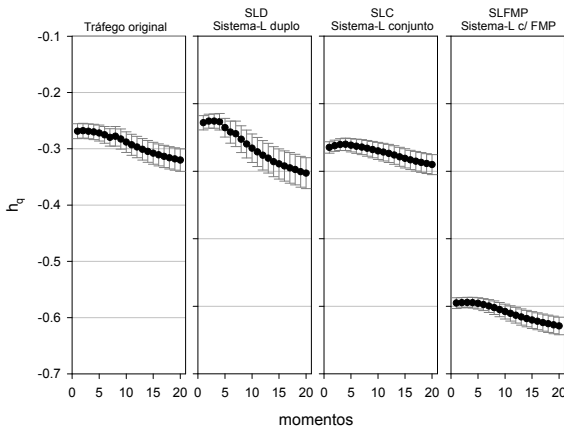


Figura 6.14: Diagramas lineares multiescalares com sistemas-L, processo TMP, UA.

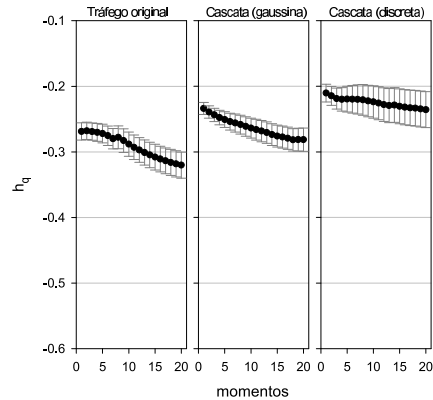


Figura 6.15: Diagramas lineares multiescalares com cascatas conservativas, processo TMP, UA.

Uma análise similar foi efectuada à captura da Bellcore (Figuras 6.16 e 6.18), revelando que este tráfego possui características de similaridade escalar multifractal não trivial quando se analisam os processos TMP e BPS. De notar no entanto que o processo BPS possui características de similaridade escalar multifractal não trivial pouco acentuadas estando no limiar da similaridade trivial (alinhamento horizontal de h_q). O modelo SLC e o modelo SLD obtiveram muito bons resultados no ajuste de ambos os processos de tráfego analisados, com relevância mais uma vez para o primeiro modelo. No que respeita ao modelo SLFMP verifica-se mais uma vez a dificuldade deste modelo em capturar as características multiescalares do processo TMP. Da Figura 6.17 verifica-se que os modelos baseados em cascatas geraram um processo de chegadas em bytes por segundo com características multiescalares completamente diferentes das do tráfego original. Da análise do processo TMP (Figura 6.19) verifica-se que neste caso as cascatas foram capazes de capturar com uma boa precisão as características multiescalares do tráfego original.

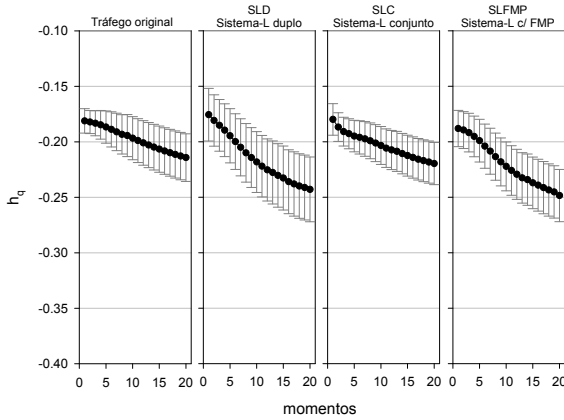


Figura 6.16: Diagramas lineares multiescalares com sistemas-L, processo BPS, Bellcore.

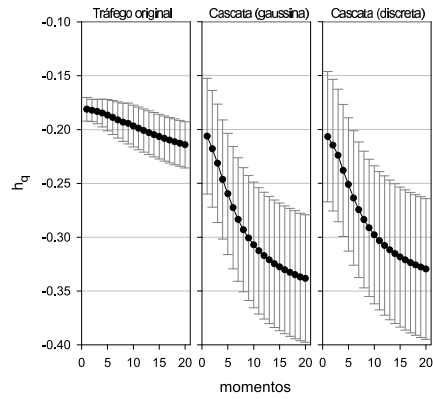


Figura 6.17: Diagramas lineares multiescalares com cascatas conservativas, processo BPS, Bellcore.

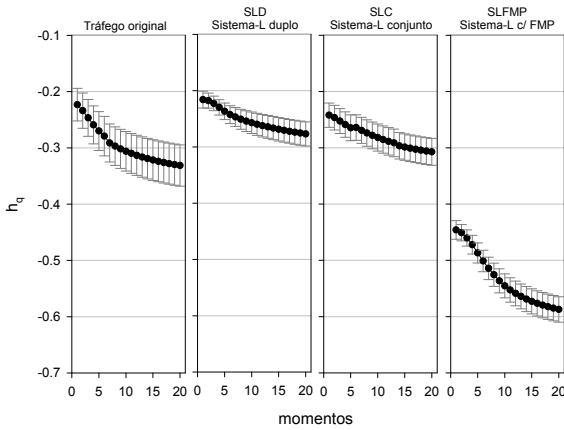


Figura 6.18: Diagramas lineares multiescalares com sistemas-L, processo TMP, Bellcore.

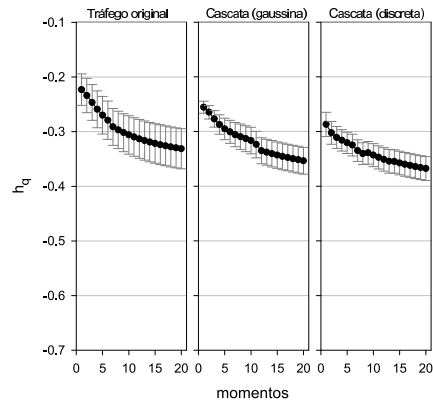


Figura 6.19: Diagramas lineares multiescalares com cascatas conservativas, processo TMP, Bellcore.

Na Tabela 6.1 e na Tabela 6.2, estão resumidos os resultados qualitativos relativos ao ajuste das características multiescalares (aproximação da curva h_q em função dos momentos), conseguidos pelos diferentes modelos aqui estudados.

Da análise das estatísticas de primeira e segunda ordem do processo de BPS da captura Bellcore, pode verificar-se na Figura 6.20 que o ajuste da FMP por parte do modelo SLC foi superior ao dos restantes modelos, incluindo os obtidos pelos modelos baseados em cascatas (Figura 6.22). No ajuste da FAC do processo de BPS (Figuras 6.21 e 6.23), o modelo SLC conseguiu igualmente melhores resultados. Estes resultados podem ser explicados pelo facto de o SLC ser o único modelo em que o procedimento de inferência de parâmetros ajusta conjuntamente o processo de chegadas e o tamanho dos pacotes. Embora o modelo SLFMP correlacione o processo de chegadas com o processo do tamanho dos pacotes, a correlação entre os tamanhos dos pacotes não é capturada (Figura 6.25), resultando num ajuste defeituoso do processo conjunto de chegadas e

Tráfego	Sistema-L duplo	Sistema-L conjunto	Sistema-L c/ FMP	Cascata (gaussiana)	Cascata (discreta)
Bellcore UA	aproximado aproximado	muito bom muito bom	aproximado aproximado	mau deficiente	mau deficiente

Tabela 6.1: Resumo da comparação qualitativa dos resultados relativos ao ajuste das características multiescalares do processo de chegada em bytes por segundo.

Tráfego	Sistema-L duplo	Sistema-L conjunto	Sistema-L c/ FMP	Cascata (gaussiana)	Cascata (discreta)
Bellcore UA	muito bom muito bom	muito bom muito bom	mau mau	muito bom aproximado	muito bom aproximado

Tabela 6.2: Resumo da comparação qualitativa dos resultados relativos ao ajuste das características multiescalares do processo do tamanho médio de pacote.

tamanho de pacotes (processo de BPS). O ajuste da FMP do processo TMP (Figura 6.24) feito pelo SLFMP é inferior ao dos restantes modelos. Este resultado pode ser explicado pelo facto de o SLFMP não ajustar directamente o processo TMP mas sim o tamanho individual dos pacotes. Os restantes modelos baseados em sistemas-L fizeram todos um ajuste relativamente bom da FMP e da FAC do processo de TMP (Figuras 6.24 e 6.25). Quanto aos modelos baseados em cascatas, estes fizeram um ajuste defeituoso da FMP (Figura 6.26) mas conseguiram fazer um bom ajuste da FAC do processo TMP (Figura 6.27). Apesar dos modelos SLD e baseados em cascatas terem capturado as correlações individuais dos processos de chegada de pacotes (Figura 5.18) e de tamanho de pacotes (Figuras 6.25 e 6.27), o facto de não existir qualquer modelação da correlação existente entre os dois processos explica os piores resultados do ajuste das estatísticas de segunda ordem do processo BPS (Figuras 6.21 e 6.23). O modelo que pior se comportou foi o SLFMP, o qual produziu um processo de BPS com uma autocovariância muito inferior à do tráfego original, revelando que existe uma elevada correlação entre os tamanhos dos pacotes que deve ser modelada.

A análise das estatísticas de primeira e segunda ordem da captura UA mostra, que de uma forma geral, mais uma vez o modelo SLC tem os melhores resultados no ajuste do processo de BPS (Figuras 6.28, 6.29, 6.30 e 6.31). A maior parte das observações e conclusões relativas à captura Bellcore aplicam-se igualmente à captura UA. Uma das diferenças é o desempenho muito superior dos modelos no ajuste da FAC do processo de BPS. Face às observações anteriores, este resultado pode ser explicado por uma menor correlação entre os processos de chegada e tamanho dos pacotes. Outra das diferenças relativamente à captura Bellcore é o ajuste superior das estatísticas de primeira e segunda ordem do processo de TMP, por parte dos modelos baseados em cascatas

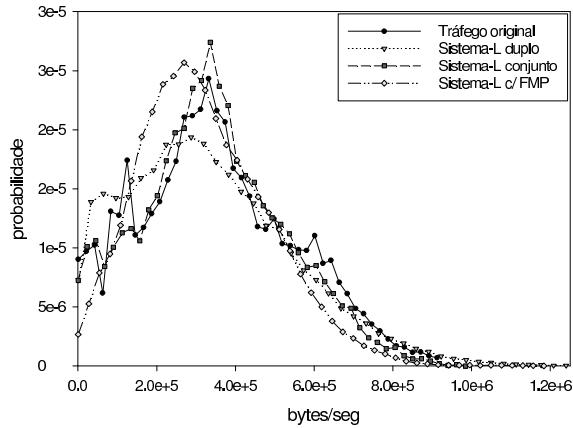


Figura 6.20: Ajuste da função massa de probabilidade com sistemas-L, processo BPS, Bellcore.

