



Marcus Vinícius dos Santos Fernandes

**Modelagem de Canais de Comunicações Digitais
Sujeitos a Erros em Surtos**

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio como parte dos requisitos parciais para obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Prof. Marco Antônio Grivet Mattoso Maia
Co-orientador: Prof. Ernesto Leite Pinto

Rio de Janeiro

Abril de 2010

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



Marcus Vinícius dos Santos Fernandes

**Modelagem de Canais de Comunicações Digitais
Sujeitos a Erros em Surtos**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Marco Antonio Grivet Mattoso Maia
Orientador

Centro de Estudos de Telecomunicações – PUC-Rio

Prof. Ernesto Leite Pinto
Co-Orientador
IME

Prof. Pedro Henrique Gouvêa Coelho
UERJ

Prof. José Mauro Pedro Fortes
Centro de Estudos de Telecomunicações – PUC-Rio

Prof. Raimundo Sampaio Neto
Centro de Estudos de Telecomunicações – PUC-Rio

Prof. Marcelo Roberto Baptista Pereira Luis Jimenez
Centro de Estudos de Telecomunicações – PUC-Rio

Prof. Rodolfo Sabóia Lima de Souza
Inmetro

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico – PUC-Rio

Rio de Janeiro, 16 de abril de 2010

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Marcus Vinícius dos Santos Fernandes

Graduado como Engenheiro em Engenharia Eletrônica pelo Instituto Militar de Engenharia em 1994, e Mestre em Engenharia Elétrica - Processamento de Sinais, pelo Instituto Militar de Engenharia em 2002.

Ficha Catalográfica

Fernandes, Marcus Vinícius dos Santos

Modelagem de canais de comunicações digitais sujeitos a erros em surtos / Marcus Vinícius dos Santos Fernandes ; orientador: Marco Antônio Grivet Mattoso Maia; co-orientador: Ernesto Leite Pinto. – 2010.

103 f. ; 30 cm

Tese (Doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2010.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Erros em surtos. 3. Modelos escondidos de Markov. 4. Canal de comunicações. I. Maia, Marco Antônio Grivet Mattoso. II. Pinto, Ernesto Leite. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Aos meus pais, esposa e filha.

Agradecimentos

Aos Professores Marco Antônio Grivet Mattoso Maia e Ernesto Leite Pinto, pela orientação técnico-científica e apoio no desenvolvimento deste trabalho.

À minha esposa Maria Inês e filha Marina Letícia, pelo incentivo, apoio, compreensão e amor dedicados.

Aos meus pais, Walter Fernandes e Maria Bernadete, pelo amor e pela sólida educação que me proporcionaram.

Aos amigos e colegas de trabalho pelo incentivo.

Aos professores do CETUC, pela amizade e ensinamentos.

Aos funcionários do CETUC, e do Departamento de Engenharia pela amizade e apoio administrativo.

Ao Comando e companheiros do Instituto Militar de Engenharia pelo apoio à realização da presente Tese.

A todos aqueles que de alguma forma contribuíram para a realização desse trabalho.

Resumo

Fernandes, Marcus Vinícius dos Santos; Maia, Marco Antonio Grivet Mattoso (Orientador); Pinto, Ernesto Leite (Co-orientador). **Modelagem de Canais de Comunicações Digitais Sujeitos a Erros em Surto**. Rio de Janeiro, 2010. 103p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

A ocorrência de erros em surto é observada principalmente em canais sem fio. Para a análise e melhor entendimento deste tipo de erro, a fim de se melhorar os projetos de sistemas de comunicações digitais, uma modelagem mais precisa, de canais com esta característica, torna-se necessária. Uma diversidade de métodos de estimação de parâmetros tem sido estudada, principalmente aquelas baseadas nos Modelos Escondidos de Markov (HMM do inglês). Em geral cada método é focado em um sistema de comunicações específico, sobre uma camada específica. Neste trabalho é proposto um novo método baseado em um HMM com uma estrutura particular, que permite a dedução de expressões analíticas para todas as estatísticas de interesse. A estrutura do modelo proposto permite a geração de eventos que ocorrem numa sequência binária de dados sujeita a surtos de erro, de acordo com a definição de surtos de erro do CCITT. O modelo proposto possui um número fixo de apenas sete parâmetros, mas o seu número de estados cresce com um de seus parâmetros, que aumenta a precisão, mas não a complexidade. Este trabalho adotou técnicas de otimização, associadas aos métodos de Máxima Verossimilhança e *Particle Swarm Optimization* (PSO) a fim de realizar a estimação dos parâmetros do modelo proposto. Os resultados demonstram que o modelo proposto permite a caracterização precisa de canais com memória de diversas origens.

Palavras-chave

Erros em surtos, Modelos Escondidos de Markov, canal de comunicações.

Abstract

Fernandes, Marcus Vinícius dos Santos; Maia, Marco Antonio Grivet Mattoso (Advisor); Pinto, Ernesto Leite (Co-advisor). **Modeling of Digital Communication Channels Under Burst of Errors**. Rio de Janeiro, 2010. 103p. Doctorate Thesis – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

The occurrence of error busts is mainly observed in wireless channels. For analysis and a better understanding of such errors, in order to improve the design of communication systems, an accurate modeling of channels with this characteristic is necessary. A lot of parameter estimation methods have been studied, mainly the ones based on Hidden Markov Models (HMM). In general each method is focused in a specific communication system, on a specific layer. In this work it is proposed a new method based on a HMM with particular structure that allows the deduction of analytical expressions for all statistics of interest. The structure of the proposed model permits the generation of events that occur in a binary data sequence subject to bursts of error concerning CCITT error burst definition. The proposed model has a fixed number of only seven parameters but its number of states increase with one of those parameters that increase the accuracy but not the complexity. This work adopted techniques of optimization associated to Maximum Likelihood (ML) and Particle Swarm Optimization (PSO) to perform the parameter estimation to the proposed model. The results show that the proposed model achieves accurate characterization of channels with memory from many different sources.

Keywords

Burst errors, Hidden Markov Models, communication channel.

Sumário

1. Introdução	15
1.1. Modelagem de erros	16
1.2. Surto de Erro	18
1.3. Proposta deste trabalho	19
1.4. Apresentação do trabalho	20
2. Conceitos básicos	21
2.1. Modelos de Markov	21
2.2. Modelos Escondidos de Markov	22
2.3. HMM aplicado a canais sujeitos a erros em surtos	23
2.4. Modelo de Gilbert-Elliot	24
2.5. Modelo de Fritchman	25
2.6. Estimação de parâmetros do HMM	27
2.6.1. Cálculo da probabilidade de geração da sequência observada	27
2.6.2. Determinação da sequência de estados mais provável	30
2.6.3. Estimação ML de Parâmetros	31
2.6.4. Algoritmo de Baum-Welch	32
2.7. Ferramentas de otimização	36
2.7.1. Particle Swarm Optimization (PSO)	37
2.7.2. A técnica “Branch and Bound” (BB)	39
3. Modelo de surtos proposto	42
3.1. Objetivos	42
3.2. Descrição do modelo	42
3.2.1. Parâmetros	44
3.2.2. Estados	45
3.2.3. Matriz de Probabilidades Transição	46
3.2.4. Probabilidades dos estados em regime permanente	46
3.3. Estatísticas de interesse	48

3.3.1. Probabilidade de erro	49
3.3.2. Probabilidades $P[10^m1]$ e $P[01^m0]$	50
3.3.2.1. Cálculo de $P[10^m1]$	50
3.3.2.2. Cálculo de $P[01^m0]$	51
3.3.3. Distribuição de gaps	52
3.3.4. Tamanho médio dos gaps	52
3.3.5. Distribuição dos clusters de erro	54
3.3.6. Tamanho médio dos clusters de erro	54
3.3.7. Distribuição de intervalos entre surtos	55
3.3.8. Distribuição de gaps no interior de surtos	55
3.3.9. Função de Auto-correlação	56
3.3.10. Probabilidade $P[1^m 0]$	58
3.3.11. Probabilidade $P[0^m 1]$	59
4. Metodologia para ajuste dos parâmetros do modelo proposto	61
4.1. Função Verossimilhança do modelo	62
4.2. Metodologia de otimização	65
5. Resultados	66
5.1 Estimação do parâmetro L	68
5.2 Validação do método de ajuste do modelo	70
5.3. Equalizador DFE (caso I)	71
5.4. Canal com desvanecimento Rayleigh (caso II)	74
5.4.1 Canal com desvio Doppler de 10^{-1} e E_b/N_0 de 10 dB	75
5.4.2 Canal com desvio Doppler de 10^{-2} e E_b/N_0 de 15 dB	78
5.4.3 Canal com desvio Doppler de 10^{-3} e E_b/N_0 de 15 dB	82
5.5. Decodificador de Viterbi (caso III)	84
5.5.1. Decodificador com probabilidade de erro de entrada de 0,1	85
5.5.2. Decodificador com probabilidade de erro de entrada de 0,05	88
5.6. Influência do emprego do PSO	91
6. Conclusão	94

7. Referências bibliográficas	96
Apêndice A. Notação matricial	99
Apêndice B. Técnicas de otimização	102

Lista de figuras

Figura 2.1 – Cadeia de Markov de 3 estados.	22
Figura 2.2 – Modelo de Gilbert-Elliott.	25
Figura 2.3 – Modelo de Fritchman-SES de N estados.	26
Figura 3.1 – Diagrama de estados do Modelo Proposto.	43
Figura 3.2 – Diagrama de estados equivalente ao estado e_2	57
Figura 4.1 – Modelo de transição	62
Figura 5.1 – Levantamento do melhor valor do parâmetro L para o modelo proposto para os casos I e II	69
Figura 5.2 – Levantamento do melhor valor parâmetro L para o modelo proposto para o caso III.	69
Figura 5.3 – Resultados de validação do método de ajuste dos parâmetros do modelo proposto.	70
Figura 5.4 – Estimativas da distribuição de surtos de erros para o Caso I.	71
Figura 5.5 – Estimativas da função autocorrelação para o Caso I.	72
Figura 5.6 – Estimativas da distribuição de gaps para o Caso I.	72
Figura 5.7 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso I.	73
Figura 5.8 – Estimativas da função Probabilidade de erros por bloco para um bloco de tamanho 300 ($P(m,300)$) para o Caso I.	74
Figura 5.9 – Estimativas distribuição de surtos de erro para o Caso II.a.	75
Figura 5.10 – Estimativas da distribuição de gaps para o Caso II.a.	76
Figura 5.11 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso II.a.	76
Figura 5.12 – Estimativas da função autocorrelação (BECF) para o Caso II.a.	77

Figura 5.13 – Estimativas da função Probabilidade de erros por bloco de tamanho 300, $P(m,300)$, para o Caso II.a.	78
Figura 5.14 – Estimativas da distribuição de surtos de erro para o Caso II.b.	79
Figura 5.15 – Estimativas da distribuição de gaps para o Caso II.b.	79
Figura 5.16 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso II.b.	80
Figura 5.17 – Estimativas da função autocorrelação (BECF) para o Caso II.b.	81
Figura 5.18 – Estimativas da função Probabilidade de erros por bloco de tamanho 300 ($P(m,300)$), para o Caso II.b.	81
Figura 5.19 – Estimativas da distribuição de surtos de erro para o Caso II.c.	82
Figura 5.20 – Estimativas da distribuição de gaps para o Caso II.c.	83
Figura 5.21 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso II.c.	83
Figura 5.22 – Estimativas da função Probabilidade de erros por bloco de tamanho 300 ($P(m,300)$), para o Caso II.c.	84
Figura 5.23 – Estimativas da distribuição de comprimentos de surtos de erro para o Caso III.a.	85
Figura 5.24 – Estimativas da distribuição de gaps para o Caso III.a.	86
Figura 5.25 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso III.a.	86
Figura 5.26 – Estimativas da função autocorrelação (BECF) para o Caso III.a.	87
Figura 5.27 – Estimativas da função Probabilidade de erros por bloco de tamanho 300 ($P(m,300)$), para o Caso III.a.	88
Figura 5.28 – Estimativas da distribuição de comprimentos de surtos de erro para o Caso III.b.	89
Figura 5.29 – Estimativas da distribuição de gaps para o Caso III.b.	89
Figura 5.30 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso III.b.	90

Figura 5.31 – Estimativas da função autocorrelação (BECF) para o Caso III.b.	90
Figura 5.32 – Estimativas da função Probabilidade de erros por bloco de tamanho 300, $P(m,300)$, para o Caso III.b	91
Figura 5.33 – Estimativas obtidas sem o emprego do PSO no ajuste do modelo proposto.	92
Figura 5.34 – Estimativas obtidas com o emprego do PSO no ajuste do modelo proposto.	93

Lista de tabelas

Tabela 3.1 – Funções dos estados.

45

1 Introdução

O crescente avanço tecnológico tem possibilitado meios de transportes cada vez mais velozes e com capacidade de atingir distâncias cada vez maiores, seja na terra, no mar, no ar ou no espaço, onde se incluem os satélites artificiais, sondas e naves espaciais o que tem requerido sistemas de comunicações digitais “sem fio” cada vez mais confiáveis, tanto para o tráfego de informações como para o de controle remoto de “máquina-para-máquina”, conhecido como "M2M" (mnemônico do inglês “machine-to-machine”).

O tráfego de informações, hoje em dia constituídas de imagem, som, internet banda larga e outros, torna necessário o emprego de altas taxas de transmissão de dados associados a protocolos cada vez mais robustos, com a finalidade de se minimizar os erros que corrompem as comunicações. Estas condições dão origem a canais de comunicações com características especiais, os denominados canais com memória, os quais apresentam como problema principal a ocorrência de erros em surtos, que é um fenômeno verificado principalmente nos canais “sem fio” incluindo as comunicações móveis, via satélite e os sistemas que operam na faixa de HF, todos amplamente empregados pelos sistemas de comunicações atuais.

A qualidade destes canais pode variar aleatoriamente devido a diversos fatores tais como a mobilidade, o sombreamento e o efeito “multi-percursos”, que podem influenciar no processo de desvanecimento do sinal.

Isto ressalta a importância da modelagem destes canais visando o estudo deste fenômeno, e a possibilidade de desenvolvimento de métodos que o minimizem, tornando mais confiáveis estes importantes sistemas de comunicações.

Os erros em surtos introduzidos pelo canal com memória de um sistema de comunicações podem acarretar impactos sobre os protocolos de comunicações. Como exemplo, pode-se citar o protocolo TCP o qual associado ao IP é amplamente utilizado em aplicações Internet, tendo sido inicialmente projetado para redes por cabo, onde a taxa de erro de canal é muito baixa. Entretanto,

quando associado a canais “sem fio”, o TCP tem seu desempenho severamente afetado, devido à perda de pacotes ocasionada pelos surtos de erros [3].

Da mesma forma nos anos mais recentes aconteceu com o ATM “sem fio”, o qual inicialmente projetado para ambientes livres de erros, agora operando com canais “sem fio” requer estudos sobre o impacto que recebe pelos erros em surtos [25]. Por exemplo, os sistemas de comunicações táticas do Exército dos E.U.A. são tipicamente amparados por sistemas de satélites a fim de prover alcance além da linha de visada. Estes sistemas adotam os protocolos ATM ou IP, os quais requerem canais de alta qualidade com BER menor ou igual de 10^{-8} , o que torna importante o estudo do impacto causado pelos erros em surtos no desempenho das técnicas de controle de erros.

Neste trabalho é apresentada a proposta de um modelo de surtos abrangente no sentido de representar com precisão o comportamento de diferentes tipos de canais e eficiente no sentido de apresentar número de parâmetros e complexidade matemática moderados, possibilitando complexidade computacional não elevada.

1.1. Modelagem de erros

A modelagem de erros, foco do problema abordado por este trabalho, é aqui direcionada aos erros em surtos, buscando-se eficiência no que diz respeito à precisão associada à baixa complexidade matemática e computacional e abrangência, possibilitando a capacidade de representar dados gerados por diferentes fontes de erros em surtos.

A modelagem de canais de comunicações tem como finalidade a obtenção de modelos matemáticos capazes de representar, através do adequado ajuste de seus parâmetros, o comportamento estatístico destes canais, sendo de grande importância para o levantamento de estatísticas de interesse direcionadas ao estudo e análise dos erros introduzidos por estes canais.

Modelos nos quais os erros são independentes e identicamente distribuídos (IID) são inadequados quando se espera as condições de um canal sujeito ao fenômeno dos erros em surtos [26], pois estes possuem significativo grau de correlação, não podendo, portanto, ser desprezado o efeito memória destes canais.

Os modelos matemáticos mais empregados para representar processos com memória têm sido os modelos de Markov, em que se supõe que a informação atual depende exclusivamente da informação do instante de tempo imediatamente anterior sendo desnecessária a informação de todo o passado restante [27] e [8]. Dentre eles destacam-se os amplamente utilizados modelos de Gilbert-Elliott e Fritchman, que por serem simples aproximações, apresentam limitações em representar o efeito memória dos canais com erros em surtos [26].

Como exemplo destas limitações, Chu e Sweeney [4] citam que para o canal UHF em satélites de baixa órbita, o modelo de Gilbert-Elliott é inadequado para representar as estatísticas de surto de erros inerentes a este canal e que para ângulos de elevação de 23° e 52° , as mesmas podem ser adequadamente representadas por um modelo de Fritchman de 3 estados, enquanto que para os outros ângulos empregados é necessário um Fritchman de 4 estados.

Métodos empíricos e ajuste de curvas também são empregados para representação de distribuições estatísticas, como em [5], que procura ajustar a distribuição dos comprimentos de surtos por uma distribuição geométrica. Em [10] mapas caóticos são aplicados para o ajuste de curvas de intervalos entre surtos e [2] e [31] apresentam modelos baseados em processos determinísticos, porém estes são aplicados apenas a determinados tipos de canais, deixando a desejar no aspecto abrangência, mencionado no primeiro parágrafo.

Modelos mais precisos podem ser encontrados, porém em muitos casos, resultando em um aumento de complexidade matemática, não compensador em termos de implementação computacional [26]. Assim, na escolha do modelo mais adequado ao mecanismo gerador de surtos o qual desejamos representar, devemos levar em conta os fatores precisão e complexidade, assim como o tempo de processamento computacional para os cálculos nos ajustes de estatísticas de interesse. O fator precisão significa a capacidade do modelo de representar, em termos estatísticos, o comportamento de um determinado canal de comunicações. O fator complexidade diz respeito à representação matemática do modelo, implicando diretamente na complexidade computacional, em termos de memória e do tempo de processamento, o qual se relaciona à velocidade com que os resultados da estimação de parâmetros e ajustes de estatísticas de interesse são gerados.

O modelo proposto tem a vantagem de possibilitar a dedução de expressões analíticas simples para a função de verossimilhança para estimação ML dos seus parâmetros e para algumas estatísticas de interesse, abordadas no Capítulo 3. Assim, a complexidade computacional é relativamente baixa, sendo possível de se realizar ajustes em tempo muito inferior ao de outros modelos com mesmo número de parâmetros.

1.2. Surto de Erro

Para o estudo dos surtos de erro produzidos por canais com memória, algumas definições tornam-se necessárias. Em [2], uma *sequência de erros* é representada por uma seqüência binária de uns e zeros onde o um representa bit com erro e o zero representa bit sem erro. Um *gap* é definido como uma seqüência de bits zeros consecutivos entre dois bits uns, cujo comprimento é igual ao número de bits zeros. Um *cluster* é uma seqüência de bits uns consecutivos, cujo comprimento é igual ao número de bits uns. Para estas duas definições considerou-se os termos em inglês *gap* e *cluster*, devido ao amplo uso destas palavras.

Da definição do CCITT em [1] : “Um surto de erro é um grupo de bits no qual dois bits errados sucessivos estão separados por menos do que um dado número X de bits corretos”. Ou seja, segundo esta definição, no interior de um surto de erro, podem existir seqüências contínuas de bits sem erro (*gaps*), desde que com comprimentos menores do que X .

Finalmente, um *surto livre de erros* é definido como uma seqüência de bits zeros consecutivos com comprimento de pelo menos X bits, onde X é um número inteiro positivo. Um *surto livre de erro* não precisa estar necessariamente entre dois bits uns como o caso do *gap*.

Observe que por estas definições, uma mesma *sequência de erros* pode ser representada por diferentes distribuições de comprimentos de *surtos de erro*, já que os comprimentos dos surtos de erro dependem diretamente do valor de X , o mesmo acontecendo para distribuição de *surtos livres de erro*.

Observamos também que à medida que diminuimos o valor de X , *surtos de erro* maiores vão se dividindo em surtos menores e para $X=1$, os *surtos de erro* se

reduzem a *clusters*. A literatura pertinente e o próprio CCITT não apresentam maiores esclarecimentos sobre a origem do referido parâmetro X . Supomos que tenha a finalidade de adequar a definição de erro em surto às limitações relativas aos protocolos de tratamento de erros encontradas nas diversas camadas dos sistemas de comunicações.

1.3. Proposta deste trabalho

O novo modelo de surtos proposto neste trabalho tem a finalidade de modelar os erros em surtos produzidos por canais com memória, para representação de determinadas estatísticas de interesse à análise do impacto destes erros nos sistemas de comunicações, de modo a contornar algumas deficiências encontradas em modelos já existentes, tais como a falta de flexibilidade em se ajustar dados gerados por diferentes fontes e a grande complexidade no tratamento analítico e computacional. Esta complexidade decorre principalmente da tentativa de se buscar maior precisão nos ajustes por meio de modelos mais complexos.

Trata-se de um Modelo Escondido de Markov, também conhecido pela abreviatura desta expressão em inglês HMM (Hidden Markov Model), com características especiais. Sua estrutura foi baseada na própria definição de surto de erro do CCITT [1], descrito no item 1.2, onde os seus estados são dispostos de modo a formarem grupos, cada qual destinado a reproduzir um dos eventos possíveis de ocorrerem em uma sequência binária afetada por surtos de erro. Estes eventos são o *surto de erros (error burst)*; a sequência de bits errados (*cluster*); o *intervalo entre surtos (error free burst)* e erros aleatórios isolados.

O modelo proposto facilita a dedução de expressões analíticas de estatísticas de interesse e da função de verossimilhança. Estas expressões são funções dos sete parâmetros do modelo, que são seis números reais e um inteiro positivo. Assim se possibilita a estimação dos mesmos pelo método de Máxima Verossimilhança utilizando-se ferramentas de otimização.

Com o objetivo de capturar eficientemente o comportamento estatístico de massas de dados geradas por diferentes fontes, a presente proposta é de um modelo abrangente em relação à origem dos dados a serem ajustados e que exija um tempo computacional para os cálculos relativos à estimação ML dos

parâmetros, bem menor do que o necessário a outros métodos que envolvam a mesma quantidade de parâmetros.

1.4. Apresentação do trabalho

Os tópicos abordados por este trabalho estão distribuídos da seguinte maneira: no Capítulo 2 serão apresentados conceitos básicos e teorias utilizadas neste trabalho, relacionadas ao modelo proposto e aos modelos e técnicas utilizados como referências de comparação com o mesmo; no Capítulo 3, será descrito de forma detalhada, o modelo de surtos proposto e seus objetivos, apresentando como o mesmo foi concebido e como se deu sua evolução; no Capítulo 4 serão abordadas as técnicas utilizadas para estimação dos parâmetros do modelo proposto; no Capítulo 5 serão apresentados os resultados obtidos da aplicação do modelo proposto a dados produzidos por diferentes fontes geradoras de erros em surtos; e finalmente o Capítulo 6 apresenta algumas conclusões e sugestões para trabalhos futuros.

2 Conceitos básicos

O objetivo deste Capítulo é abordar teoricamente os assuntos que formam a base para o desenvolvimento do modelo proposto e a descrição do modelo de Fritchman, que devido sua frequente aplicação em trabalhos encontrados na literatura da área, foi utilizado como modelo de referência para o ajuste dos dados segundo as estatísticas de interesse. Assim, este Capítulo está dividido em duas partes. A primeira trata dos modelos de Markov, incluindo os Modelos Escondidos de Markov (HMM) e o método de Baum-Welch para estimação de seus parâmetros. A segunda parte aborda as ferramentas de otimização empregadas para estimação dos parâmetros do modelo proposto, incluindo uma técnica de otimização clássica e o processo heurístico conhecido como PSO (Particle Swarm Optimization).

2.1. Modelos de Markov

Estes modelos tem como elemento básico a Cadeia de Markov, que é um processo estocástico de tempo discreto em que a transição para um estado no instante t , depende somente do estado visitado no instante $t-1$.

Define-se como probabilidade de transição aquela que caracteriza a transição para o estado J_t (estado no instante t) dado que o estado anterior foi I_{t-1} .

Esta é a definição do que é denominado de cadeia de Markov de ordem um, pois é possível conceber a situação onde a transição para um estado no instante t depende dos k últimos estados visitados. Neste caso a cadeia de Markov é denominada de ordem k .

Neste trabalho, concentraremos nossa atenção nas cadeias de Markov homogêneas e estacionárias de primeira ordem caracterizada por um número finito de estados, onde o estado no instante t é denotado por s_t .

Assim uma cadeia de Markov pode ser caracterizada pela trinca $\{S, \pi, P\}$ onde:

- $S = \{1, 2, \dots, u\}$ – Conjunto dos estados do modelo
 $\pi = [p_1 \ p_2 \ \dots \ p_u]$, $p_j = P(s_0=j)$ – Vetor das probabilidades iniciais de estado
 $P = [p_{ij}]_{u \times u}$, $p_{ij} = P(s_t=j \mid s_{t-1}=i)$ – Matriz de probabilidades de transição

A Figura abaixo ilustra uma cadeia de Markov de três estados, sendo que entre eles, todas as transições são possíveis.

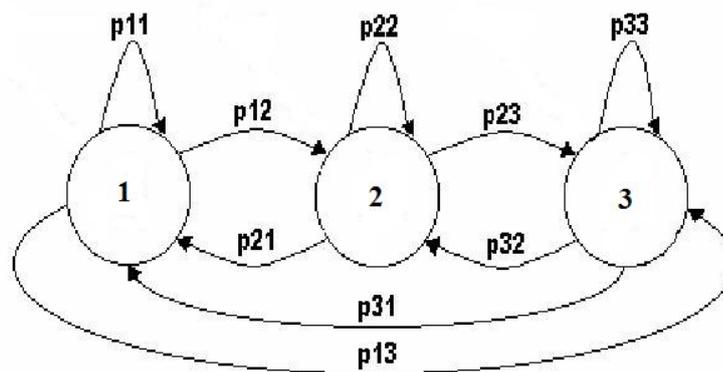


Figura 2.1 – Cadeia de Markov de 3 estados

2.2. Modelos Escondidos de Markov

Denomina-se de Modelo Escondido de Markov (HMM) a um modelo baseado em uma cadeia de Markov onde:

- Uma observação é produzida ao se acessar cada estado. Esta observação é de natureza estocástica e particular de cada estado.
- Os estados da cadeia de Markov subentendida não são observáveis (daí o nome “escondido”).