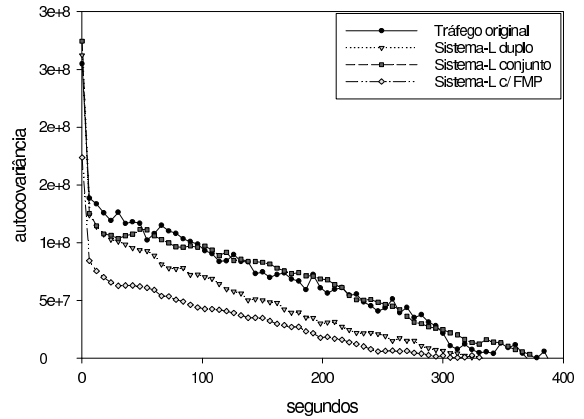


Figura 6.21: Ajuste da função de autocovariância com sistemas-L, processo BPS, Bellcore.

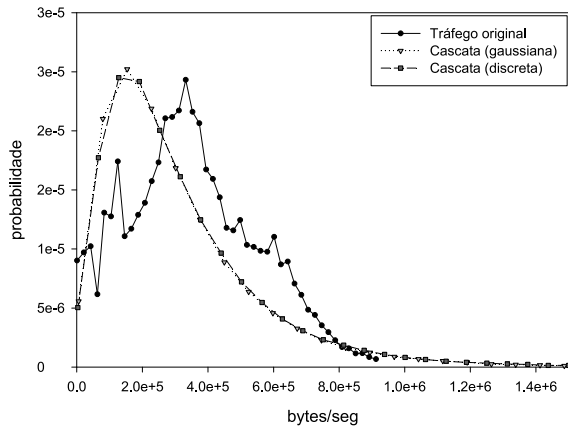


Figura 6.22: Ajuste da função massa de probabilidade com cascatas conservativas, processo BPS, Bellcore.

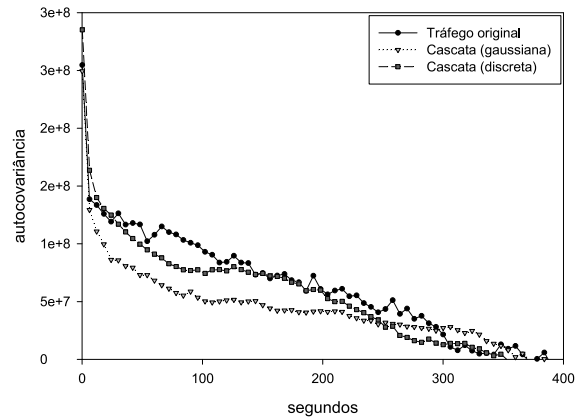


Figura 6.23: Ajuste da função de autocovariância com cascatas conservativas, processo BPS, Bellcore.

(Figuras 6.34 e 6.35). Estas diferenças poderão ser explicadas pelas características inerentes ao gerador das cascatas que tendem a conseguir melhores ajustes quando as FMP são simétricas em relação à média.

De modo a avaliar o comportamento da captura Bellcore quando submetida a uma fila de espera, fez-se variar a unidade de armazenamento da fila de espera de 10 Kbytes até 8 Mbytes para uma taxa de serviço de 518 Kbytes/s (que corresponde a um factor de utilização de 0.7). Fez-se ainda variar a unidade de armazenamento da fila de espera de 100 Kbytes até 70 Mbytes para uma taxa de serviço de 403 Kbytes/s (que corresponde a um factor de utilização de 0.9). Para a captura UA utilizaram-se factores de utilização de 0.9 e 0.98, os quais correspondem respectivamente a taxas de serviço de 726 Kbytes/s e 666 Kbytes/s. Em ambos os casos fez-se variar a unidade de armazenamento da fila de espera de 10 Kbytes até 40 Mbytes.

Começando a análise dos resultados das simulações pela captura UA, podem-se re-

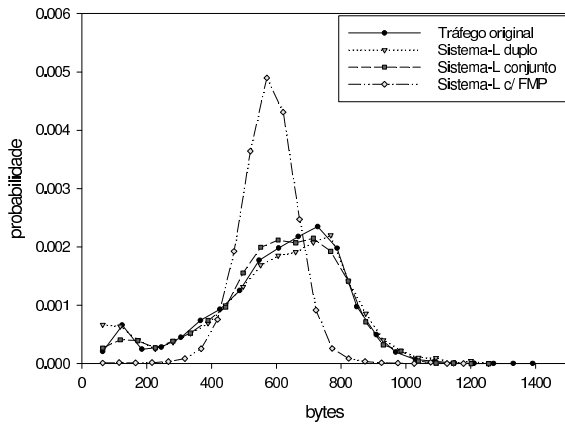


Figura 6.24: Ajuste da função massa de probabilidade com sistemas-L, processo TMP, Bellcore.

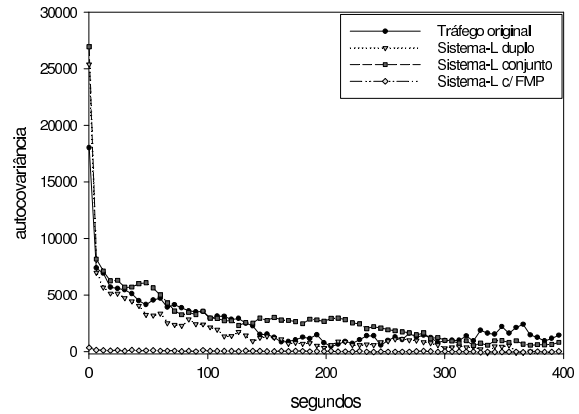


Figura 6.25: Ajuste da função de autocovariância com sistemas-L, processo TMP, Bellcore.

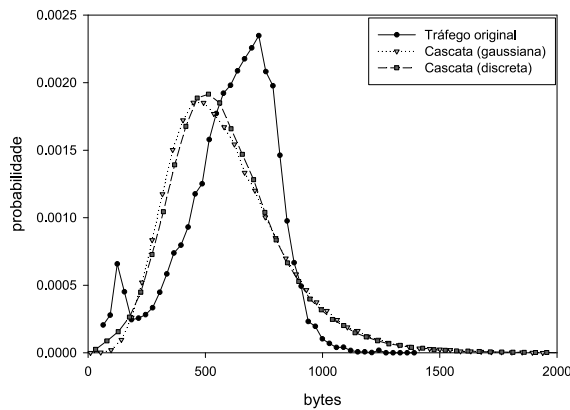


Figura 6.26: Ajuste da função massa de probabilidade com cascatas conservativas, processo TMP, Bellcore.

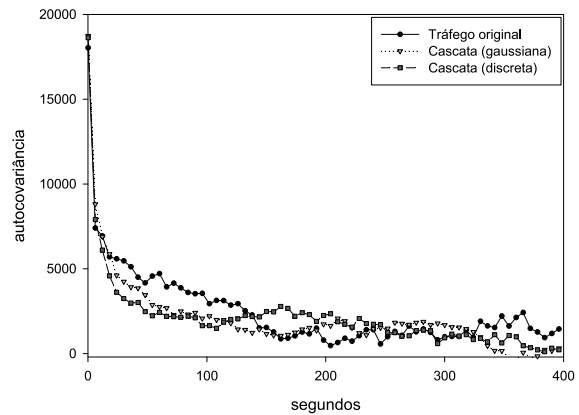


Figura 6.27: Ajuste da função de autocovariância com cascatas conservativas, processo TMP, Bellcore.

tirar várias conclusões. O modelo que obteve resultados mais próximos dos do tráfego original foi o SLC. Este modelo conseguiu resultados muito bons para ambas as taxas de utilização tanto para o RPP como para o AMFE (Figuras 6.36, 6.37, 6.40 e 6.41). Os modelos SLD e SLFMP apresentam resultados muito semelhantes entre eles, mas falharam na captura do comportamento do tráfego original, no que respeita ao RPP, para unidades de armazenamento da fila de espera pequenas, sendo este facto mais notório para a menor taxa de utilização (Figura 6.36). Este facto pode ser explicado pela falta de modelação da correlação do processo TMP no caso do SLFMP e da correlação entre os processos de chegada e de tamanho dos pacotes no caso do SLD. Num sistema de fila de espera quanto menor for o tamanho da unidade de armazenamento da fila de espera mais susceptível é o sistema à existência de correlação nos processos de chegada. A reduzida ocupação média e menor tempo de espera médio na fila de espera não permite ao sistema reduzir o impacto da correlação dos processos. Assim

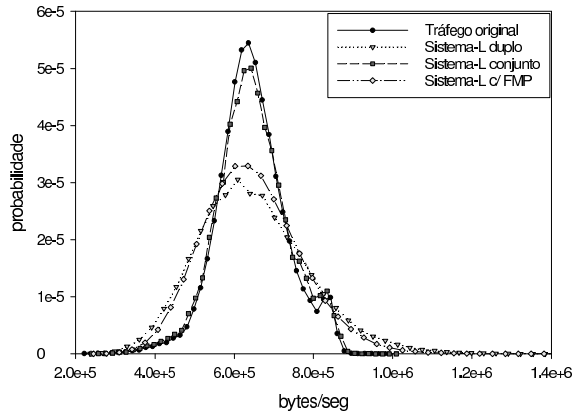


Figura 6.28: Ajuste da função massa de probabilidade com sistemas-L, processo BPS, UA.

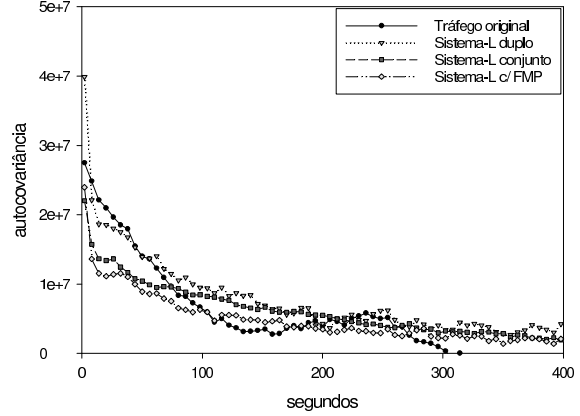


Figura 6.29: Ajuste da função de autocovariância com sistemas-L, processo BPS, UA.