Assim um Modelo Escondido de Markov é caracterizado por uma cadeia de Markov subjacente e por distribuições de probabilidade das observações produzidas em cada estado. Mais formalmente, um HMM pode ser descrito pelo seguinte conjunto de parâmetros:

$$\{ S, A, \pi, P, B \}$$

onde:

- $S = \{1, 2, \dots, u\}$ – Conjunto dos estados do modelo;
 $A = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$ – Conjunto dos símbolos de saída;

$\pi = [p_1 p_2 \dots p_u]$, $p_j = P(s_1=j)$ – Vetor das probabilidades iniciais de estado;
 $P = [p_{ij}]_{u \times u}$, $p_{ij} = P(s_t=j | s_{t-1}=i)$ – Matriz de probabilidades de transição;
 $B = [P(O | j)]$ – Vetor com a distribuição de probabilidades das observações (O) por estado (j).

Em resumo, o HMM é um modelo composto de uma cadeia de Markov subjacente cujos estados não são observáveis, onde uma observação é emitida quando da permanência em qualquer um de seus estados.

Como veremos mais adiante, os modelos HMM tem se mostrado extremamente úteis em Telecomunicações, particularmente no que diz respeito à modelagem de canais de comunicações em processos de transmissão digital [16], [17]. Esta discussão é o objetivo da próxima seção.

2.3. HMM aplicado a canais sujeitos a erros em surtos

Em comunicações digitais, podemos supor, sem perda de generalidade, que as imperfeições existentes no canal de comunicações farão com que a sequência recebida seja diferente da sequência transmitida, o que caracteriza o que é chamado de *erro de transmissão*. Assim todo o efeito danoso do canal de comunicações pode ser igualmente representado por uma sequência binária onde por convenção, usaremos 1 e 0 para respectivamente representar a ocorrência e não ocorrência de erro num particular instante do processo de transmissão de dados.

Assim caracterização do comportamento de um canal de comunicações pode ser descrita por uma longa sequência binária, majoritariamente constituída por 0's, onde os 1's representam a ocorrência de erros no processo de transmissão.

Tem sido o objetivo de vários pesquisadores nesta área, a busca por modelos teóricos de natureza estocástica que representem probabilisticamente e descrevam adequadamente este fenômeno de ocorrência de erros provocados pelo canal de comunicações. A utilidade de tais modelos é indiscutível e dentre as suas principais vantagens podemos citar a capacidade de simular novos sistemas de comunicações em canais de comunicações cujos efeitos sobre os dados transmitidos podem ser avaliados sem a necessidade de práticas experimentais.

Com relação aos efeitos danosos causados pelos canais de comunicações com memória constata-se que os erros não podem ser considerados estatisticamente independentes. Observa-se em muitos casos que erros acontecem na forma que é comumente denominada de *surtos*, onde longos períodos de ausência de erros são seguidos por curtos períodos de alta incidência de erros.

Assim, a busca de modelos teóricos que sejam capazes de explicar estatisticamente este fenômeno de surtos, pode ser considerada extremamente importante. Observa-se que a literatura tem revelado que os modelos escondidos de Markov tem sido usados na explicação estatística de alguns fenômenos de ocorrência de erros em surtos [3, 4, 7, 16, 17,19, 20, 25, 26]. Porém estes não são adequados a muitos dos casos de geração de surtos, e ainda estão sujeitos a limitações impostas pelo impraticável aumento de complexidade matemática quando se busca maior precisão de ajustes por meio do aumento do número de estados.

Esta tese é basicamente motivada por essa constatação, e visa propor e discutir um modelo que apresente adequada capacidade para representar tal fenômeno através de seu comportamento estatístico, superando as dificuldades acima mencionadas, de um novo HMM com estrutura especial, direcionada a geração de sequências de erros.

A título de ilustração, serão apresentados a seguir, alguns modelos clássicos de HMM encontrados na literatura para a caracterização estatística de canais de comunicações.

2.4. Modelo de Gilbert-Elliott

O modelo denominado de Gilbert-Elliott, em referência a seus autores, é um HMM de dois estados. O diagrama de estados correspondente encontra-se na Figura 2.2.

Este modelo é constituído de um estado onde a ocorrência de erros possui baixa probabilidade (daí o estado ser chamado de “bom” ou “good” (G)) e por um segundo estado , onde a ocorrência de erros possui alta probabilidade (estado chamado de “ruim” ou “bad” (B)). A frequência com que se permanece em cada um destes estados é controlada pelas probabilidades de transição entre estados.

Assim, este modelo é caracterizado por cinco parâmetros e descrito por:

$$\{ S, A, \pi, P, B_e \}$$

onde:

$$S = \{ 1, 2 \}$$

$$A = \{ 0, 1 \}$$

$$\pi = [p_1 \quad 1-p_1]$$

$$P = \begin{bmatrix} p_{11} & 1-p_{11} \\ 1-p_{22} & p_{22} \end{bmatrix}$$

$$B_e = [p_{e1} \quad p_{e2}] \quad (\text{Vetor das probabilidades de erro por estado})$$

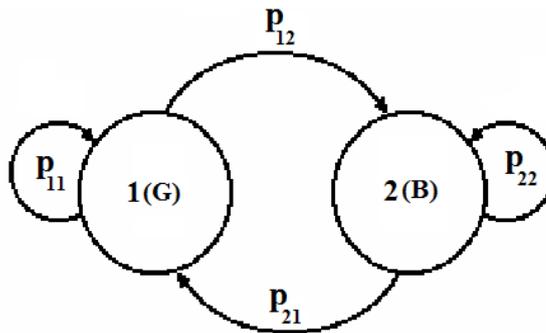


Figura 2.2 Modelo de Gilbert-Elliott

2.5.

Modelo de Fritchman

Este modelo representa uma generalização do anterior e é um modelo consagrado e amplamente empregado na modelagem de erros em surtos. Por ser particularmente importante, será tomado como referência de comparação com o modelo proposto nesta tese.

Num modelo de Fritchman de N estados, estes são particionados em dois grupos, um com k estados de erro (ruins), que geram “1”s com probabilidade 1, e o outro com $N-k$ estados livres de erros (bons), que geram “0”s com probabilidade 1.

Somente são permitidas auto-transições ou transições entre estes grupos[7]. Assim, a distribuição do comprimento dos *gaps* (intervalo entre surtos) é descrita por uma combinação linear de $N-k$ distribuições exponenciais, enquanto a distribuição de comprimentos de surtos é descrita por uma combinação linear de k distribuições exponenciais [4]. Apesar de este ser modelo mais rico do que o Gilbert-Elliott, tem-se verificado que é inadequado para muitos casos de interesse [26, 32]. Este modelo possui um número maior de estados e conseqüentemente a estimação de seus parâmetros apresenta complexidade em geral elevada.

Uma particularização do modelo de Fritchman é o denominado de modelo Fritchman-SES (SES – *Single-Error-State*), selecionado em [9] para modelar o “HF SchEMe” (*Skywave Channel Error Model*), e corresponde ao caso em que $k=1$.

A Figura 2.3 ilustra um modelo de Fritchman-SES de N estados, onde o estado ruim é indicado por (B) e os $N-1$ estados bons são respectivamente indicados por (G_1) , (G_2) ... (G_{N-1}) . Este modelo é descrito pela matriz de probabilidades de transição mostrada em (2.1), onde o estado ruim é o de índice N :

$$\mathbf{P} = \begin{bmatrix} p_{11} & 0 & 0 & \cdots & 0 & p_{1N} \\ 0 & p_{22} & 0 & \cdots & 0 & p_{2N} \\ 0 & 0 & p_{33} & \cdots & 0 & p_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p_{N-1,N-1} & 0 \\ p_{N1} & p_{N2} & p_{N3} & \cdots & p_{N,N-1} & p_{NN} \end{bmatrix} \quad (2.1)$$

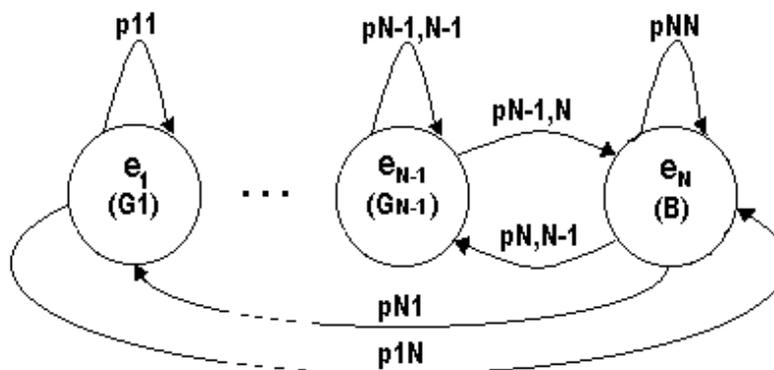


Figura 2.3 - Modelo de Fritchman-SES de N estados

Outros modelos de Fritchman específicos são também utilizadas em diversas aplicações. Estudos sobre modelagem de canais de satélites de baixa órbita (LEO) descritos em [4] empregam, além dos modelos de Fritchman-SES de três e quatro estados, o modelo de Fritchman ($k=2, N=4$). O primeiro modelo mencionado acima é utilizado em [4] para descrever a estatística de surtos de erros do sistema LEO-UHF para os ângulos de elevação de 23° e 52° ; o segundo para os ângulos de elevação restantes, e o terceiro para canais sujeitos a interferências.

2.6. Estimação de parâmetros do HMM

A estimação de parâmetros dos HMM's a partir de uma sequência de dados observados dá origem a três problemas clássicos e de grande interesse:

- i) o cálculo do estimador de máxima verossimilhança (ML) das observações condicionado aos parâmetros do HMM;
- ii) a estimação da sequência de estados mais provável e
- iii) a estimação ML de parâmetros. A seguir, será apresentada uma breve descrição de cada um desses três problemas [14].

2.6.1. Cálculo da probabilidade de geração da sequência observada

Com o objetivo de simplificar as notações, considere um HMM $\lambda=(P,B,\pi)$, onde, P representa sua matriz de probabilidades de transição; B a matriz das distribuições de probabilidades das emissões em cada estado e π o vetor das probabilidades iniciais de cada estado.

Seja a sequência $O=\{ o_1, o_2, \dots o_T \}$ uma sequência de observações. Deseja-se calcular a probabilidade das observações terem sido geradas pelo modelo λ :

$$V(O|\lambda); \quad (2.2)$$

Se este cálculo for feito de maneira direta, a sua complexidade o torna computacionalmente inviável, mesmo para seqüências de pequeno comprimento, pois envolve um número de operações da ordem N^T , que cresce exponencialmente com o tamanho da seqüência (T), sendo N o número de estados. Assim devemos

buscar um método de cálculo que reduza esta complexidade, de modo a tornar o problema computacionalmente tratável. Para tanto, faremos uso de duas variáveis auxiliares: $\alpha_t(i)$ – Variável Progressiva e $\beta_t(i)$ – Variável Regressiva.

Variável Progressiva

A Variável Progressiva $\alpha_t(i)$ é definida como a probabilidade de ocorrência da seqüência parcial de observações $O_1^t = \{o_1, o_2, \dots, o_t\}$, sendo o_t gerada pelo estado i . ou seja,

$$\alpha_t(i) = P(o_1, o_2, \dots, o_t, s_t = i | \lambda) \quad (2.3)$$

Analisando todas as possibilidades de transição para um estado j , podemos trivialmente, deduzir a seguinte relação recursiva:

$$\alpha_{t+1}(j) = b_j(o_{t+1}) \sum_{i=1}^N \alpha_t(i) p_{ij}, \quad 1 \leq j \leq N, \quad 1 \leq t \leq T-1, \quad (2.4)$$

onde:

p_{ij} é o elemento da matriz de probabilidades de transição P correspondente à transição do estado i para o estado j . $b_j(o_{t+1})$ é a probabilidade de emitir a observação o_{t+1} , no instante $t+1$, dado que se esteja no estado j ;

Os cálculos podem ser iniciados por:

$$\alpha_1(j) = \pi_j b_j(o_1), \quad 1 \leq j \leq N. \quad (2.5)$$

onde:

π_j é a probabilidade inicial de se estar no estado j .

Através das relações acima se calcula esta variável para toda a seqüência:

$$\alpha_T(i), \quad 1 \leq j \leq N \quad (2.6)$$

Por fim, a probabilidade de geração da sequência observada é dada por

$$V(O | \lambda) = \sum_{i=1}^N \alpha_T(i) \quad (2.7)$$

Note-se que a complexidade deste cálculo é proporcional a N^2T , ou seja, tem crescimento linear com o tamanho da sequência T .

Variável Regressiva

A Variável Regressiva $\beta_t(i)$ é definida como a probabilidade condicional de ocorrência da sequência parcial de observações $O_{t+1}^T = \{o_{t+1}, o_{t+2}, \dots, o_T\}$, dado o estado s_T é igual a i , ou seja,

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, \dots, o_T | s_T = i, \lambda) \quad (2.8)$$

Como no caso de $\alpha_t(i)$, podemos facilmente chegar à seguinte relação recursiva para o cálculo de $\beta_t(i)$:

$$\beta_t(i) = \sum_{j=1}^N \beta_{t+1}(j) p_{ij} b_j(o_{t+1}), \quad 1 \leq i \leq N, \quad 1 \leq t \leq T-1, \quad (2.9)$$

onde:

$$\beta_T(i) = 1, \quad 1 \leq i \leq N \quad (2.10)$$

Das relações acima temos:

$$\alpha_t(i) \beta_t(i) = P(O, s_t = i | \lambda), \quad 1 \leq i \leq N, \quad 1 \leq t \leq T \quad (2.11)$$

Assim sendo, obtemos outro modo para o cálculo de $V(O|\lambda)$, como a seguir:

$$V(O|\lambda) = \sum_{i=1}^N P(O, s_t = i | \lambda) = \sum_{i=1}^N \alpha_t(i) \beta_t(i) \quad (2.12)$$

2.6.2.

Determinação da sequência de estados mais provável

Dado um HMM $\lambda=(P,B,\pi)$ e uma sequência $O=\{o_1, o_2, \dots, o_T\}$ de observações, deseja-se saber qual a sequência de estados $S = \{s_1, s_2, \dots, s_T\}$ com maior probabilidade de ter gerado as observações O . Este problema equivale à maximização da probabilidade $P(S,O|\lambda)$ e pode ser resolvido pelo *Algoritmo de Viterbi*, como se mostra a seguir:

Primeiramente definimos a variável $\delta_t(i)$, que indica a sequência de estados mais provável de comprimento t , relativa às t primeiras observações, e que termina no estado i :

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P(s_1 s_2 \dots s_{t-1}, s_t = i, o_1 o_2 \dots o_t | \lambda) \quad (2.13)$$

Por indução pode-se mostrar que:

$$\delta_{t+1}(j) = \max_i [\delta_t(i) p_{ij}] b_j(o_{t+1}) \quad (2.14)$$

Assim encontramos a sequência de estados $s_1 s_2, \dots, s_t$, que satisfaz a equação (2.14) para cada t e j .

Para obter a sequência de estados mais provável podemos seguir os seguintes passos:

a) Inicialização:

$$\delta_1(i) = \pi_i b_i(o_1) \quad (2.15)$$

$$\psi_1(i) = 0 \quad 1 \leq i \leq N \quad (2.16)$$

b) Recursividade:

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) p_{ij}] b_j(o_t) \quad (2.17)$$

$$\psi_t(j) = \arg \max_{1 \leq i \leq N} [\delta_{t-1}(i) p_{ij}] \quad 2 \leq t \leq T \quad e \quad 1 \leq j \leq N \quad (2.18)$$

c) Término:

$$P^* = \max_{1 \leq i \leq N} [\delta_T(i)] \quad (2.19)$$

$$q_T^* = \arg \max_{1 \leq i \leq N} [\delta_T(i)] \quad (2.20)$$

d) Seqüência de estados:

$$q_t^* = \psi_{t+1}(s_{t+1}^*) \quad t=T-1, T-2, \dots, 1 \quad (2.21)$$

2.6.3. Estimação ML de Parâmetros

Dada uma seqüência $O = \{ o_1, o_2, \dots, o_T \}$ de observações, deseja-se ajustar os parâmetros do modelo HMM $\lambda = (A, B, \pi)$, de modo a maximizar:

$$V_{tot} = P(O/\lambda) \quad (2.22)$$

Não existe nenhuma maneira conhecida de se resolver este problema analiticamente, ou seja, encontrar os parâmetros do HMM λ que maximize V_{tot} [14]. No entanto podemos encontrar parâmetros que maximizem localmente V_{tot} , através do uso do *Algoritmo de Baum-Welch*, o qual está descrito a seguir.

2.6.4. Algoritmo de Baum-Welch

Primeiramente definimos a função auxiliar Q :

$$Q(\lambda', \lambda) = \sum_q P(O, S | \lambda') \log P(O, S | \lambda) \quad (2.23)$$

a qual deverá ser maximizada em λ a fim de atualizar λ' no sentido de aumentar $V(O|\lambda)$, pois mostra-se em [14] que:

$$Q(\lambda', \lambda) \geq Q(\lambda', \lambda') \Rightarrow P(O | \lambda) \geq P(O | \lambda') \quad (2.24)$$

Uma característica muito importante deste algoritmo é a garantia da convergência para um ponto de máximo local [14].

Definem-se ainda, mais duas variáveis auxiliares em adição às *variáveis, progressiva e regressiva*, definidas anteriormente. A primeira delas é dada pela probabilidade de se estar no estado i em t e no estado j em $t+1$ dada a seqüência de observações, ou seja:

$$\xi_t(i, j) = P(s_t = i, s_{t+1} = j | O, \lambda) \quad (2.25)$$

ou ainda:

$$\xi_t(i, j) = \frac{P(s_t = i, s_{t+1} = j, O | \lambda)}{P(O | \lambda)} \quad (2.26)$$

Em função das variáveis progressiva e regressiva, temos:

$$\xi_t(i, j) = \frac{\alpha_t(i) p_{ij} \beta_{t+1}(j) b_j(o_{t+1})}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) p_{ij} \beta_{t+1}(j) b_j(o_{t+1})} \quad (2.27)$$

A segunda variável auxiliar é dada pela probabilidade de se estar no estado i em t , dada a seqüência de observações, ou seja:

$$\gamma_t(i) = P(s_t = i | O, \lambda) \quad (2.28)$$

Em função das variáveis progressiva e regressiva, podemos obter:

$$\gamma_t = \left[\frac{\alpha_t(i)\beta_t(i)}{\sum_{i=1}^N \alpha_t(i)\beta_t(i)} \right]. \quad (2.29)$$

A relação entre $\gamma(i)$ e $\xi(i,j)$ é dada por:

$$\gamma_t = \sum_{j=1}^N \xi_t(i,j), \quad 1 \leq i \leq N, \quad 1 \leq t \leq T \quad (2.30)$$

Os parâmetros de um HMM são atualizados a cada iteração, no sentido de maximizar a probabilidade $P(O|\lambda)$, supondo um modelo inicial $\lambda_1 = (P_1, B_1, \pi_1)$.

Inicialmente calculamos os ' α 's utilizando as equações (2.4) e (2.5), e os ' β 's utilizando as equações (2.9) e (2.10). Em seguida calculamos ξ e γ através das equações (2.27) e (2.29), respectivamente. Finalmente atualizamos os parâmetros do HMM seguindo as equações a seguir, conhecidas como *fórmulas de reestimação*:

$$\bar{\pi}_i = \gamma_1(i), \quad 1 \leq i \leq N \quad (2.31)$$

$$\bar{p}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad 1 \leq i \leq N, \quad 1 \leq j \leq N \quad (2.32)$$

$$\bar{b}_j(v_k) = \frac{\sum_{t=1}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}, \quad 1 \leq j \leq N, \quad 1 \leq k \leq m \quad (2.33)$$

onde:

v_k pertence ao alfabeto de símbolos $\{v_1, v_2, \dots, v_m\}$

Em função das variáveis progressiva e regressiva, temos:

$$\bar{\pi}_i = \frac{\alpha_i(i)\beta_1(i)}{\sum_{j=1}^N \alpha_T(j)} \quad (2.34)$$

$$\bar{p}_{ij} = \frac{\sum_{t=1}^T \alpha_{t-1}(i) p_{ij} b_j(o_t) \beta_t(j)}{\sum_{t=1}^T \alpha_{t-1}(i) \beta_{t-1}(i)} \quad (2.35)$$

$$\bar{b}_i(v_k) = \frac{\sum_{t=1}^T \alpha_t(i) \beta_t(i) \delta(o_t, v_k)}{\sum_{t=1}^T \alpha_t(i) \beta_t(i)} \quad (2.36)$$

sendo:

$$\delta(o_t, v_k) = \begin{cases} 1 & \text{se } o_t = v_k \\ 0 & \text{se } o_t \neq v_k \end{cases}$$

Problemas relativos à implementação do algoritmo

De posse das fórmulas de reestimação do algoritmo de Baum-Welch, verifica-se que ainda existem vários detalhes em nível de implementação, para que este possa ser executado em computador. Serão comentados em seguida os mais significativos para este trabalho, acompanhado da solução adotada.

Limite de ordem de grandeza inferior

As operações recursivas realizadas durante a estimação dos parâmetros de um HMM fazem com que estes tendam a zero com o número de iterações. Isto acontece devido ao fato de que as variáveis $\alpha(i)$ e $\beta(i)$ são produzidas por um cálculo recursivo e que cada iteração envolve produtos de probabilidades, como observado nas equações (2.4) e (2.9).

Assim, mesmo que a seqüência não seja muito grande, o limite de precisão de qualquer computador é rapidamente atingido ao se executar este algoritmo.

Para combater este problema utilizaremos uma técnica de ponderar estas variáveis, ou seja, multiplicá-las por um fator de escala ' c_t ' dependente do tempo, mas que não dependa de ' i ' [14], de forma que os valores destas variáveis sejam mantidos dentro dos limites de precisão do computador, quando o valor de t varia de 1 até T . Os cálculos deste procedimento estão descritos a seguir:

- para $t=1$, calcula-se $\alpha_1(i)$ como na equação (2.4);

- cálculo do fator c_1 :

$$c_1 = \frac{1}{\sum_{i=1}^N \alpha_1(i)} \quad \text{e} \quad \hat{\alpha}_1(i) = c_1 \alpha_1(i) \quad (2.37)$$

onde, $\hat{\alpha}_1(i)$ representa o $\alpha_1(i)$ ponderado pelo fator c_1 ;

- cálculo recursivo de $\hat{\alpha}_t(i)$:

$$\hat{\alpha}_t(i) = \sum_{j=1}^N \hat{\alpha}_{t-1}(j) p_{ji} b_i(o_t); \quad (2.38)$$

- após calculado, $\hat{\alpha}_t(i)$ deve ser aplicado nas equações (2.39), para o cálculo dos novos valores de c_t e $\hat{\alpha}_t(i)$, que deverão ser novamente aplicados na equação (2.38), e assim sucessivamente, sendo:

$$\hat{\alpha}_t(i) = c_t \hat{\alpha}_t(i) \quad \text{e} \quad c_t = \frac{1}{\sum_{i=1}^N \hat{\alpha}_t(i)} \quad (2.39)$$

Como $\beta_t(i)$ possui a mesma ordem de grandeza de $\alpha_t(i)$, utilizaremos o mesmo fator de escala c_t para ponderar estas variáveis, que também serão mantidas em limites razoáveis para o cálculo computacional. Assim, temos:

$$\hat{\beta}_t(i) = c_t \beta_t(i). \quad (2.40)$$

Seqüência de observações de tamanho insuficiente

Outro problema relativo à implementação deste algoritmo é o efeito causado por uma pequena seqüência de observação.

No caso de $b_i(v_k)$, que representa a probabilidade de ocorrer a observação v_k no estado i , observa-se que o numerador da equação de reestimação (2.36) é uma soma cujos termos em que a observação o_t é diferente do símbolo v_k , são nulos.

Se esta probabilidade é pequena e a amostra da seqüência utilizada na reestimação for também pequena, pode não ocorrer nenhuma vez o evento o_t igual a v_k no estado i . Com isso, o valor de $b_i(k)$ reestimado e os seguintes serão nulos, levando os cálculos a valores irrealis.

Possíveis soluções seriam aumentar o tamanho da seqüência observada, diminuir o número de estados, ou diminuir o número de símbolos do modelo utilizado, o que geralmente não pode ser feito. Uma solução prática seria estipular limiares mínimos para os parâmetros do modelo, como no exemplo a seguir [14]:

$$b_i(k) = \begin{cases} b_i(k), & b_i(k) \geq \delta_b; \\ \delta_b, & b_i(k) < \delta_b; \end{cases} \quad (2.41)$$

onde δ_b é o valor mínimo fixado.

Estimação inicial dos parâmetros

As equações de reestimação produzem valores dos parâmetros de um HMM, correspondentes à convergência para o máximo global ou para qualquer máximo local da função verossimilhança, se houver. Não existe nenhuma maneira simples e direta para resolver este problema.

2.7. Ferramentas de Otimização

Neste trabalho foram utilizadas duas ferramentas de otimização para a estimação dos parâmetros, a otimização clássica e *Particle Swarm Optimization*

(PSO). Ambas puderam ser empregadas, devido ao modelo proposto ter possibilitado a dedução de uma função de verossimilhança, matematicamente simples.

A otimização clássica, foi inicialmente empregada para a estimação ML dos parâmetros do modelo proposto. Observou-se a existência de diversos máximos locais na função de verossimilhança, levando à necessidade de uma apropriada inicialização dos parâmetros que aumentasse as chances de convergência para o máximo global da função de verossimilhança.

Para tanto se adotou o PSO, que apesar de ser um método que consome um tempo de processamento relativamente alto para ser empregado como método principal de otimização, se mostrou bastante eficiente na estimação de valores iniciais para a estimação ML dos parâmetros do modelo proposto. Um resumo da teoria de otimização clássica encontra-se no Apêndice B e os detalhes sobre o emprego deste método neste trabalho serão abordados no Capítulo 4.

2.7.1. Particle Swarm Optimization (PSO)

O PSO é um método de otimização global, que difere de outros métodos conhecidos como Algoritmos Evolucionários (AE) [30]. Como nos AE, um conjunto de possíveis soluções é utilizado para sondar o espaço de pesquisa adequado ao problema, porém nenhum tipo de operação baseada em rotinas de evolução é aplicado ao conjunto de soluções, para atualizá-lo.

No caso do PSO, fazendo-se uma analogia no espaço tridimensional, é como se um “enxame” de partículas se locomovessem neste espaço, em busca de um posicionamento adequado a cada uma, que gere a solução desejada. Cada elemento, aqui chamado *partícula*, do conjunto de soluções chamado *enxame*, atualiza sua trajetória considerando a sua melhor posição anterior e a melhor posição anterior alcançada por qualquer membro de sua *vizinhança* topológica [30].

Deste modo, sendo todo o *enxame* considerado como a *vizinhança*, ocorre o compartilhamento global de informações e as partículas se valem das descobertas e experiência de todas as outras, durante a busca. Muitas variações da técnica de PSO tem sido propostas [30].