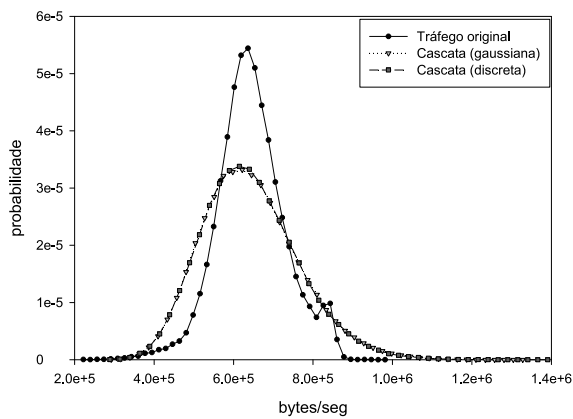


Figura 6.30: Ajuste da função massa de probabilidade com cascatas conservativas, processo BPS, UA.

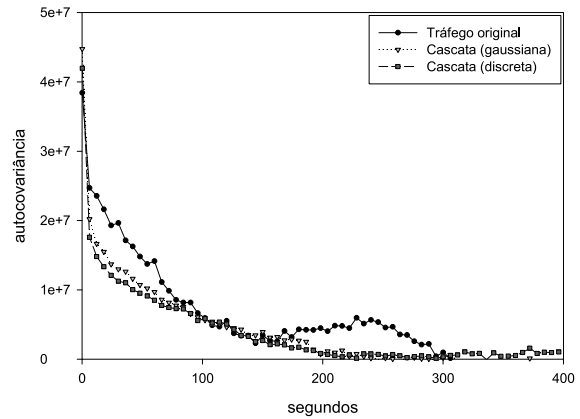


Figura 6.31: Ajuste da função de autocovariância com cascatas conservativas, processo BPS, UA.

nestas condições, os modelos que não conseguem modelar algumas das componentes de correlação obtêm piores resultados. Verifica-se igualmente que os resultados obtidos para o AMFE com o tráfego gerado por estes modelos (SLD e SLFMP) para a taxa de utilização de 90% foram decepcionantes (Figura 6.37), apesar de para uma taxa de utilização mais elevada estes modelos se terem mostrado eficazes (Figura 6.41). Mais uma vez este resultado pode ser explicado pela falta de modelação de uma componente de correlação. Para uma taxa de utilização superior a ocupação média da fila de espera é superior, fazendo com que a influência da correlação do tráfego original diminua. Como consequência os resultados obtidos na fila de espera pelo tráfego original e pelo tráfego gerado pelos modelos SLD e SLFMP aproximam-se.

Os modelos baseados em cascatas obtiveram resultados de RPP muito maus (Figuras 6.38 e 6.42), mas os obtidos para o AMFE foram surpreendentemente bons, no caso de 98% de utilização (Figura 6.43) e para unidades de armazenamento da fila de

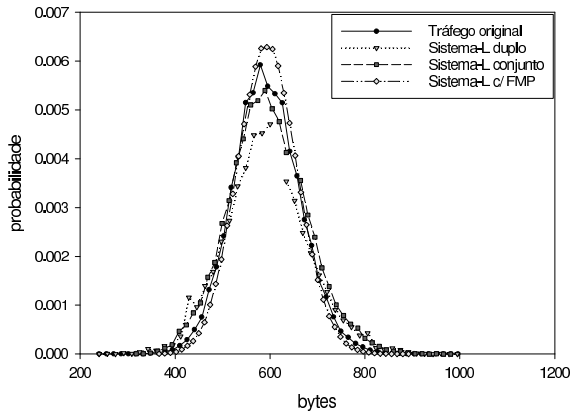


Figura 6.32: Ajuste da função massa de probabilidade com sistemas-L, processo TMP, UA.

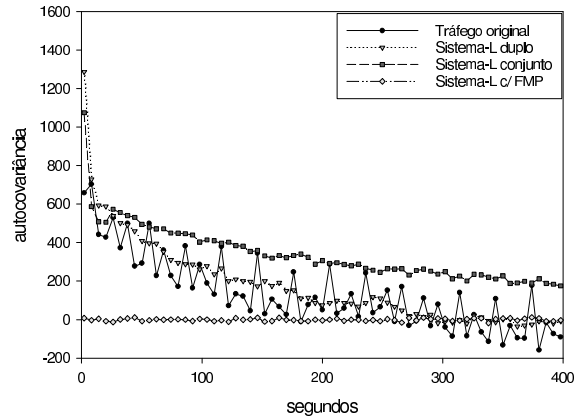


Figura 6.33: Ajuste da função de autocovariância com sistemas-L, processo TMP, UA.

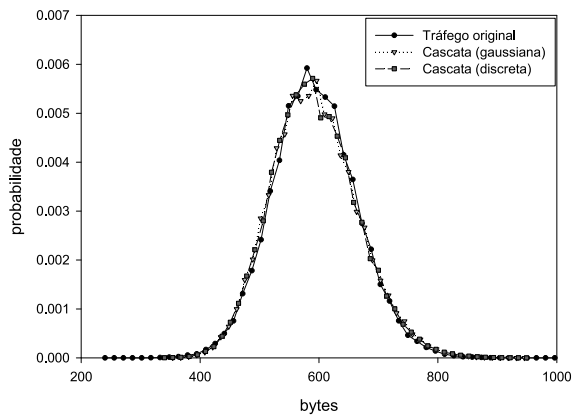


Figura 6.34: Ajuste da função massa de probabilidade com cascatas conservativas, processo TMP, UA.

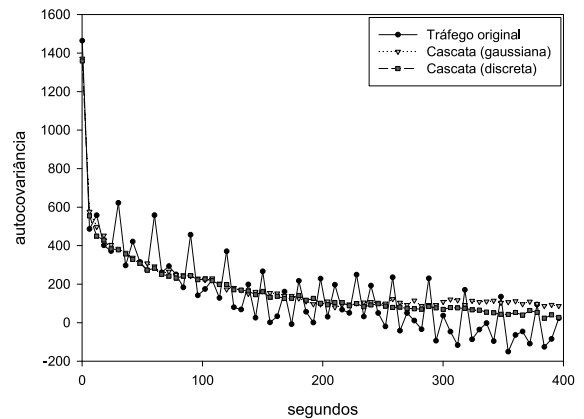


Figura 6.35: Ajuste da função de autocovariância com cascatas conservativas, processo TMP, UA.

espera mais pequenas com uma utilização de 90% (Figura 6.39).

No que respeita aos resultados de simulação da captura Bellcore, mais uma vez se verifica que o modelo SLC conseguiu reproduzir de forma precisa o comportamento do tráfego original, quer se analise o RPP ou o AMFE (Figuras 6.44, 6.45, 6.48 e 6.49). Também os modelo SLD e SLFMP obtiveram bons resultados de RPP, apesar de haver algumas discrepâncias em relação aos resultados obtidos com o tráfego original, para valores da unidade de armazenamento mais altas a uma taxa de utilização de 90% (Figura 6.48) e para valores menores da unidade de armazenamento a uma taxa de 70% (Figura 6.44). Em relação ao AMFE todos os modelos baseados em sistemas-L obtiveram resultados semelhantes e bons, se bem que o modelo SLC seja ligeiramente melhor que os restantes modelos (Figuras 6.45 e 6.49). Com os modelos baseados em cascatas obtiveram-se resultados de AMFE muito próximos dos que se conseguiram com os modelos SLD, SLC e SLFMP (Figuras 6.47 e 6.51). No que respeita ao RPP, os

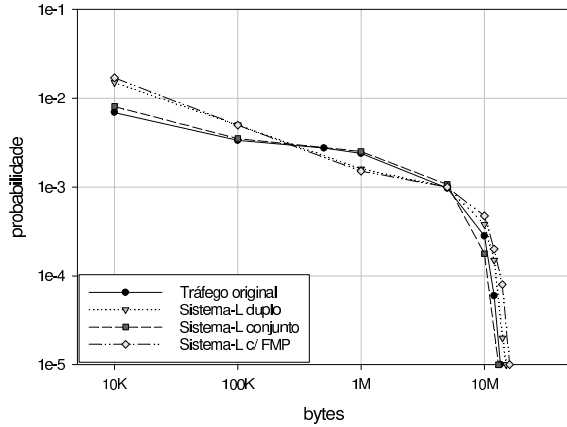


Figura 6.36: Rácio de perda de pacotes com sistemas-L e uma utilização de 90%, UA.

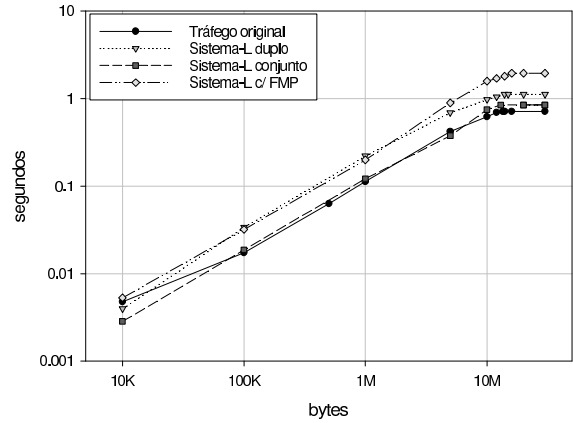


Figura 6.37: Atraso médio na fila de espera com sistemas-L e uma utilização de 90%, UA.

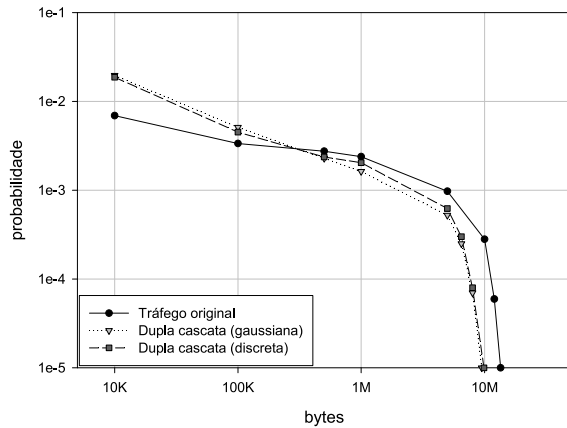


Figura 6.38: Rácio de perda de pacotes com cascatas conservativas e uma utilização de 90%, UA.

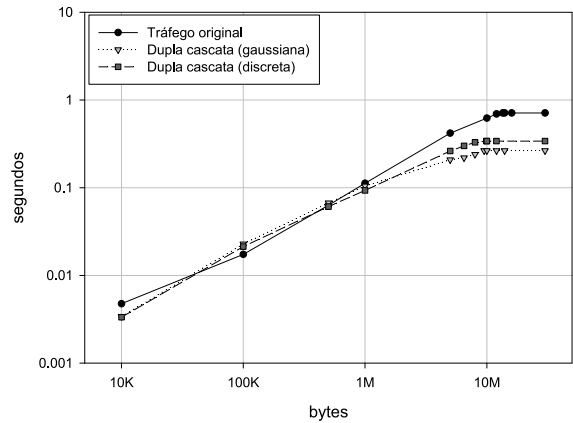


Figura 6.39: Atraso médio na fila de espera com cascatas conservativas e uma utilização de 90%, UA.

modelos com base em cascatas obtiveram resultados muito inferiores aos conseguidos com os sistemas-L (Figuras 6.46 e 6.50). Nota-se, no entanto, que para a captura Bellcore estes modelos conseguiram muito melhores resultados do que para a captura UA. Isto pode ser explicado por dois factores: (i) a captura da Bellcore era mais pequena (cerca de 7 vezes) do que a captura UA e (ii) o processo de BPS da captura Bellcore tem características de similariedade escalar multifractal não trivial menos acentuadas que a captura UA.

6.6 Conclusões

Neste capítulo foram apresentados três novos modelos de tráfego baseados em sistemas-L com capacidade de modelar simultaneamente os processos de chegada e do tamanho dos pacotes, designados por sistema-L duplo, sistema-L conjunto e sistema-L com

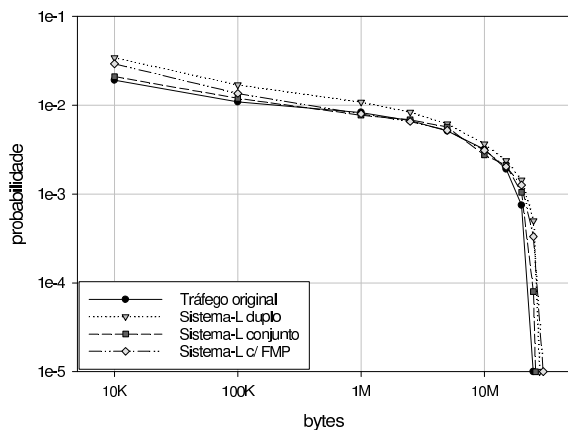


Figura 6.40: Rácio de perda de pacotes com sistemas-L e uma utilização de 98%, UA.

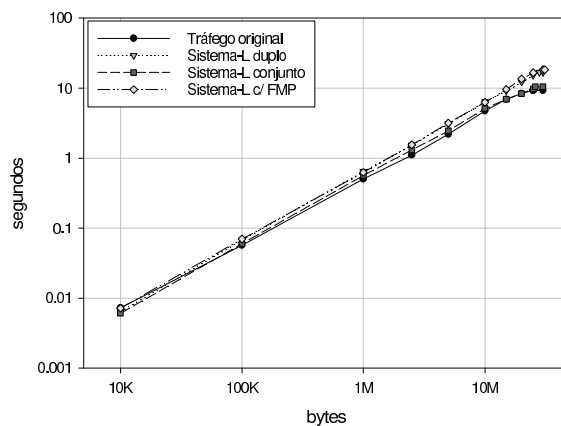


Figura 6.41: Atraso médio na fila de espera com sistemas-L e uma utilização de 98%, UA.

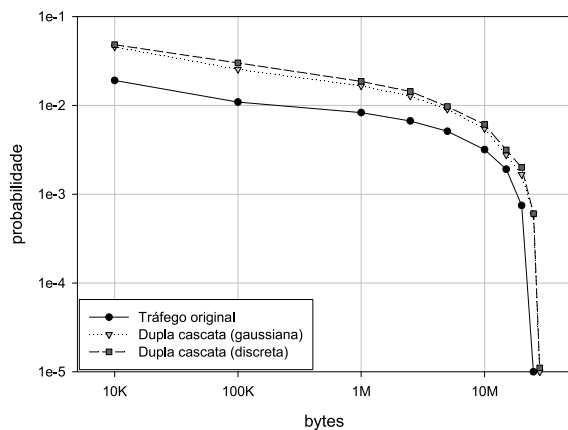


Figura 6.42: Rácio de perda de pacotes com cascatas conservativas e uma utilização de 98%, UA.

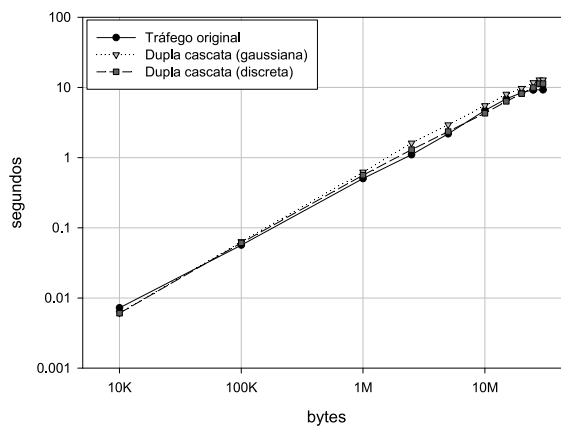


Figura 6.43: Atraso médio na fila de espera com cascatas conservativas e uma utilização de 98%, UA.

funções massa de probabilidade. Estes modelos conseguem modelar com diferentes graus de aproximação as componentes fundamentais de tráfego com tamanho de pacotes variável: estatísticas de primeira e segunda ordem e propriedades multiescalares, ao nível dos processos de chegada de pacotes e de tamanho de pacotes.

O sistema-L duplo (SLD) é constituído por dois sistemas-L independentes, em que um modela o ritmo de chegada de pacotes e o outro o tamanho dos pacotes. O sistema-L conjunto (SLC) é um sistema-L bi-dimensional onde o ritmo de chegada de pacotes e o tamanho dos mesmos são modelados conjuntamente. O sistema-L com funções massa de probabilidade (SLFMP) conjuga um sistema-L único, que modela a chegada dos pacotes, com um conjunto de distribuições para o tamanho dos pacotes na escala temporal mais fina, dependentes da taxa de chegadas.

Os modelos apresentados neste capítulo, bem como dois modelos de cascata conservativa dupla, foram aplicados a capturas de tráfego real com propriedades multi-

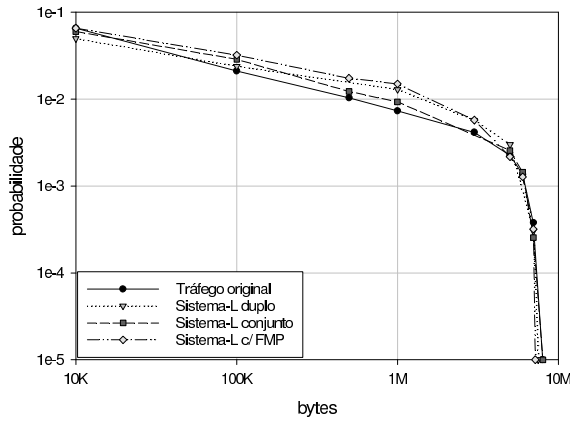


Figura 6.44: Rácio de perda de pacotes com sistemas-L e uma utilização de 70%, Bellcore.

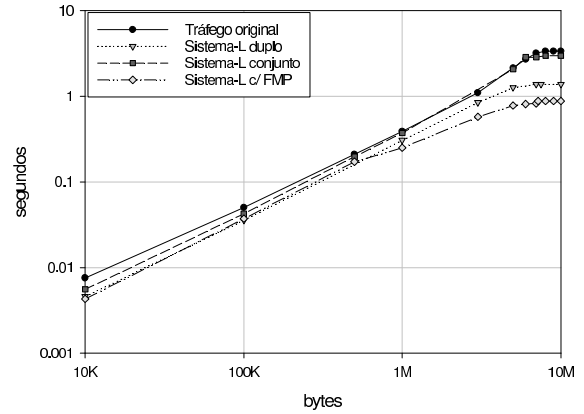


Figura 6.45: Atraso médio na fila de espera com sistemas-L e uma utilização de 70%, Bellcore.

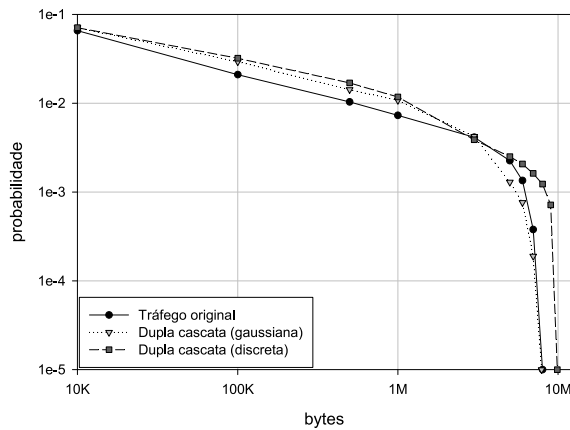


Figura 6.46: Rácio de perda de pacotes com cascatas conservativas e uma utilização de 70%, Bellcore.

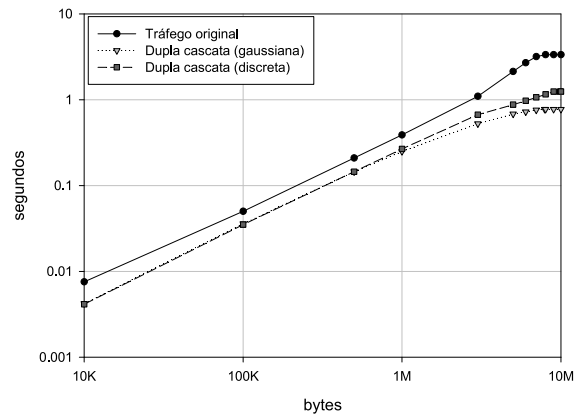


Figura 6.47: Atraso médio na fila de espera com cascatas conservativas e uma utilização de 70%, Bellcore.

fractais não triviais. A partir dos resultados obtidos foi possível concluir que o SLC obteve resultados muito bons, que se verificou serem sempre superiores aos dos restantes modelos. O SLC tem a vantagem de conseguir modelar todas as componentes de correlação do tráfego, nomeadamente, autocorrelação do ritmo de chegadas, autocorrelação do tamanho médio dos pacotes e correlação entre o processo de chegadas e o tamanho dos pacotes. Os resultados obtidos com o SLD e o SLFMP são equivalentes e, apesar de serem inferiores aos do SLC, são razoáveis. O SLD não é capaz de capturar a correlação entre as chegadas e os tamanhos dos pacotes enquanto que o SLFMP não consegue capturar a autocorrelação do tamanho dos pacotes. Em geral, verificou-se que todos os modelos com base em sistemas-L obtiveram resultados muito melhores dos que os obtidos com os modelos de cascata dupla conservativa.