Para formalização de um algoritmo PSO b considerando o espaço de pesquisa como D -dimensional, a i -ésima partícula do enxame é representada pelo vetor D -dimensional $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, e a partícula correspondente ao maior valor da função que se deseja maximizar de todo o enxame indicada pelo índice g . A melhor posição conseguida e velocidade, correspondentes a partícula X_i são, respectivamente representadas e armazenadas pelos vetores:

$$P_i = (p_{i1}, p_{i2}, \dots, p_{iD}) \text{ e } V_i = (v_{i1}, v_{i2}, \dots, v_{iD}) \quad (2.42)$$

As partículas se relacionam segundo as seguintes equações:

$$V_i^{n+1} = wV_i^n + c_1 r_{i1}^n (P_i^n - X_i^n) + c_2 r_{i2}^n (P_g^n - X_i^n) \quad (2.43)$$

$$X_i^{n+1} = X_i^n + \chi V_i^{n+1}, \quad i = 1, 2, \dots, N \quad (2.44)$$

onde:

- os super-escritos indicam a iteração;
- χ é um fator de constrição para controlar a velocidade;
- w é o *peso inercial* e sua finalidade é descrita a seguir- c_1 e c_2 são constantes positivas chamadas de parâmetro cognitivo e social respectivamente, cujo emprego é descrito abaixo
- r_{i1}^n e r_{i2}^n são dois números uniformemente distribuídos no intervalo $[0, 1]$, cuja finalidade é comentada abaixo.

A equação (2.43) é usada para calcular a nova velocidade da i -ésima partícula em cada iteração, sendo o primeiro termo desta equação, wV_i^n , representa a velocidade anterior da partícula multiplicadas pelo *peso inercial* w . O segundo termo $(P_i^n - X_i^n)$ representa a distância entre a melhor posição anterior e atual de cada partícula. O terceiro termo $(P_g^n - X_i^n)$, é a distância entre a melhor posição entre todas as partículas do enxame e a posição atual da i -ésima partícula.

Os parâmetros $c_1 r_{i1}^n$ e $c_2 r_{i2}^n$ introduzem aleatoriedade, a qual deixa a técnica menos previsível, porém mais flexível [30].

A equação (2.44) calcula a nova posição da i -ésima partícula, adicionando a sua nova velocidade com a posição atual.

O desempenho de cada partícula é medida de acordo com a função objetivo em questão.

O *peso inercial* w é considerado fundamental para a convergência do PSO, pois esse é empregado para controlar o impacto do histórico das velocidades sobre a velocidade atual. Assim, o parâmetro w regula a troca entre as habilidades de exploração global e local do *enxame*, sendo que grandes valores do *peso inercial* facilitam a exploração global, enquanto pequenos valores facilitam a exploração local, o que pode ser visto como um ajuste fino da pesquisa na região atual. Assim sendo, a escolha deste parâmetro resulta diretamente na taxa de convergência para o ponto de máximo desejado.

A população inicial de partículas e suas respectivas velocidades podem ser geradas randomicamente ou por um gerador de sequências de Sobol [30], que assegura que os vetores D -dimensionais serão uniformemente distribuídos no espaço de pesquisa. O conceito de recombinação está relacionado ao movimento estocástico de cada partícula em direção à sua melhor posição anterior, assim como em direção a melhor posição global de todo enxame ou à melhor posição de sua vizinhança, dependendo de qual das variações de PSO é utilizada. Além de tudo.

A técnica de PSO tem se mostrado eficaz para resolver problemas de otimização global, em ambientes continuamente variantes ou em treinamento de redes neurais e ainda em problemas de otimização multi-objetivo [30].

2.7.2. Técnica “Branch and Bound” (BB)

A técnica *Branch and Bound* é amplamente utilizada para a resolução de problemas de PSO. Nesta técnica, a região pertinente ao problema é particionada em várias sub-regiões, isso é chamado de *branch*. Ao longo destas sub-regiões, limites (*bounds*) inferiores e superiores para os valores da função podem ser

determinados. A técnica BB pode ser esquematizada em forma de algoritmo, como se segue [30]:

1. Inicialmente, define-se uma região $M_0 \supset S$ e seu particionamento em um número finito de subconjuntos M_i , $i = 1, 2, \dots, m$, onde S é a região pertinente ao problema.

2. Para cada subconjunto M_i , determinam-se os limites, inferior e superior da função, $\beta(M_i)$ e $\alpha(M_i)$, respectivamente, satisfazendo a seguintes inequações:

$$\beta(M_i) \leq \inf (f(M_i \cap S)) \leq \alpha(M_i) \quad (2.45)$$

onde f é a função objetivo em consideração

Define-se os limites globais, inferior e superior, pelas equações (2.46) e (2.47), respectivamente.

$$\beta = \min_{i=1,2,\dots,m} \beta(M_i) \quad (2.46)$$

$$\alpha = \min_{i=1,2,\dots,m} \alpha(M_i) \quad (2.47)$$

sendo:

$$\beta \leq \min f(S) \leq \alpha \quad (2.48)$$

3. Define-se o critério de parada como:

$$\alpha = \beta \text{ ou } \alpha - \beta \leq \varepsilon, \quad \varepsilon > 0$$

4. Caso contrário, escolhem-se alguns dos subconjuntos M_i a fim de particioná-los e se obter um particionamento mais refinado de M_0 . Determinam-se limites sobre os novos elementos da nova partição, e repete-se o procedimento.

Uma vantagem da técnica BB é que durante o processo de iteração, pode-se normalmente excluir subconjuntos de S , nos quais o mínimo de f não pode ser alcançado. Esta técnica tem sido aplicada com sucesso em problemas de programação inteira [30].

3

Modelo de surtos proposto

Neste Capítulo o modelo proposto será abordado, apresentando-se seus objetivos; a descrição completa de sua estrutura e suas características especiais relativas à transição entre estados e emissão de símbolos.

3.1.

Objetivos

O objetivo principal do novo modelo de surtos é o de representar canais com memória, minimizando as dificuldades e deficiências comumente encontradas em modelos já existentes, tais como a falta de abrangência no que diz respeito à capacidade de ajustar satisfatoriamente dados gerados por diversas fontes diferentes, ilustrado pelo exemplo de [4] descrito na subseção 1.1. Também pode ser considerado como um exemplo de deficiência, a grande complexidade Matemática, quando se busca melhorias no ajuste através do aumento do número de parâmetros, acarretando longo tempo de processamento computacional para execução dos cálculos para estimação de parâmetros [26].

Para a concretização deste objetivo, buscou-se desenvolver um HMM, com uma estrutura especial, que facilitasse a dedução da expressão analítica da função de verossimilhança e no qual o aumento do número de estados não alterasse a quantidade de parâmetros a serem estimados. Obtivemos um modelo com seis probabilidades de transição, sendo o número de estados o seu sétimo parâmetro. Como apresentado ao longo deste texto, este modelo se mostra bastante flexível para modelagem de erros em surtos oriundos de canais com características físicas diferentes.

3.2.

Descrição do modelo

A idéia central da nossa abordagem é impor uma estrutura a uma cadeia de Markov a fim de reproduzir a geração de padrões típicos que ocorrem em uma sequência afetada por surtos de erros e dissociar o número de parâmetros a serem

ajustados do número de estados da cadeia. O modelo proposto está ilustrado na Figura 3.1 e é composto de $L+3$ estados e dividido em dois sub-modelos denominados *gaps de comprimento limitado* e *gaps de comprimento ilimitado*. Apesar da nomenclatura acima mencionada, os sub-modelos poderão estar envolvidos tanto na geração de *gaps* como de *clusters*, com a diferença básica de que os *gaps* gerados pelo primeiro sub-modelo são limitados ao comprimento máximo de $L-1$. Os estados diretamente envolvidos na geração destes *gaps* (e_4, e_5, \dots, e_{L+2}) formam o grupo denominado *geração de gaps I* na Figura 3.1.

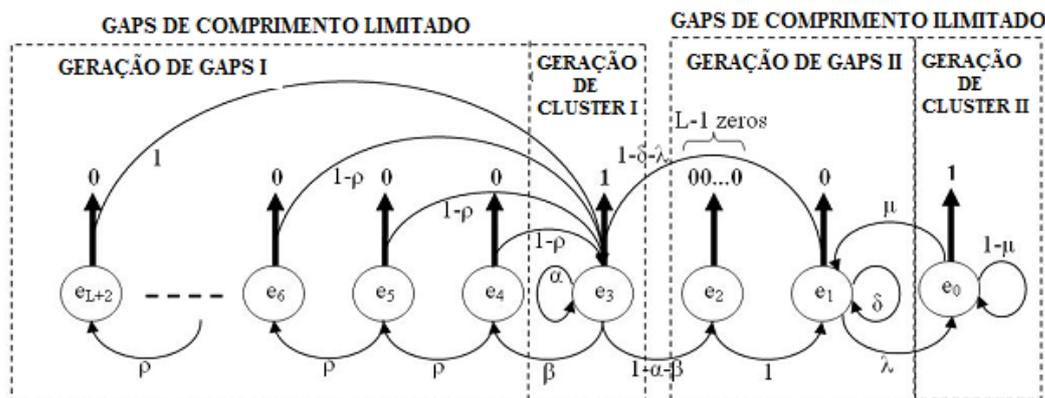


Figura 3.1 – Diagrama de estados do Modelo Proposto

Por outro lado, *gaps* sem restrição de comprimento podem ser gerados pelo sub-modelo então denominado como *gaps de comprimento ilimitado*. Em particular, *gaps* de qualquer comprimento podem ser originado do número de auto-transições realizadas pelo estado e_1 . Além disso, *gaps* de comprimento maior ou igual a L podem ser gerados pelo grupo denominado *geração de gaps II* na Figura 3.1, o qual é composto pelos estados e_1 e e_2 .

Com relação à geração de *clusters* no sub-modelo *gaps de comprimento ilimitado*, é importante notar que *clusters* de qualquer comprimento podem ser obtidos com transições para o estado e_0 e suas auto-transições. *Clusters* sem restrições de comprimento também podem ser gerados a partir do outro sub-modelo, com transições para o estado e_3 seguidas de suas auto-transições.

Cabe notar que existem diferenças significativas nos papéis desempenhados pelos estados e_0 e e_3 , na geração das sequências de erros. Em particular, deve ser observado que o estado e_3 é conectado diretamente a ambos os mecanismos de produção de *gaps*, e nesse sentido ele conecta os sub-modelos

gaps de comprimento limitado e *gaps de comprimento ilimitado*. Por outro lado, e_0 é um estado gerador de clusters com conexão direta apenas com o grupo *geração de gaps II*. A fim de destacar melhor as diferenças entre as gerações de clusters realizadas pelos estados e_3 e e_0 , estes foram rotulados como gerador de cluster I e gerador de cluster II, respectivamente, na Figura 3.1.

Podemos ainda observar na Figura 3.1, que o estado e_2 ao ser alcançado, apresenta como saída uma seqüência de $L-1$ bits sem erro, ou seja, este estado representa o encapsulamento de $L-1$ estados sem erro, todos com probabilidade de transição para o próximo estado igual a 1, como esquematizado na Figura 3.2.

A Figura 3.1 representa o diagrama de estados da última versão do modelo proposto, que foi inicialmente concebido sem o estado e_0 , de forma que o único estado responsável pela emissão de *clusters* de erro era o estado e_3 . O objetivo da inclusão do estado e_0 foi o de permitir maior flexibilidade ao modelo, principalmente no ajuste de determinadas seqüências de dados, que apesar de oriundas de canais com memória, apresentam, além dos surtos de erro, grande ocorrência de bits de erro isolados. Note-se que o estado e_0 , apesar de também ser capaz de gerar *clusters* de erro, pode ser direcionado para, com maior probabilidade, emitir bits de erro isolados, através da atribuição de valores adequados para os parâmetros λ e μ .

3.2.1. Parâmetros

Este modelo é parametrizado pelas probabilidades de transição $\alpha, \beta, \delta, \lambda, \mu, \rho$ e pelo número inteiro positivo L . O número reduzido de parâmetros se deve principalmente pelo fato de as probabilidades transições dos estados e_4 até e_{L+2} serem idênticas, como mostrado na Figura 3.1. Observa-se que o parâmetro L influi no número de bits sem erro emitidos pelo estado e_2 e no número de estados total do modelo, que é de $L+3$ estados.

Portanto, as seguintes restrições podem ser estabelecidas quando se tratar do ajuste destes parâmetros:

$$\begin{aligned}
0 &\leq \alpha, \beta, \delta, \lambda, \mu, \rho \leq 1 \\
0 &\leq \alpha + \beta \leq 1 \\
0 &\leq \delta + \lambda \leq 1 \\
L &= 1, 2, 3, \dots
\end{aligned}
\tag{3.1}$$

3.2.2. Estados

A Tabela 3.1 descreve qual a participação de cada estado na geração da seqüência de saída do modelo. A primeira coluna apresenta a nomenclatura dos estados da direita para a esquerda, em relação a representação gráfica da Figura 3.1, e a segunda coluna descreve os eventos que podem ser gerados em parte ou totalmente pelo correspondente estado na primeira coluna.

Exemplificando, a Tabela 3.1 mostra que um erro isolado, um *cluster* ou os bits com erro de um surto de erro, encontrados em uma seqüência de dados, podem ter sido gerados pelos estados e_0 ou e_3 . Por outro lado, um intervalo entre surtos somente pode ter sido gerado pelos estados e_1 e e_2 em conjunto e os *gaps* internos aos surtos de erro podem ter sido gerados pelos estados e_1 ou pelos estados de e_4 até e_{L+2} .

Tabela 3.1 - Funções dos estados.

Estado	Evento no qual o estado participa da geração
e_0	<i>clusters</i>
e_1	<i>gaps</i> de qualquer tamanho
e_2	<i>gaps</i> tamanho $L-1$
e_3	<i>Clusters</i>
$e_i, 4 \leq i \leq L+2$	<i>gaps</i> de tamanho g , onde $i-3 \leq g \leq L-1$

3.2.3. Matriz de Probabilidades de Transição

A matriz de probabilidades de transição correspondente ao modelo proposto tem dimensão $N=L+3$ e é dada por:

$$P = \begin{bmatrix} 1-\mu & \mu & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \lambda & \delta & 0 & 1-\delta-\lambda & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1-\alpha-\beta & \alpha & \beta & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1-\rho & 0 & \rho & 0 & \dots & 0 \\ 0 & 0 & 0 & 1-\rho & 0 & 0 & \rho & \dots & 0 \\ \dots & 0 \\ 0 & 0 & 0 & 1-\rho & 0 & 0 & 0 & \dots & \rho \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \quad (3.2)$$

Como se pode observar, esta matriz possui muitos elementos nulos e muitos elementos iguais, o que contribui para diminuir a complexidade da análise matemática deste modelo.

3.2.4. Probabilidades dos estados em regime permanente

A seguir, serão deduzidas as expressões das probabilidades de cada estado em regime estacionário, as quais são representadas pelos elementos do vetor $\underline{x} = [x_0 \ x_1 \ \dots \ x_{L+2}]^T$, onde $x_k = \lim_{n \rightarrow \infty} P[\text{estado } k \text{ no instante } n]$.

Assim, o vetor: \underline{x} é solução da equação $\underline{x}^T = \underline{x}^T \cdot P$, onde P representa a matriz de transição do modelo proposto.

Da equação matricial acima decorre que:

$$x_0 = (1-\mu)x_0 + \lambda x_1 \Rightarrow x_1 = \left(\frac{\mu}{\lambda}\right)x_0 \quad (3.3)$$

$$x_1 = \mu x_0 + \delta x_1 + x_2 \Rightarrow x_2 = (1 - \delta - \lambda)x_1 \quad (3.4)$$

$$x_2 = (1 - \alpha - \beta)x_3 \Rightarrow x_3 = \left(\frac{1}{1 - \alpha - \beta} \right) x_2 \quad (3.5)$$

$$x_3 = (1 - \delta - \lambda)x_1 + \alpha x_3 + (1 - \rho) \sum_{k=4}^{L+1} x_k + x_{L+2} \quad (3.6)$$

$$\left. \begin{array}{l} x_4 = \beta x_3 \\ x_5 = \rho x_4 \\ x_6 = \rho x_5 \\ \vdots \\ x_{L+2} = \rho x_{L+1} \end{array} \right\} \Rightarrow x_{4+i} = \beta \rho^i x_3, \quad i = 0, 1, \dots, L-2 \quad (3.7)$$

A soma dos elementos do vetor x tem que ser 1, logo:

$$\sum_{k=0}^{L+2} x_k = x_0 + x_1 + x_2 + x_3 + \sum_{k=4}^{L+2} x_k = 1 \quad (3.8)$$

Manipulando os limites do somatório; calculando a soma dos termos da progressão geométrica formada e chamando a fração resultante, entre parêntesis de $\bar{\rho}$, temos:

$$\sum_{k=4}^{L+2} x_k = \sum_{k=0}^{L-2} x_{4+k} = \sum_{k=0}^{L-2} \beta \rho^k x_3 = \left(\frac{1 - \rho^{L-1}}{1 - \rho} \right) \beta x_3 = \bar{\rho} \beta x_3 \quad (3.9)$$

então:

$$x_0 + x_1 + x_2 + (1 + \bar{\rho} \cdot \beta) x_3 = 1 \quad (3.10)$$

Substituindo sucessivamente as equações (3.5), (3.4) e (3.3) em (3.10), temos o seguinte desenvolvimento:

$$1 = x_0 + x_1 + \left(1 + \frac{1 + \bar{\rho} \cdot \beta}{1 - \alpha - \beta} \right) x_2 \quad (3.11)$$

$$= x_0 + \left[1 + (1 - \delta - \lambda) \left(1 + \frac{1 + \bar{\rho} \cdot \beta}{1 - \alpha - \beta} \right) \right] x_1 \quad (3.12)$$

$$= \left\{ 1 + \left(\frac{\mu}{\lambda} \right) \left[1 + (1 - \delta - \lambda) \left(1 + \frac{1 + \bar{\rho} \cdot \beta}{1 - \alpha - \beta} \right) \right] \right\} x_0 \quad (3.13)$$

Logo:

$$x_0 = \left\{ 1 + \left(\frac{\mu}{\lambda} \right) \left[1 + (1 - \delta - \lambda) \left(1 + \frac{1 + \bar{\rho} \cdot \beta}{1 - \alpha - \beta} \right) \right] \right\}^{-1} \quad (3.14)$$

$$x_1 = \left(\frac{\mu}{\lambda} \right) x_0 \quad (3.15)$$

$$x_2 = (1 - \delta - \lambda) x_1 \quad (3.16)$$

$$x_3 = \left(\frac{1}{1 - \alpha - \beta} \right) x_2 \quad (3.17)$$

$$x_{4+i} = \beta \rho^i x_3, \quad i = 0, 1, \dots, L-2 \quad (3.18)$$

Assim com a equação (3.14), obtivemos a probabilidade no regime estacionário x_0 do estado e_0 em função dos parâmetros do modelo, e por substituições sucessivas, nas equações de (3.15) a (3.18), obtemos as correspondentes probabilidades dos outros estados do modelo.

3.3. Estatísticas de interesse

A seguir serão apresentadas as deduções das expressões analíticas para as principais estatísticas de interesse, consideradas neste trabalho, as quais foram escolhidas, por serem bastante empregadas em outros trabalhos para representação de erros em surtos: probabilidade de erro (P_1); distribuição de gaps (GD); tamanho médio dos gaps (G); distribuição dos clusters de erro (ECD); tamanho médio de

“cluster” de erro (E); distribuição de error free bursts (EFBD); função de correlação de erro em bit (BECF); probabilidade $P(I^m|0)$, que é probabilidade de ocorrerem “m” 1’s dado a ocorrência de um “0” e a Probabilidade $P(0^m|1)$ que é a probabilidade de ocorrerem “m” 0’s dado a ocorrência de um “1”.

3.3.1. Probabilidade de erro

Esta probabilidade é estimada a partir de uma amostra simplesmente pela razão entre o número de bits com erro e o número total de bits na mesma.

Do ponto de vista teórico, entende-se a probabilidade de erro (P_I) como o limite da razão entre o número de erros que ocorrem em uma seqüência de tamanho *limitado* e o valor de seu tamanho, quando *este* tende a infinito.

Esta probabilidade pode então ser expressa em termos das probabilidades de estado em regime estacionário, através da razão entre o somatório das probabilidades de se estar em um dos estados de erro e o somatório das probabilidades de se estar em cada estado, ou seja:

$$P_1 = \frac{x_0 + x_3}{x_0 + x_1 + (L-1)x_2 + x_3 + x_4 + x_5 + \dots + x_{L+2}} = \frac{x_0 + x_3}{1 + (L-2)x_2} \quad (3.19)$$

Substituindo x_2 pela expressão (3.5), obtém-se:

$$P_1 = \frac{x_0 + x_3}{1 + (L-2)(1-\alpha-\beta)x_3} \quad (3.20)$$

Observe que em todos os estados exceto e_2 , a cadeia emite apenas um bit, enquanto que no estado citado, a cadeia emite $L-1$ bits. Assim, na equação (3.19), a probabilidade x_2 está multiplicada pelo fator $L-1$, devido a esta probabilidade ser correspondente a um dos $L-1$ estados encapsulados no único estado e_2 , como mostrado nas Figuras 3.1 e 3.2.

3.3.2. Probabilidades $P[10^m1]$ e $P[01^m0]$

O cálculo destas probabilidades será fundamental, como ponto de partida para dedução das expressões para o cálculo de diversas probabilidades de interesse.

A probabilidade $P[10^m1]$ representa o valor da probabilidade de ocorrer sucessivamente um bit de erro, m bits sem erro, e novamente um bit com erro. Por sua vez, a probabilidade $P[01^m0]$ representa o valor da probabilidade de ocorrer sucessivamente um bit sem erro, m bits com erro, e em seguida um bit sem erro. Os cálculos destas probabilidades são apresentados a seguir.

3.3.2.1. Cálculo de $P[10^m1]$

Na notação utilizada em seguida, consideram-se os eventos ocorrendo da direita para esquerda. A probabilidade procurada é a soma das probabilidades das sequências que podem ser iniciadas nos dois estados geradores de erro e_0 e e_3 . Assim sendo, temos:

$$P[10^m1] = P[10^m e_0] + P[10^m e_3] = \quad (3.21)$$

$$x_0 P[10^m | e_0] + x_3 P[10^m | e_3]$$

A partir do diagrama de estados da Fig. 3.1, obtemos:

$$P[10^m1] = x_0 \mu (1 - \delta) \delta^{m-1} + x_3 \begin{cases} \beta (1 - \rho) \rho^{m-1} & \text{se } m \in \{1, 2, \dots, L-2\} \\ \beta \rho^{L-2} & \text{se } m = L-1 \\ (1 - \alpha - \beta) (1 - \delta) \delta^{m-L} & \text{se } m \geq L \end{cases} \quad (3.22)$$

Considerando-se, que existem *gaps* de comprimentos que variam de 1 a infinito, e usando a expressão acima, obtemos:

$$\sum_{m=1}^{\infty} P[10^m 1] = \mu x_0 + (1 - \alpha)x_3 \quad (3.23)$$

3.3.2.2. Cálculo de $P[01^m 0]$

De modo análogo à análise da probabilidade $P[10^m 1]$, consideramos todas as possibilidades para início da sequência em um dos estados sem erros. Desta forma, obtemos:

$$P[01^m 0] = P[01^m e_1] + \sum_{i=4}^{L+1} P[01^m e_i] + P[01^m e_{L+2}] = \quad (3.24)$$

$$x_1 P[01^m | e_1] + \sum_{i=4}^{L+1} x_i P[01^m | e_i] + x_{L+2} P[01^m | e_{L+2}] =$$

A análise do diagrama de estados nos mostra que:

$$P[01^m | e_k] \begin{cases} \lambda \mu (1 - \mu)^{m-1} + (1 - \delta - \lambda)(1 - \alpha) \alpha^{m-1} & \text{se } k = 1 \\ (1 - \rho)(1 - \alpha) \alpha^{m-1} & \text{se } k \in \{4, 5, \dots, L+1\} \\ (1 - \alpha) \alpha^{m-1} & \text{se } k = L+2 \end{cases} \quad (3.25)$$

$$P[01^m 0] = x_1 P[01^m | e_1] + \sum_{i=4}^{L+1} x_i P[01^m | e_i] + x_{L+2} P[01^m | e_{L+2}] = \quad (3.26)$$

$$x_1 P[01^m | e_1] + \left(\frac{1 - \rho^{L-2}}{1 - \rho} \right) \beta x_3 (1 - \rho)(1 - \alpha) \alpha^{m-1} + \beta \rho^{L-2} x_3 (1 - \alpha) \alpha^{m-1} =$$

$$x_1 [\lambda \mu (1 - \mu)^{m-1} + (1 - \delta - \lambda)(1 - \alpha) \alpha^{m-1}] + \beta x_3 (1 - \alpha) \alpha^{m-1} =$$

Colocando-se as potências em evidência temos:

$$P[01^m0] = [x_1\lambda\mu](1-\mu)^{m-1} + \{[x_1(1-\delta-\lambda) + \beta x_3](1-\alpha)\}\alpha^{m-1} \quad (3.27)$$

Considerando-se o somatório destas probabilidades para *todos os comprimentos de gaps possíveis, obtemos a partir da equação acima:*

$$\sum_{m=1}^{\infty} P[01^m0] = \sum_{m=1}^{\infty} ([x_1\lambda\mu](1-\mu)^{m-1} + \{[x_1(1-\delta-\lambda) + \beta x_3](1-\alpha)\}\alpha^{m-1}) =$$

$$x_1\lambda + x_1(1-\delta-\lambda) + \beta x_3 =$$

$$x_1(1-\delta) + \beta x_3 \quad (3.28)$$

3.3.3. Distribuição de Gaps

Cabe notar inicialmente que se trata aqui da distribuição de probabilidades condicional do comprimento de gaps.