Dos resultados obtidos pode-se concluir que um modelo que vise um ajuste preciso das características do tráfego deve ser capaz de ajustar as propriedades multiescala-

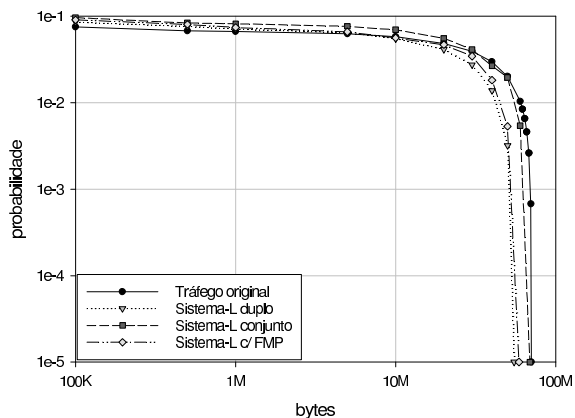


Figura 6.48: Rácio de perda de pacotes com sistemas-L e uma utilização de 90%, Bellcore.

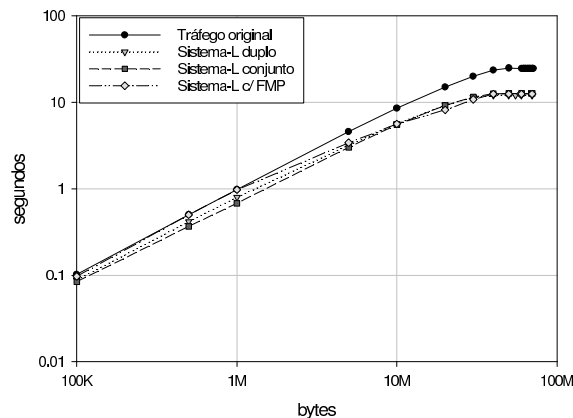


Figura 6.49: Atraso médio na fila de espera com sistemas-L e uma utilização de 90%, Bellcore.

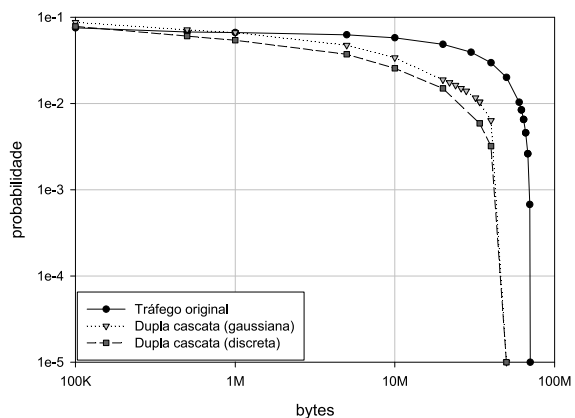


Figura 6.50: Rácio de perda de pacotes com cascatas conservativas e uma utilização de 90%, Bellcore.

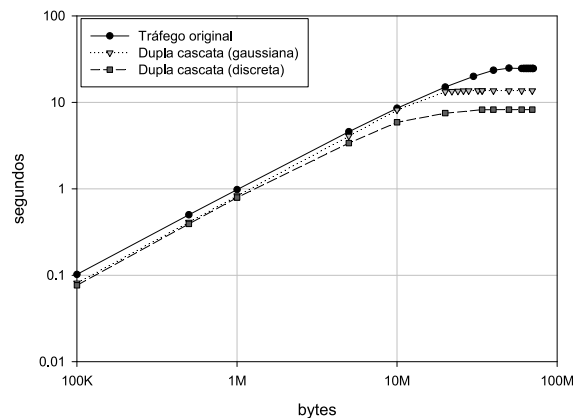


Figura 6.51: Atraso médio na fila de espera com cascatas conservativas e uma utilização de 90%, Bellcore.

res do tráfego e as três componentes de correlação: (i) autocorrelação do ritmo de chegadas, (ii) autocorrelação do tamanho médio dos pacotes e (iii) correlação entre o processo de chegadas e o tamanho dos pacotes. O SLC é um modelo que para além da sua versatilidade na modelação das propriedades multiescalares do tráfego é capaz de capturar todas estas componentes de correlação. Outro factor importante a considerar na escolha de um modelo de tráfego é a complexidade computacional tanto na inferência dos parâmetros como no pós-processamento do mesmo. A escolha de um modelo de tráfego é um compromisso entre a precisão desejada e a complexidade computacional do modelo. Perante os resultados obtidos, em circunstâncias onde uma elevada precisão no ajuste não é um factor determinante o uso dos modelos SLD e SLFMP pode ser preferível devido ao número menor de parâmetros destes.

Capítulo 7

Conclusões

Neste capítulo são apresentadas as principais conclusões do trabalho efectuado e descrito na Tese. São também apresentadas algumas propostas para trabalho futuro.

7.1 Principais conclusões

Esta Tese propôs um conjunto de modelos de tráfego, e respectivos métodos de inferência de parâmetros, com capacidade de capturar as estatísticas de primeira e segunda ordem e as características de similaridade escalar do tráfego (auto-similaridade, dependência longa, monofractalidade e multifractalidade). Foram propostos seis novos modelos de tráfego pertencentes a duas classes distintas: a dos modelos Markovianos e a dos modelos fractais (com base em sistemas de Lindenmayer). Em ambas as classes foram propostos modelos para tráfego com pacotes de tamanho fixo e com pacotes de tamanho variável.

Os modelos Markovianos propostos na Tese têm todos por base o processo de Poisson modulado à Markov (MMPP). Os métodos de inferência propostos revelaram-se capazes de capturar, com elevado grau de precisão, as estatísticas de primeira e segunda ordem do tráfego e de replicar algumas características de similaridade escalar (nomeadamente a dependência longa) numa gama de escalas temporais. Os MMPPs não são processos que possuam intrinsecamente a propriedade da dependência longa. No entanto, quando construídos de forma apropriada conseguem replicar de forma adequada esta propriedade.

O primeiro modelo Markoviano proposto foi um processo de Poisson modulado à Markov em tempo discreto (dMMPP) que caracteriza o número de chegadas em intervalos de tempo fixo. Aplica-se a tráfego com pacotes de tamanho fixo. O modelo

é obtido por sobreposição de L 2-dMMPPs e de um M -dMMPP sem memória. Os 2-dMMPP modelam a função de autocovariância e o M -dMMPP modela a função de probabilidade, dentro das restrições impostas pela modelação da autocovariância. Uma característica importante do modelo é o facto do número de estados do dMMPP ser determinado como parte do procedimento de inferência de parâmetros (não é fixado à priori), o que permite uma melhor adaptação ao tipo de tráfego que se está a caracterizar.

O segundo modelo Markoviano proposto foi uma extensão do modelo anterior, e do procedimento de inferência de parâmetros respectivo, que acrescenta a possibilidade de modelar o tamanho dos pacotes. O modelo é um processo Markoviano de chegadas em rajada em tempo discreto (dBMAP) onde as chegadas de pacotes ocorrem de acordo com um dMMPP e os tamanhos dos pacotes são caracterizados por uma distribuição que depende da fase do dMMPP subjacente. O procedimento de inferência de parâmetros do dMMPP é idêntico ao do modelo anterior. Para inferir a distribuição do tamanho dos pacotes em cada estado são usados histogramas.

Os modelos fractais propostos na Tese têm por base os chamados sistemas de Lindenmayer (sistemas-L), que foram introduzidos em 1968 pelo biólogo A. Lindenmayer para modelar o crescimento das plantas. Um sistema-L gera iterativamente sequências de símbolos progressivamente maiores, a partir de um símbolo inicial, por aplicação sucessiva de regras de produção. Para definir modelos de tráfego baseados em sistemas-L, os símbolos são interpretados como taxas de chegadas ou comprimentos médios de pacotes (em intervalos de comprimento fixo) e cada iteração é associada a uma escala temporal do tráfego. Foram propostos vários modelos de tráfego (e respectivos métodos de inferência dos parâmetros) com diferentes características, que se revelaram capazes de capturar comportamentos monofractais e multifractais do tráfego.

O primeiro modelo fractal proposto caracteriza o número de chegadas em intervalos de tempo fixo, aplicando-se portanto a tráfego com pacotes de comprimento fixo. Baseia-se num único sistema-L em que os símbolos correspondem a taxas de chegada de pacotes. Os restantes modelos fractais são um extensão do anterior e acrescentam a possibilidade de modelar o tamanho dos pacotes. O segundo modelo é formado por dois sistemas-L independentes, em que um modela as chegadas e o outro o tamanho dos pacotes. Neste caso os símbolos de cada sistema-L correspondem, respectivamente, a taxas de chegada e comprimentos médios de pacotes. O terceiro modelo é formado por um único sistema-L em que as chegadas e os tamanhos dos pacotes são modelados

conjuntamente. Neste caso os símbolos correspondem a pares de taxas de chegadas e comprimentos médios de pacotes. O quarto modelo conjuga um sistema-L, que modela as chegadas de pacotes, com distribuições gerais (uma por cada taxa de chegadas) que modelam os tamanhos dos pacotes.

Os modelos propostos nesta Tese foram testados com tráfego medido e foram avaliados comparando (i) as estatísticas de primeira e segunda ordem, (ii) o desempenho sobre uma fila de espera e (iii) as características de similaridade escalar, do tráfego medido e de tráfego gerado a partir dos modelos inferidos. Os resultados obtidos mostram que os modelos propostos são, em geral, capazes de reproduzir de forma bastante rigorosa as principais características do tráfego. Os modelos que obtiveram melhores resultados são os que conseguem capturar não só as estatísticas de primeira e segunda ordem, mas também as características de similaridade escalar (tanto ao nível do processo de chegadas como do tamanho dos pacotes) e a correlação entre os processos de chegadas e de tamanhos de pacotes. Dentro da classe dos modelos Markovianos, o modelo dBMAP conseguiu excelentes resultados, só igualados pelo modelo mais complexo baseado em sistemas de Lindenmayer. Os restantes modelos obtiveram também resultados muito bons, sendo por isso de considerar caso se pretenda modelos com um menor número parâmetros.

Os estudos efectuados nesta Tese permitiram também concluir que, na caracterização de tráfego IP, é importante modelar não apenas o processo de chegadas, mas também o processo de tamanho de pacotes e a correlação entre estes dois processos. Considerar que os pacotes têm todos o mesmo tamanho revela-se uma aproximação muito grosseira. Note-se, no entanto, que os modelos que apenas consideram o processo de chegadas dos pacotes encontram aplicação (i) em tecnologias que utilizem pacotes com tamanho fixo (por exemplo, a tecnologia ATM) e (ii) em IP na modelação de tráfego gerado por aplicações em que um tamanho de pacotes seja predominante (por exemplo, *streaming* de voz e vídeo).