Note-se também que, dado que houve um gap, a probabilidade deste gap ter tamanho m é dada pela seguinte expressão:

$$P[G = m] = \frac{P[10^m1]}{\sum_{k=1}^{\infty} P[10^k1]} \quad (3.29)$$

Substituindo a equação (3.23) em (3.29) resulta:

$$P[G = m] = \frac{P[10^m1]}{\mu x_0 + (1-\alpha)x_3} \quad \text{para } m \geq 1 \quad (3.30)$$

3.3.4. Tamanho médio dos gaps

O tamanho médio dos gaps é dado por:

$$E[G] = \sum_{m=1}^{\infty} mP[G = m] =$$

$$\frac{1}{\mu x_0 + (1 - \alpha)x_3} \sum_{m=1}^{\infty} mP[10^m 1] \quad (3.31)$$

Calculando $\sum_{m=1}^{\infty} mP[10^m 1]$, temos

$$\sum_{m=1}^{\infty} mP[10^m 1] = x_0 \mu (1 - \delta) \sum_{m=1}^{\infty} m \delta^{m-1} + x_3 \beta (1 - \rho) \sum_{m=1}^{L-2} m \rho^{m-1} +$$

$$+ x_3 (L - 1) \beta \rho^{L-2} + x_3 (1 - \alpha - \beta) (1 - \delta) \sum_{m=1}^{\infty} m \delta^{m-L} =$$

$$x_0 \mu \frac{1}{1 - \delta} + x_3 \beta \frac{1 - \rho^{L-2} [\rho + (L - 1)(1 - \rho)]}{1 - \rho} +$$

$$+ x_3 (L - 1) \beta \rho^{L-2} + x_3 (1 - \alpha - \beta) \left(L + \frac{\delta}{1 - \delta} \right) =$$

$$x_0 \mu \frac{1}{1 - \delta} + x_3 \beta \left(\frac{1 - \rho^{L-2}}{1 - \rho} \right) + x_3 (1 - \alpha - \beta) \left(L + \frac{\delta}{1 - \delta} \right) \quad (3.32)$$

Fazendo $\left(\frac{1 - \rho^{L-2}}{1 - \rho} \right) = \tau$, resulta:

$$\sum_{m=1}^{\infty} mP[10^m 1] = x_0 \frac{\mu}{1 - \delta} + x_3 \beta \tau + x_3 (1 - \alpha - \beta) \left(L + \frac{\delta}{1 - \delta} \right). \quad (3.33)$$

Substituindo a equação (3.33) em (3.31), chegamos a expressão final de $E[G]$:

$$E[G] = \frac{\left(\frac{\mu}{1-\delta}\right)x_0 + \left(\beta\tau + (1-\alpha-\beta)\left(L + \frac{\delta}{1-\delta}\right)\right)x_3}{\mu x_0 + (1-\alpha)x_3} \quad (3.34)$$

3.3.5. Distribuição dos clusters de erro

Trata-se aqui também de uma distribuição de probabilidade condicional. Dada a ocorrência de um cluster, é fácil concluir que a probabilidade condicional de que este o seu comprimento (EC) seja m é dada pela razão entre a equação (3.27) e a equação (3.28). Ou seja:

$$P[EC = m] = \frac{P[01^m0]}{\sum_{k=1}^{\infty} P[01^k0]} = \quad (3.35)$$

Substituindo o denominador acima pela expressão (3.28), obtemos:

$$P[EC = m] = \frac{P[01^m0]}{\frac{\mu(1-\delta)}{\lambda}x_0 + \beta x_3} \quad \text{para } m \geq 1 \quad (3.36)$$

3.3.6. Tamanho médio dos clusters de erro

O tamanho médio dos clusters de erro é dado por:

$$E[EC] = \sum_{m=1}^{\infty} mP[EC = m] =$$

$$\frac{1}{\frac{\mu(1-\delta)}{\lambda}x_0 + \beta x_3} \sum_{m=1}^{\infty} mP[01^m0] =$$

$$\frac{x_0 + x_3}{\frac{\mu(1-\delta)}{\lambda}x_0 + \beta x_3}$$

Considerando que x_0 é dada por (3.3), obtemos:

$$E[EC] = \frac{x_0 + x_3}{(1 - \delta)x_1 + \beta x_3} \quad (3.37)$$

3.3.7. Distribuição de intervalos entre surtos

Cabe notar inicialmente que um intervalo entre surtos pode ser visto como um *gap* de tamanho superior ou igual a um determinado parâmetro X .

Na análise que se segue considera-se apenas o caso em que o valor de X é igual ao parâmetro L do modelo proposto. Além disso, a variável aleatória comprimento de intervalo entre surtos é denotada por EFB, em consonância com o termo “error-free burst” frequentemente usado na literatura internacional.

Desta forma, temos:

$$P[EFB = m] = P[G = m | G \geq L] = \frac{P[G = m]}{P[G \geq L]} \quad (3.38)$$

onde:

$$P[G \geq L] = \frac{1}{\mu x_0 + (1 - \alpha)x_3} \sum_{m=L}^{\infty} P[10^m 1] = \frac{\mu \delta^{L-1} x_0 + (1 - \alpha - \beta)x_3}{\mu x_0 + (1 - \alpha)x_3} \quad (3.39)$$

Substituindo as equações (3.30) e (3.39) em (3.38), temos:

$$P[EFB = m] = \frac{P[10^m 1]}{\mu \delta^{L-1} x_0 + (1 - \alpha - \beta)x_3}, \quad m \geq L \quad (3.40)$$

3.3.8. Distribuição de gaps no interior de surtos

Considera-se aqui o que poderia ser denominado em Inglês “error-free cluster”, ou seja, o gap no interior de um surto, cujo comprimento será denotado por *EFC*.

Considerando a definição de surto para o caso particular em que o parâmetro X é igual a L , a variável aleatória EFC poderá assumir valores entre 1 e L com probabilidades dadas por:

$$P[EFC = m] = P[G = m | G < L] = \frac{P[G = m]}{P[G < L]} \quad (3.41)$$

Usando a expressão (3.39), verificamos que:

$$P[G < L] = \frac{1}{\mu x_0 + (1 - \alpha)x_3} \sum_{m=1}^{L-1} P[10^m 1] = \frac{\mu(1 - \delta^{L-1})x_0 + \beta x_3}{\mu x_0 + (1 - \alpha)x_3}$$

Substituindo a expressão acima juntamente com a equação (3.30) na equação (3.41), obtemos finalmente:

$$P[EFC = m] = \frac{P[10^m 1]}{\mu(1 - \delta^{L-1})x_0 + \beta x_3} \quad (3.42)$$

3.3.9. Função Auto-correlação

A função auto-correlação do processo de erros será aqui denotada por $\phi(m)$. Considerando que se trata de um processo discreto assumindo valores no conjunto $\{0, 1\}$, é fácil concluir que $\phi(m) = P[1X^{m-1}1]$.

Por outro lado, levando em conta que no modelo proposto os erros são gerados apenas nos estados e_0 e e_3 , temos:

$$\begin{aligned} \phi(m+1) &= P[1X^m 1] = P[1X^m e_0] + P[1X^m e_3] = \\ &= x_0 P[1X^m | e_0] + x_3 P[1X^m | e_3] \end{aligned} \quad (3.43)$$

Podemos estabelecer também a seguinte lei de recorrência nos termos do lado direito da expressão acima:

$$P[1X^m | e_k] = P[1X^{m-1}0 | e_k] + P[1X^{m-1}1 | e_k] =$$

$$P[0 | e_k]P[1X^{m-1} | 0, e_k] + P[1 | e_k]P[1X^{m-1} | 1, e_k] \quad (3.44)$$

Cabe notar que o modelo possibilita que saibamos quais os estados “vizinhos” ao estado e_k cujas saídas sejam 0 ou 1. Denominando estes estados respectivamente de e_k^0 e e_k^1 , temos:

$$P[1X^m | e_k] = P[0 | e_k]P[1X^{m-1} | e_k^0] + P[1 | e_k]P[1X^{m-1} | e_k^1] \quad (3.45)$$

Convém lembrar, neste ponto, que e_2 corresponde ao encapsulamento de L-1 estados com transições sucessivas de probabilidade um como apresentado na Figura 3.2:

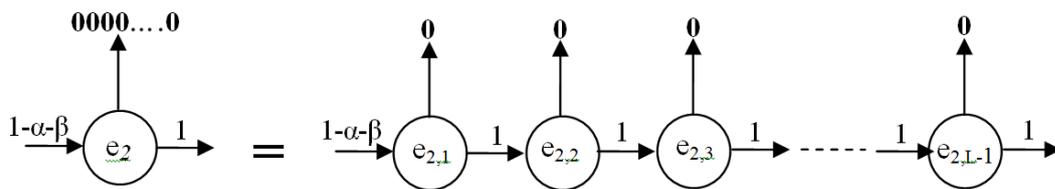


Figura 3.2 – Diagrama de estados equivalente ao estado e_2

Analisando agora as probabilidades condicionais dos eventos ($1X^m$), relativas a cada um dos estados do modelo “aberto”, temos:

$$P[1X^m | e_0] = \mu P[1X^{m-1} | e_1] + (1 - \mu)P[1X^{m-1} | e_0] \quad (3.46)$$

$$P[1X^m | e_1] = \delta P[1X^{m-1} | e_1] + \lambda P[1X^{m-1} | e_0] + (1 - \delta - \lambda)P[1X^{m-1} | e_3] \quad (3.47)$$

$$P[1X^m | e_{2,k}] = P[1X^{m-1} | e_{2,k+1}] \quad \text{para } k = 1, 2, \dots, L-2 \quad (3.48)$$

$$P[1X^m | e_{2,L-1}] = P[1X^{m-1} | e_1] \quad (3.49)$$

$$P[1X^m | e_3] = \alpha P[1X^{m-1} | e_3] + \beta P[1X^{m-1} | e_4] + (1 - \alpha - \beta) P[1X^{m-1} | e_{2,1}] \quad (3.50)$$

$$P[1X^m | e_k] = \rho P[1X^{m-1} | e_{k+1}] + (1 - \rho) P[1X^{m-1} | e_3] \quad \text{se } k = 4, 5, \dots, L+1 \quad (3.51)$$

$$P[1X^m | e_{L+2}] = P[1X^{m-1} | e_3] \quad (3.52)$$

Pode-se verificar que as condições iniciais para as recursões acima apresentadas são:

$$P[1X^0 | e_k] = \begin{cases} 1 - \mu & \text{se } k = 0 \\ 1 - \delta & \text{se } k = 1 \\ 0 & \text{se } k = 2 \\ \alpha & \text{se } k = 3 \\ 1 - \rho & \text{se } k = 4, 5, \dots, L+1 \\ 1 & \text{se } k = L+1 \end{cases} \quad (3.53)$$

3.3.10.

Probabilidade $P[1^m|0]$

Esta é a probabilidade de ocorrerem “m” 1’s após a ocorrência de um “0”. Esta estatística tem sido amplamente empregada em diversos trabalhos nesta área. Uma vantagem obtida no seu uso em relação ao da distribuição de comprimento de surtos, é a não dependência do número X , da definição de surto de erros do CCITT [1].

Considerando novamente o diagrama de estados da Figura 3.1, pode-se concluir que

$$P[1^m | 0] = \frac{x_1 P[1^m | 0, e_1] + \sum_{k=4}^{L+2} x_k P[1^m | 0, e_k]}{x_1 + (L-1)x_2 + \sum_{k=4}^{L+2} x_k} \quad (3.54)$$

$$\begin{aligned}
&= \frac{x_1 \left[(1-\delta-\lambda)\alpha^{m-1} + \lambda(1-\mu)^{m-1} \right] + \sum_{k=4}^{L+2} x_k (1-\rho)\alpha^{m-1} + x_{L+2}\alpha^{m-1}}{1-x_0 + (L-2)x_2 - x_3} \\
&= \frac{x_1 \left[(1-\delta-\lambda)\alpha^{m-1} + \lambda(1-\mu)^{m-1} \right] + (1-\rho)\alpha^{m-1} \left(\frac{1-\rho^{L-2}}{1-\rho} \right) \beta x_3 + \beta \rho^{L-2} x_3 \alpha^{m-1}}{1-x_0 + (L-2)x_2 - x_3} \\
&= \frac{x_1 \left[(1-\delta-\lambda)\alpha^{m-1} + \lambda(1-\mu)^{m-1} \right] + \beta x_3 \alpha^{m-1}}{1-x_0 + (L-2)x_2 - x_3} \\
&= \frac{x_1 \left[(1-\delta-\lambda) + \beta x_3 \right] \alpha^{m-1} + [x_1 \lambda] (1-\mu)^{m-1}}{1-x_0 + (L-2)x_2 - x_3}
\end{aligned}$$

Explicitando as potências chegamos a seguinte expressão:

$$P[1^m | 0] = \left[\frac{(1-\alpha)x_3}{1-x_0 + (L-2)x_2 - x_3} \right] \alpha^{m-1} + \left[\frac{\mu x_0}{1-x_0 + (L-2)x_2 - x_3} \right] (1-\mu)^{m-1} \quad (3.55)$$

3.3.11. Probabilidade $P[0^m/1]$

Esta é a probabilidade de ocorrerem “m” 0’s após a ocorrência de um “1”. Em conjunto com a anterior, esta estatística também tem sido amplamente empregada em trabalhos sobre erros em surtos. De maneira similar, a vantagem obtida com o seu uso, em relação ao da distribuição de intervalos entre surtos, vem também do fato de não requerer a atribuição de um valor ao parâmetro X , da definição de surto de erro do CCITT [1].

Esta probabilidade dá origem a uma função composta, onde se deve considerar dois casos: para valores de m menores e maiores ou igual a L , sendo que em cada caso considera-se a soma das probabilidades de se gerar m zeros partindo-se de ambos os estados geradores de erros (x_0 e x_3).

$$P[0^m | 1] = \begin{cases} \frac{x_0 \mu \delta^{m-1} + x_3 [(1 - \alpha - \beta) + \beta \rho^{m-1}]}{x_0 + x_3}, & m < L \\ \frac{x_0 \mu \delta^{m-1} + x_3 (1 - \alpha - \beta) \delta^{m-L}}{x_0 + x_3}, & m \geq L \end{cases} \quad (3.56)$$

4

Metodologia para ajuste dos parâmetros do modelo proposto

Neste Capítulo, será apresentado o método adotado para o ajuste do modelo proposto às sequências de dados que se deseja representar.

A estrutura especial do modelo proposto possibilitou a dedução das expressões analíticas para a função de verossimilhança e para as estatísticas de interesse, o que por sua vez possibilitou o desenvolvimento de um método para a estimação dos parâmetros, usando o critério da máxima verossimilhança (ML). A estimação de parâmetros foi realizada utilizando-se as ferramentas de otimização do MatLab.

Em princípio, poderíamos arbitrar parâmetros iniciais para a estimação ML, porém como se trata uma função não linear de seis parâmetros e conforme observado nos experimentos, a existência de diversos máximos locais pode levar o processo de otimização a convergir para ótimos locais distantes do ótimo global procurado, produzindo resultados inadequados.

Esta deficiência foi minimizada pelo emprego do PSO, que por ser uma técnica exaustiva de busca é menos suscetível a convergir para pontos de máximo locais.

Cabe notar que o método PSO poderia em princípio vir a consumir considerável esforço computacional, caso se buscasse melhores resultados por meio do aumento do número de “partículas” utilizadas, o que reduziria uma das principais vantagens obtidas com o uso do modelo proposto que é o tempo reduzido de cálculo com a utilização da otimização clássica.