7.2 Sugestões para trabalho futuro

Existem várias linhas de investigação futura que decorrem do trabalho efectuado na Tese.

No âmbito de novos modelos de tráfego, uma hipótese de trabalho futuro passa pela elaboração de modelos de tráfego Markovianos cujo processo de inferência ajuste

directamente as propriedades multiescalares a partir da análise escalar por onduletas, em detrimento do ajuste das escalas temporais da função de autocovariância. Um modelo Markoviano assim construído seria em princípio capaz de ajustar as eventuais características multifractais do tráfego, pelo menos dentro de uma gama limitada de escalas temporais. A modelação de tráfego multifractal através de MMPPs tem como vantagem a possibilidade de usar as ferramentas de avaliação de desempenho em fila de espera já disponíveis na literatura.

O conjunto de modelos baseados em sistemas de Lindenmayer apresentados na Tese constituem uma classe de modelos recente que ainda não foi suficientemente estudada. Em particular, não existem resultados teóricos, por exemplo, para o desempenho em fila de espera e para a largura de banda efectiva. A existência destes resultados viria aumentar substancialmente a flexibilidade de utilização dos modelos.

Os modelos de tráfego desenvolvidos ao longo da Tese, dirigiram-se essencialmente a agregados de tráfego e à explicação dos comportamentos peculiares neles observados. Não foi endereçada de forma directa a modelação do tráfego gerado por aplicações específicas. Também não foi endereçada a modelação do tráfego gerado individualmente pelos utilizadores, nem a forma como o comportamento destes (por exemplo, a preferência por determinadas aplicações) influencia o tráfego na rede. Estes aspectos são importantes para uma compreensão completa do tráfego e para a utilização dos modelos no contexto da previsão de tráfego (a curto e longo prazo).

A previsão de tráfego (a longo prazo) pode ser aplicada no contexto do dimensionamento. Uma abordagem a explorar é a previsão do número de utilizadores e do tipo e peso relativo das aplicações por eles usadas, partindo de uma caracterização de tráfego por utilizador e por aplicação. Preferencialmente, a previsão deve ter em conta os factores sócio-económicos que condicionam o comportamento dos utilizadores, nomeadamente as preferências destes por tipos específicos de aplicações. Esta abordagem permitiria estimar o tráfego agregado futuro com maior precisão, permitindo um dimensionamento da rede mais eficiente.

A previsão de tráfego (a curto prazo) pode ser também aplicada no contexto do controle de admissão com base em medições de tráfego. Os procedimentos de inferência de parâmetros dos modelos propostos na Tese são, por um lado, bastante precisos, e por outro, eficientes do ponto de vista computacional. Deste modo, espera-se que possam conduzir a estimativas de tráfego futuro mais rigorosas que as actualmente consideradas nos métodos de controle de admissão (que recorrem, em geral, a estatísticas sumárias

do tráfego medido em um ou mais intervalos de tempo anteriores) e passíveis de serem calculadas em tempo real, permitindo uma utilização de recursos mais eficiente.

Apêndice A

Capturas de tráfego

A.1 Introdução

Este apêndice tem como objectivo a apresentação dos dados reais usados na avaliação dos diferentes modelos e das ferramentas utilizadas na sua captura. Embora nas diversas publicações relativas ao trabalho apresentado nesta Tese tenham sido usadas uma grande panóplia de capturas de tráfego, nesta Tese optámos por utilizar apenas duas. Esta opção visou permitir uma comparação mais justa entre os diferentes modelos, sem contudo sobrecarregar desnecessariamente a Tese com resultados.

A organização deste apêndice é a seguinte. Na secção A.2 são descritos o equipamento e as ferramentas usados na captura do tráfego e na secção A.3 são apresentadas as capturas propriamente ditas.

A.2 Medições efectuadas

A.2.1 Equipamento

O equipamento escolhido para efectuar as medições de tráfego foi um computador pessoal (PC). Optou-se por este tipo de equipamento, em detrimento de equipamentos comerciais desenvolvidos especificamente para o efeito, devido à pouca versatilidade que os mesmos hoje em dia apresentam em termos de capacidade de armazenamento de dados.

Foi usado um PC equipado com um processador AMD Thunderbird 1.2 GHz e 1.5 GBytes de RAM. O equipamento estava ainda equipado com duas placas de rede e usava o Windows 2000 Professional como sistema operativo.

De referir que não foi capturada informação referente ao campo de dados dos pacotes IP. Este procedimento teve dois objectivos: por um lado poupar espaço de armazenamento em disco e por outro garantir a privacidade dos dados.

A.2.2 Ferramentas de captura e análise

A ferramenta usada para capturar tráfego IP foi o analisador de tráfego WinDump disponível gratuitamente na Internet, no servidor do "Politecnico di Torino"¹. Este software tem por base a arquitectura WinPcap para captura e análise de tráfego. Foi ainda usada como ferramenta auxiliar o Ethereal². O Ethereal é um analisador de tráfego com um interface em janelas tendo como vantagem, em relação ao WinDump, a capacidade de permitir uma fácil descodificação e análise do conteúdo dos pacotes. O Ethereal é gratuito e está disponível para Linux e Windows. Esta ferramenta permite analisar tráfego directamente da rede ou a partir de uma captura guardada em disco. É possível percorrer os dados em causa visualizando sumários da informação ou a informação detalhada de cada um dos pacotes capturados. Esta ferramenta foi por isso usada quando se pretendia analisar particularidades do tráfego capturado.

Com base na arquitectura WinPcap foram desenvolvidas pelo autor algumas ferramentas de software adicionais para facilitar o tratamento dos dados obtidos com o WinDump. Uma dessa ferramentas permite retirar a informação pretendida (instante de chegada e tamanho do pacote) e outra cria ficheiros com dados amostrados (número de chegadas e tamanho médio dos pacotes por intervalo de amostragem).

A.3 Apresentação e caracterização das capturas de tráfego usadas nesta Tese

Ao longo desta Tese são usadas duas capturas na avaliação dos modelos de tráfego. Uma das capturas usadas é a captura Bellcore de Outubro de 1994 [LTWW94], por vezes designada na literatura de "pOct.TL". Esta captura embora relativamente antiga, continua a ser muito usada na avaliação de modelos de tráfego. A captura inclui os instantes de chegada e os tamanhos de pacotes Ethernet. Cerca de 99.5% dos pacotes Ethernet transportavam pacotes IP, sendo por isso possível considerar que o tráfego

¹<http://windump.polito.it/>

²<http://www.ethereal.com/>

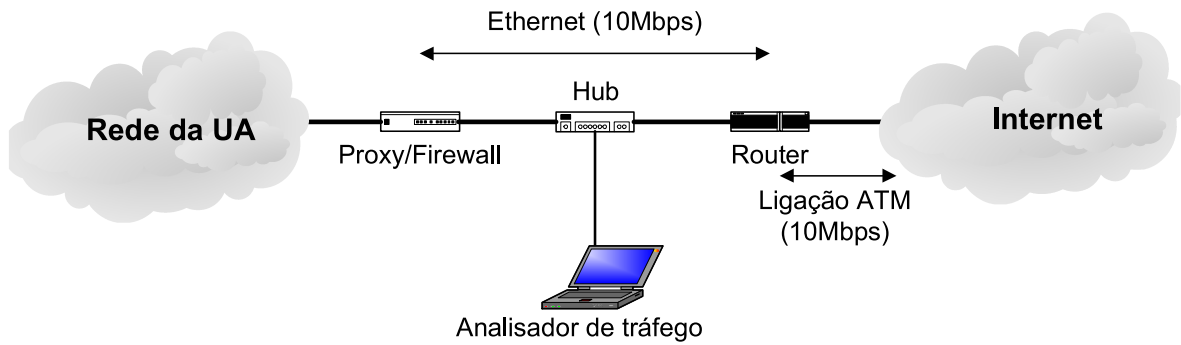


Figura A.1: Esquema das medições na rede da UA.

tem o comportamento típico de uma captura de tráfego IP.

A segunda captura foi medida na Universidade de Aveiro (UA). O cenário da medição está representado na figura A.1. O tráfego foi medido à saída da rede local da UA no troço entre o *proxy* e o *router* fronteira. Foram registados os instantes de chegada e os tamanhos dos pacotes IP que passaram por esse troço da rede, que agrega o tráfego de entrada e saída da rede.

As tabelas A.1 e A.2 apresentam as principais características destas duas capturas: a data de captura, no caso da captura UA também o período de captura, o tipo de pacotes, o tamanho da captura em número de pacotes, a média e a variância do processo de chegadas em bytes por segundo e o tamanho médio dos pacotes.

Nome captura	Período/data de captura	Tipo
Bellcore	5 de Outubro de 1989	Ethernet
UA	12.41pm a 14.27pm, 6 de Julho de 2001	IP

Tabela A.1: Características das capturas usadas.

Nome captura	Tamanho (pacotes)	Ritmo médio (bytes/seg)	Variância (bytes ² /seg ²)	Pacote médio (bytes)
Bellcore	1 milhão	362750	3.97×10^{10}	568
UA	7 milhões	654780	7.51×10^9	600

Tabela A.2: Características das capturas usadas (continuação).

As figuras A.2 e A.3, apresentam as variações temporais dos ritmos de chegada, em pacotes por segundo e em bytes por segundo, da captura Bellcore. É possível observar que, por vezes, as chegadas ocorrem em rajada com amplitude elevada. A variação temporal dos processos de chegada (pacotes e bytes) da captura da UA podem ser observados nas figuras A.4 e A.5.

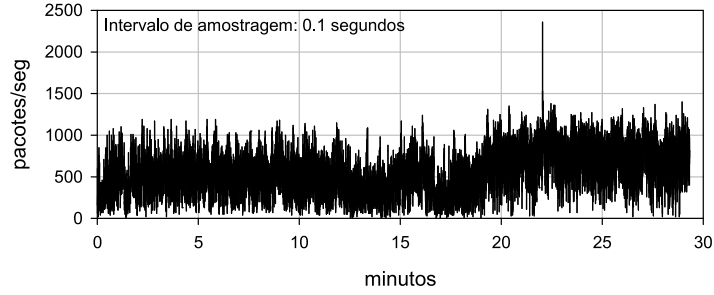


Figura A.2: Variação temporal do ritmo de chegadas em pacotes por segundo da captura da Bellcore.

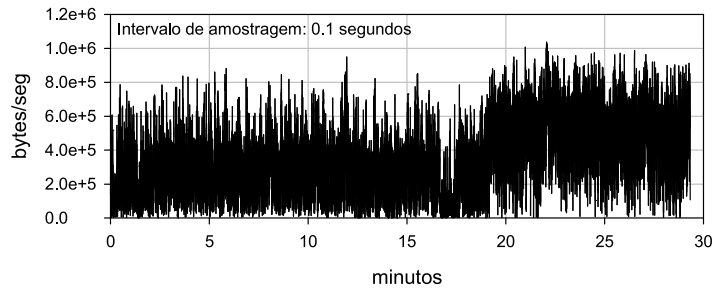


Figura A.3: Variação temporal do ritmo de chegadas em bytes por segundo da captura da Bellcore.

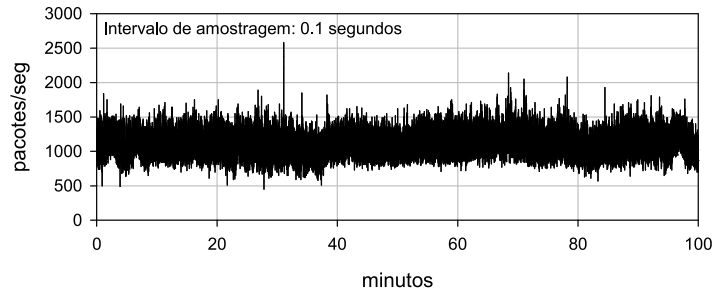


Figura A.4: Variação temporal do ritmo de chegadas em pacotes por segundo da captura da UA.

Nas Figuras A.6, A.7, A.8 e A.9 estão representados os diferentes processos de tráfego para vários níveis de agregação. Para cada uma das capturas e para os processos de chegada de pacotes e de bytes, representaram-se os primeiros 512 pontos das sequências de tráfego com diferentes intervalos de amostragem (0.1, 0.2, 0.4 e 0.8 segundos). Verifica-se que à medida que se aumenta o nível de agregação (maior intervalo de amostragem) a variabilidade do tráfego diminui muito lentamente, mostrando que estamos perante processos de tráfego com características de similaridade escalar. Este aspecto é alvo de um tratamento mais completo no corpo da Tese.

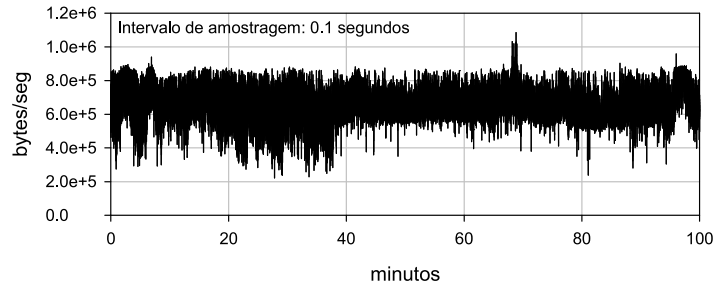


Figura A.5: Variação temporal do ritmo de chegadas em bytes por segundo da captura da UA.

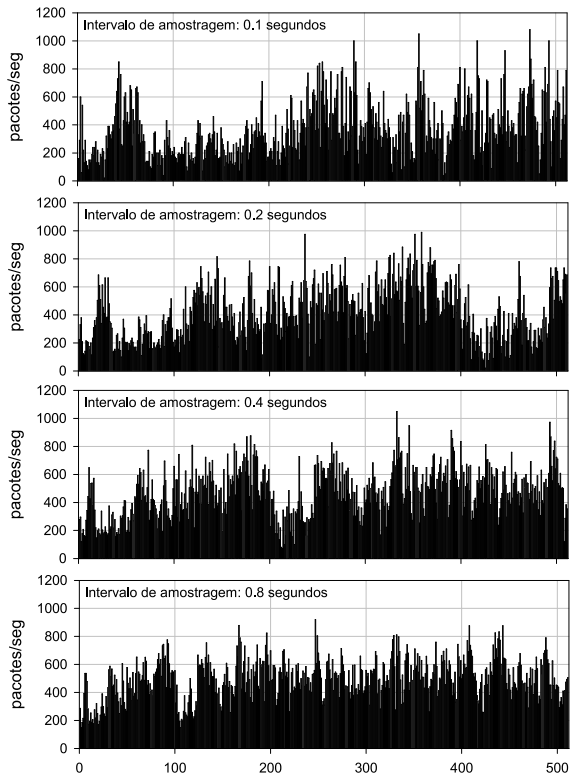


Figura A.6: Representação do processo de chegadas em pacotes por segundo para diferentes níveis de agregação, Bellcore.

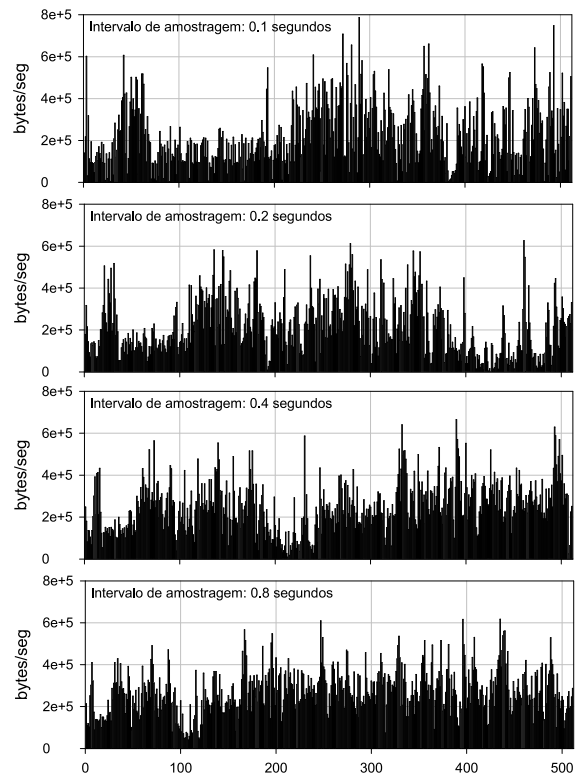


Figura A.7: Representação do processo de chegadas em bytes por segundo para diferentes níveis de agregação, Bellcore.

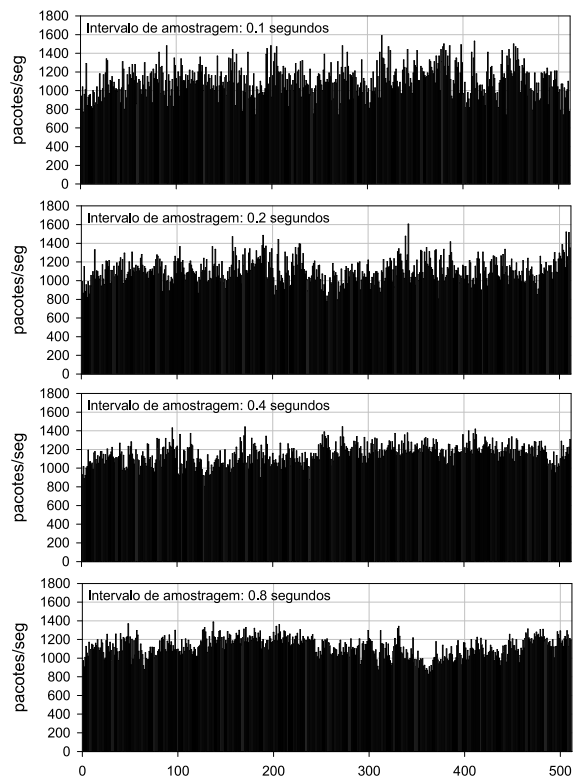


Figura A.8: Representação do processo de chegadas em pacotes por segundo para diferentes níveis de agregação, UA.

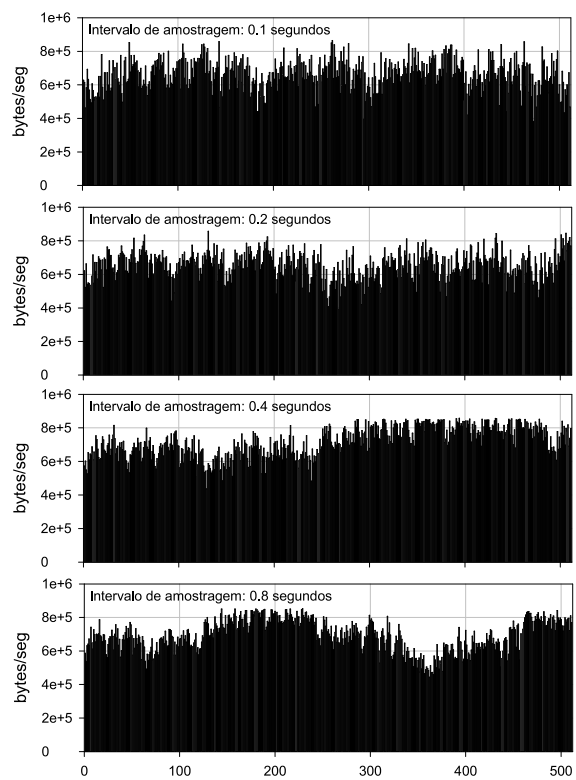


Figura A.9: Representação do processo de chegadas em bytes por segundo para diferentes níveis de agregação, UA.

Referências

- [ABF⁺02] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi e D. Veitch. The multiscale nature of network traffic: Discovery, analysis, and modelling. *IEEE Signal Processing Magazine*, 19(3):28–46, 2002.
- [AC94] A.S. Alfa e S. Chakravarthy. A discrete queue with the Markovian arrival process and phase type primary and secondary services. *Stochastic Models*, 10(2):437–451, 1994.
- [Ada97] A. Adas. Traffic models in broadband networks. *IEEE Communications Magazine*, 35(7):82–89, July 1997.
- [AFTV00] P. Abry, P. Flandrin, M.S. Taqqu e D. Veitch. *Self Similar Network Traffic Analysis and Performance Evaluation*, capítulo Wavelets for the analysis, estimation and synthesis of scaling data, páginas 39–88. Wiley, 2000.
- [AN98] A. Andersen e B. Nielsen. A Markovian approach for modeling packet traffic with long-range dependence. *IEEE Journal on Selected Areas in Communications*, 16(5):719–732, Junho 1998.
- [BC92] C. Blondia e O. Casals. Statistical multiplexing of VBR sources: A matrix-analytic approach. *Performance Evaluation*, 16(1-3):5–20, 1992.
- [BJ76] G. Box e G. Jenkins. *Time Series Analysis: Forecasting and Control*. Prentice-Hall, 1976.
- [Blo93] C. Blondia. A discrete-time batch Markovian arrival process as B-ISDN traffic model. *Belgian Journal of Operations Research, Statistics and Computer Science*, 32:3–23, 1993.

- [CB97] M. Crovella e A. Bestavros. Self-similarity in World Wide Web traffic: Evidence and possible causes. *IEEE/ACM Transactions on Networking*, 5(6):835–846, Dezembro 1997.
- [DMM02] T. Dang, S. Molnár e I. Maricza. Capturing the complete multifractal characteristics of network traffic. *Proceedings of GLOBECOM 2002*, Novembro 2002.
- [FCM97] A. Fisher, L. Calvet e B. Mandelbrot. Multifractality of Deutsche mark/US dollar exchange rates. *Cowles Foundation Discussion Paper No. 1166*, 1997.
- [FGW98] A. Feldmann, A. Gilbert e W. Willinger. Data networks as cascades: Investigating the multifractal nature of internet WAN traffic. *Proceedings of SIGCOMM*, páginas 42–55, 1998.
- [FW97] A. Feldmann e W. Whitt. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance Evaluation*, 31(3-4):245–279, 1997.
- [GB96] F. Geerts e C. Blondia. Superposition of Markov sources and long range dependence. F.A. Aagesen, H. Botnevik e D. Khakhar, editors, *Information Network and Data Communications: Proceedings of the IFIP/ICCC International Conference on Information Network and Data Communication*. Chapman & Hall, 1996.
- [GR94] R. Grünenfelder e S. Robert. Which arrival law parameters are decisive for queueing system performance. *Proceedings of ITC 14*, 1994.
- [GR01] J. Gao e I. Rubin. Multifractal modeling of counting processes of long-range-dependent network traffic. *Journal of Computer Communications*, 24:1400–1410, 2001.
- [GWF99] A. Gilbert, W. Willinger e A. Feldmann. Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Transactions on Information Theory*, 45(3):971–992, Abril 1999.
- [HL86] H. Heffes e D. Lucantoni. A markov modulated characterization of packetized voice and data traffic and related statistical multiplexer perfor-

- mance. *IEEE Journal on Selected Areas in Communications, Special Issue on Network Performance Evaluation*, SAC-4(6):856–868, 1986.
- [HTL92] D. Heyman, A. Tabatabai e T. Lakshman. Statistical analysis and simulation study of video teletraffic in ATM networks. *IEEE Transactions on Circuits and Systems for Video Technology*, 2:49–59, 1992.
- [JM92a] D. Jagerman e B. Melamed. The transition and autocorrelation structure of tes processes; part I: General theory. *Stochastic Models*, 8(2):193–219, 1992.
- [JM92b] D. Jagerman e B. Melamed. The transition and autocorrelation structure of tes processes; part II: Special cases. *Stochastic Models*, 8(3):499–527, 1992.
- [JMW96] D. Jagerman, B. Melamed e W. Willinger. *Frontiers in Queueing: Models, Methods and Problems*, capítulo Stochastic modeling of traffic processes. CRC Press, 1996.
- [KLL03] A. Klemm, C. Lindemann e M. Lohmann. Modeling IP traffic using the batch Markovian arrival process. *Performance Evaluation*, 54:149–173, 2003.
- [KS95] S. Kang e D. Sung. Two-state MMPP modelling of ATM superposed traffic streams based on the characterisation of correlated interarrival times. *IEEE GLOBECOM'95*, páginas 1422–1426, Novembro 1995.
- [LH97] S. Li e C. Hwang. On the convergence of traffic measurement and queuing analysis: A statistical-match and queuing (SMAQ) tool. *IEEE/ACM Transactions on Networking*, páginas 95–110, Fevereiro 1997.
- [Lin68] A. Lindenmayer. Mathematical models for cellular interactions in development II. Simple and branching filaments with two-sided inputs. *Journal of Theoretical Biology*, 18:300–315, 1968.
- [LMHN90] D. Lucantoni, K. Meier-Hellstern e M. Neuts. A single-server queue with server vacations and a class of non-renewal arrival processes. *Advances in Applied Probability*, 22:676–705, 1990.

- [LTWW94] W. Leland, M. Taqqu, W. Willinger e D. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, Fevereiro 1994.
- [Luc91] D. Lucantoni. New results on the single server queue with a batch Markovian arrival process. *Stochastic Models*, 7(1):1–46, 1991.
- [Luc93] D.M. Lucantoni. The BMAP/G/1 queue: A tutorial. L. Donatiello e R. Nelson, editors, *Models and Techniques for Performance Evaluation of Computer and Communication Systems*, páginas 330–358. Springer Verlag, 1993.
- [Man74] B. Mandelbrot. Intermittant turbulence in self-similar cascades: Divergence of high moments and dimensions of the carrier. *Journal of Fluid Mechanics*, 62:331–358, 1974.
- [Man77] B. Mandelbrot. *Fractal Geometry of Nature*. W.H. Freeman, New York, 1977.
- [MH87] K. Meier-Hellstern. A fitting algorithm for Markov-modulated Poisson process having two arrival rates. *European Journal of Operational Research*, 29, 1987.
- [MN68] B. Mandelbrot e J. Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10:422–437, 1968.
- [MS92] B. Melamed e B. Sengupta. Tes modeling of video traffic. *IEICE Transactions on Communications*, E75-B(12):1291–1300, 1992.
- [NP98] C. Nunes e A. Pacheco. Parametric estimation in MMPP(2) using time discretization. *Proceedings of the 2nd International Symposium on Semi-Markov Models: Theory and Applications*, Dezembro 1998.
- [OS95] M. Osborne e G. Smyth. A modified prony algorithm for fitting sums of exponential functions. *SIAM Journal of Scientific Computing*, 16:119–138, 1995.
- [PJS92] H. Peitgen, H. Jurgens e D. Saupe. *Chaos and Fractals: New Frontiers of Science*. Springer-Verlag, 1992.

- [PW00] K. Park e W. Willinger. Self-similar network traffic: an overview. K. Park e W. Willinger, editors, *Self-Similar Network Traffic and Performance Evaluation*. Wiley-Interscience, 2000.
- [Ran95] N. Rananand. Markov approximations to D-BMAPs: information-theoretic bounds on queueing performance. *Stochastic Models*, 11(4):713–734, 1995.
- [RCRB99] R. Riedi, M. Crouse, V. Ribeiro e R. Baraniuk. A multifractal wavelet model with application to network traffic. *IEEE Transactions on Information Theory*, 45(4):992–1018, Abril 1999.
- [Rid98] Ad Ridder. Fast simulation of discrete time queues with Markov modulated batch arrivals and batch departures. *AEU International Journal of Electronics and Communications*, 52:127–132, 1998.
- [Rie02] R. Riedi. *Long Range Dependence: Theory and Applications*, capítulo Multifractal processes, páginas 625–715. Birkhäuser, 2002.
- [RV97] R. Riedi e J. L. Véhel. Multifractal properties of TCP traffic: a numerical study. *Technical Report No 3129, INRIA Rocquencourt, France*, Fevereiro 1997. Disponível em www.dsp.rice.edu/~riedi.
- [SNV02a] P. Salvador, A. Nogueira e R. Valadas. Modeling multifractal IP traffic: Characterization of packet arrivals and packet sizes using stochastic L-Systems. *Proceedings of the 10th International Conference on Telecommunication Systems, Modeling and Analysis*, páginas 577–587, Outubro 2002.
- [SNV02b] P. Salvador, A. Nogueira e R. Valadas. Modeling multifractal traffic with stochastic L-Systems. *Proceedings of the GLOBECOM 2002*, Novembro 2002.
- [SNV03a] P. Salvador, A. Nogueira e R. Valadas. Framework based on stochastic L-Systems for modeling IP traffic with multifractal behavior. *Proceedings of the SPIE Conference on Performance and Control of Next Generation Communication Networks, ITCOM 2003*, Setembro 2003.

- [SNV03b] P. Salvador, A. Nogueira e R. Valadas. Joint characterization of the packet arrival and packet size processes of multifractal traffic based on stochastic L-Systems. *Proceedings of the 18th International Teletraffic Congress, ITC 18*, Setembro 2003.
- [SNV04] P. Salvador, A. Nogueira e R. Valadas. Framework based on stochastic l-systems for modeling ip traffic with multifractal behavior. *Computer Communications Journal*, 27(18):1799–1811, Outubro 2004.
- [SNVP04] P. Salvador, A. Nogueira, R. Valadas e A. Pacheco. Multi-time-scale traffic modeling using Markovian and L-Systems models. *Lecture Notes in Computer Science*, 3262:297–306, Setembro 2004.
- [SPV02] P. Salvador, A. Pacheco e R. Valadas. Multiscale fitting procedure using Markov modulated Poisson processes. *Proceedings of the 10th International Conference on Telecommunication Systems, Modeling and Analysis*, páginas 304–318, Outubro 2002.
- [SPV03] P. Salvador, A. Pacheco e R. Valadas. Multiscale fitting procedure using Markov modulated Poisson processes. *Telecommunications Systems*, 23(1-2):123–148, Junho 2003.
- [SPV04] P. Salvador, A. Pacheco e R. Valadas. Modeling IP traffic: Joint characterization of packet arrivals and packet sizes using BMAPs. *Computer Networks Journal*, 44(3):335–352, Fevereiro 2004.
- [SSD93] P. Skelly, M. Schwartz e S. Dixit. A histogram-based model for video traffic behaviour in an atm multiplexer. *IEEE/ACM Transactions on Networking*, páginas 446–458, Agosto 1993.
- [SV01a] P. Salvador e R. Valadas. A fitting procedure for Markov modulated Poisson processes with an adaptive number of states. *Proceedings of the 9th IFIP Working Conference on Performance Modelling and Evaluation of ATM & IP Networks*, páginas 62–73, Junho 2001.
- [SV01b] P. Salvador e R. Valadas. Framework based on markov modulated poisson processes for modeling traffic with long-range dependence. Robert D. van der Mei e Frank Huebner-Szabo de Bucs, editors, *Proceedings of the*

- SPIE Conference - Internet Performance and Control of Network Systems II, Vol. 4523*, páginas 221–232, Agosto 2001.
- [VA99] D. Veitch e P. Abry. A wavelet based joint estimator for the parameters of LRD. *IEEE Transactions on Information Theory*, 45(3), Abril 1999.
- [VAFC00] D. Veitch, P. Abry, P. Flandrin e P. Chainais. Infinitely divisible cascade analysis of network traffic data. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Junho 2000.
- [VR97] J. Véhel e R. Riedi. *Fractals in Engineering*, capítulo Fractional Brownian Motion and Data Traffic Modeling: the End of the Spectrum, páginas 185–202. Springer, 1997.
- [YKT01] T. Yoshihara, S. Kasahara e Y. Takahashi. Practical time-scale fitting of self-similar traffic with Markov-modulated Poisson process. *Telecommunication Systems*, 17(1-2):185–211, 2001.