No entanto, adotando-se um relativamente reduzido número de “partículas”, verificou-se que, com esforço computacional consideravelmente baixo, se pode obter valores de parâmetros adequados para inicialização da ferramenta de otimização do Matlab, aumentando significativamente as chances de convergência para o valor máximo global da função de verossimilhança.

4.1. Função de Verossimilhança do modelo

Apresentaremos a dedução de uma expressão analítica da função de verossimilhança dos parâmetros do modelo proposto, numa forma que facilita o uso de métodos de otimização de funções não lineares, para sua maximização.

Um fator que contribui para o aumento de eficiência da estimação é que, ao invés de se tratar a amostra de erros bit a bit, trabalha-se com a sequência de tamanhos de *gaps* e *clusters* correspondentes.

Sem perda de generalidade, admite-se que a sequência de observações se inicia por um cluster e é composta de $K+1$ blocos, sendo cada bloco correspondente a um par *gap* e *cluster* cujos comprimentos são denotados por m_k e n_k , respectivamente. Denota-se por S_k o estado em que o modelo se encontra depois de emitido o k -ésimo bloco, como mostrado na Figura 4.2.

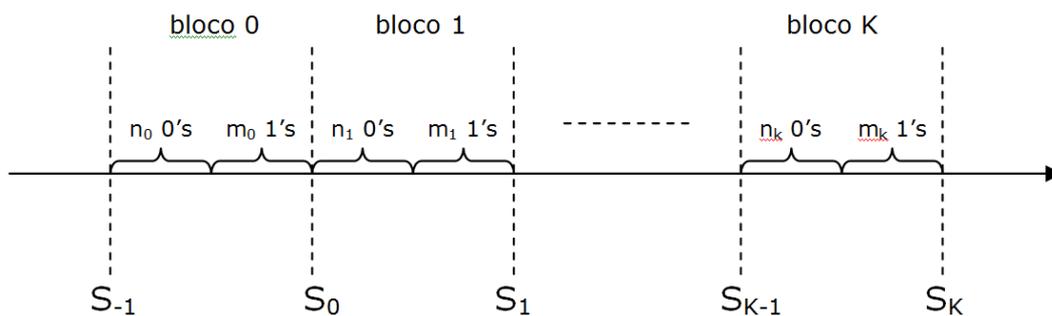


Figura 4.1 – Modelo de transição

Considerando-se as características do modelo, é fácil concluir que S_k é igual a e_0 ou e_3 , para qualquer valor de k .

Do modelo de transição da Figura 4.1, podemos definir dois vetores:

$$\underline{m}_K = (m_0, m_1, \dots, m_K) \quad \text{e} \quad \underline{n}_K = (n_0, n_1, \dots, n_K)$$

cujos elementos são os sucessivos comprimentos de clusters e comprimentos de gaps, respectivamente, retirados da sequência.

A função verossimilhança pode então ser dada por:

$$V(\underline{m}_K, \underline{n}_K) = P\left[\bigcap_{k=0}^K (1^{m_k} 0^{n_k}) \mid 1\right] = \sum_{i \in \{0,3\}} P\left[S_K = e_i; \bigcap_{k=0}^K (1^{m_k} 0^{n_k}) \mid 1\right] \quad (4.1)$$

Observe-se na equação (4.1) que o somatório representa a soma das probabilidades, considerando os dois casos de emissão de 1 dos estados e_0 e e_3 , ao final de cada bloco.

Definindo a variável auxiliar $a_t(i)$ como

$$a_t(i) = P\left[S_t = e_i; \prod_{k=0}^t (1^{m_k} 0^{n_k}) \mid 1\right] \quad (4.2)$$

Tem-se de imediato que:

$$V(\underline{m}_K, \underline{n}_K) = \sum_{i \in \{0,3\}} a_K(i) \quad (4.3)$$

Considerando-se o estado S_{t-1} , é fácil verificar que a variável $a_t(i)$ pode ser expressa como se segue:

$$\begin{aligned} a_t(i) &= \sum_{j \in \{0,3\}} P\left[S_t = e_i; (1^{m_t} 0^{n_t}); S_{t-1} = e_j; \prod_{k=0}^{t-1} (1^{m_k} 0^{n_k}) \mid 1\right] \\ &= \sum_{j \in \{0,3\}} P\left[S_{t-1} = e_j; \prod_{k=0}^{t-1} (1^{m_k} 0^{n_k}) \mid 1\right] \cdot P\left[S_t = e_i; (1^{m_t} 0^{n_t}) \mid S_{t-1} = e_j; \prod_{k=0}^{t-1} (1^{m_k} 0^{n_k}); 1\right] \end{aligned} \quad (4.4)$$

Obtém-se assim uma recursão forma recursiva para $a_t(i)$, expressa a seguir:

$$a_t(i) = \sum_{j \in \{0,3\}} a_{t-1}(j) \cdot P\left[S_t = e_i; (1^{m_t} 0^{n_t}) \mid S_{t-1} = e_j\right] \quad (4.5)$$

Mostra-se no Apêndice A, que a função de verossimilhança acima apresentada também pode ser reescrita em forma matricial como:

$$V(\underline{m}_K, \underline{n}_K) = [1 \quad 1] \cdot \left\{ \prod_{k=0}^K P(m_k, n_k) \right\} \cdot \begin{bmatrix} \bar{P} \\ 1 - \bar{P} \end{bmatrix}, \quad (4.6)$$

$$\text{sendo } \bar{P} = \frac{1}{1 + \left(\frac{\mu}{\lambda}\right) \cdot \left(\frac{1 - \delta - \lambda}{1 - \alpha - \beta}\right)} \quad (4.7)$$

e $P(m,n)$ é uma matriz quadrada de ordem 2 expressa como

$$P(m,n) = \begin{bmatrix} P_{00}(m,n) & P_{03}(m,n) \\ P_{30}(m,n) & P_{33}(m,n) \end{bmatrix}, \quad (4.8)$$

cujos elementos são expressos a seguir:

$$P_{ij}(m,n) = \begin{cases} \mu \cdot \delta^{n-1} \cdot \lambda \cdot (1 - \mu)^{m-1} & i = 0, j = 0 \\ \mu \cdot \delta^{n-1} \cdot (1 - \delta - \lambda) \cdot \alpha^{m-1} & i = 3, j = 0 \\ (1 - \alpha - \beta) \cdot \delta^{n-L} \cdot \lambda \cdot (1 - \mu)^{m-1} & i = 0, j = 3, n \geq L \\ 0 & i = 0, j = 3, n < L \\ (1 - \alpha - \beta) \cdot \delta^{n-L} \cdot (1 - \delta - \lambda) \cdot \alpha^{m-1} & i = 3, j = 3, n \geq L \\ \beta \cdot \rho^{n-1} \cdot \alpha^{m-1} & i = 3, j = 3, n = L - 1 \\ \beta \cdot \rho^{n-1} \cdot (1 - \rho) \cdot \alpha^{m-1} & i = 3, j = 3, n < L - 1 \end{cases}. \quad (4.9)$$

4.2. Metodologia de otimização

O método de otimização do Matlab implementado pela função *fmincon* foi utilizado para maximizar a função de verossimilhança acima, na sua dependência dos parâmetros contínuos do modelo proposto $\alpha, \beta, \delta, \lambda, \mu, \rho$, ou seja, para valores pré-fixados do parâmetro inteiro L .

Como esta ferramenta do Matlab lida diretamente com problemas de minimização, estabeleceu-se o seguinte problema de otimização:

$$\underset{\theta \in R_6}{Min} -V(\underline{m}_K, \underline{n}_K | \underline{\theta}), \quad \text{onde } \underline{\theta} = [\alpha, \beta, \delta, \lambda, \mu, \rho] , \quad (4.10)$$

sujeito às restrições abaixo:

$$\begin{aligned} 0 &\leq \alpha, \beta, \delta, \lambda, \mu, \rho \leq 1 \\ 0 &\leq \delta + \lambda \leq 1 \\ 0 &\leq \alpha + \beta \leq 1 \end{aligned} \quad (4.11)$$

Este procedimento de otimização foi realizado diversas vezes, variando-se o parâmetro L em uma faixa pré-determinada de valores, para em seguida buscar também o melhor valor deste parâmetro, no sentido de máxima verossimilhança, e assim completar o processo de estimação dos parâmetros do modelo.

Deve ser ressaltado que esta operação de busca do melhor L se torna simples tendo-se em vista que este é um parâmetro inteiro e positivo. Por outro lado, o processamento relacionado à otimização nos parâmetros contínuos, para cada valor de L , consome tempo reduzido, devido à compactação e recursividade no cálculo da função verossimilhança. Note-se que estas características da função de verossimilhança se devem à estrutura especial do modelo proposto e também à representação dos dados em termos de sequências de tamanhos de *gaps* e de *clusters*.

Como mencionado anteriormente, um ponto crítico no método acima descrito foi a escolha dos valores dos parâmetros a serem utilizados como partida do algoritmo de otimização, que acabou sendo obtida através da técnica PSO.

5 Resultados

Foi avaliada a capacidade do modelo proposto de capturar as propriedades estatísticas das seqüências afetadas por erros em surtos produzidas por alguns sistemas de comunicações simulados. Baseado nestes dados realizou-se a estimação ML dos parâmetros do modelo proposto, e com o modelo assim ajustado gerou-se uma nova seqüência de erros. Em seguida, utilizou-se esta seqüência para estimar diversas estatísticas de grande interesse no estudo de canais com memória [31], comparando-as com as mesmas estatísticas levantadas dos dados originais. Para facilitar a apresentação gráfica dos resultados, as definições das primeiras três estatísticas foram ligeiramente alteradas, conforme mostrado abaixo, mas seus nomes foram preservados. As estatísticas empregadas são descritas a seguir:

- a. Distribuição de *gaps* (GD, do termo em Inglês *gap distribution*) – é expressa através da função distribuição complementar do comprimentos de *gaps* G , a qual é dada por $P(G \geq m)$;
- b. Distribuição de surtos de erros (EBD, do termo em Inglês *error burst distribution*) – expressa pela função distribuição complementar do comprimento de surtos EB , dada por $P(EB \geq m)$;
- c. Distribuição de comprimentos de intervalos entre surtos (EFBD, da nomenclatura em Inglês *error-free burst distribution*) – expressa pela função distribuição complementar do comprimentos de i intervalos entre surtos EFB , sendo dada por $P(EFB \geq m)$;
- d. Função autocorrelação de erros de bit (BECF, do termo em Inglês *bit error correlation function*) – É a probabilidade condicional de que o m -ésimo bit recebido após um bit de erro esteja errado.
- e. Distribuição de probabilidade de erro por bloco, $P(m,n)$ probabilidade de ocorrerem m erros em um bloco de tamanho n .

f. Probabilidade $P(0^m|1)$ (EFRD, do termo em Inglês *error-free run distribution*) – probabilidade condicional de ocorrerem m bits corretos, após a ocorrência de um bit errado.

Para efeitos de comparação, também foi avaliado um modelo consagrado na literatura, que é o *Single Error State* (SES) de Fritchman. Em particular, utilizou-se um modelo SES de 5 estados, que tem oito parâmetros, por ser este número o valor mais próximo e maior do que o número de parâmetros do modelo proposto. Com esta escolha procurou-se a igualdade de condições ao se comparar os dois modelos, com vantagem de um parâmetro a mais para o modelo de Fritchman.

Para maior efetividade das comparações, o critério utilizado para ajuste dos parâmetros do modelo SES de Fritchman foi a mesma aplicada ao modelo proposto. De forma mais específica, os parâmetros deste modelo foram estimados utilizando-se o bem conhecido Algoritmo de Baum-Welch, abordado no capítulo 2. Doravante este modelo será mencionado apenas como modelo de Fritchman.

Os dados utilizados nos experimentos foram simulados em computador, sendo que três casos típicos de fontes geradoras de erros em surtos foram empregados: o equalizador DFE (*decision feedback equalizer*), o qual será referido como Caso I, na apresentação dos resultados; a transmissão de sinais PSK-4 num canal variante no tempo com desvanecimento plano, que será referido como Caso II, e o decodificador de Viterbi, referido como Caso III.

Para cada estatística de interesse será apresentada uma comparação entre os dados originais; os dados gerados pelo modelo proposto e os dados gerados pelo modelo de Fritchman.

As comparações apresentadas a seguir mostram que em alguns casos tanto o modelo proposto como o modelo de Fritchman apresentam bons resultados, com pequena vantagem para o primeiro, enquanto que em diversos outros casos o modelo proposto demonstrou ser visivelmente melhor.

Deve ser ressaltado de antemão que o modelo proposto apresentou ainda uma vantagem em relação ao de Fritchman, no que diz respeito à velocidade de realização dos cálculos para estimação de seus parâmetros, aproximadamente 10 a 30 vezes maior.

Observou-se também que o modelo proposto propicia uma menor dependência do tamanho da amostras de dados para obter ajustes satisfatórios, sendo possível com este modelo atingir ajustes compatíveis com os produzidos pelo modelo de Fritchman utilizando amostras de dados menores.

5.1. Estimação do parâmetro L

Como mencionado antes, a função objetivo utilizada na estimação dos parâmetros foi definida como o simétrico do logaritmo da função de verossimilhança, e lidou-se com um problema de minimização desta função. Assim sendo, são apresentadas a seguir curvas correspondentes aos valores desta função objetivo otimizada em relação aos parâmetros contínuos do modelo proposto, correspondentes a diferentes valores do parâmetro inteiro L pertencentes a um conjunto pré-fixado. Desta forma tomamos como ótimos o valor de L correspondente ao melhor valor da métrica ML e os correspondentes parâmetros contínuos. Na Figura 5.1 são mostrados os resultados obtidos nos casos I e II e a Figura 5.2 mostra os resultados do Caso III. Uma rápida comparação entre os três casos indica haver maior sensibilidade da métrica ML com as variações do valor de L no Caso I.

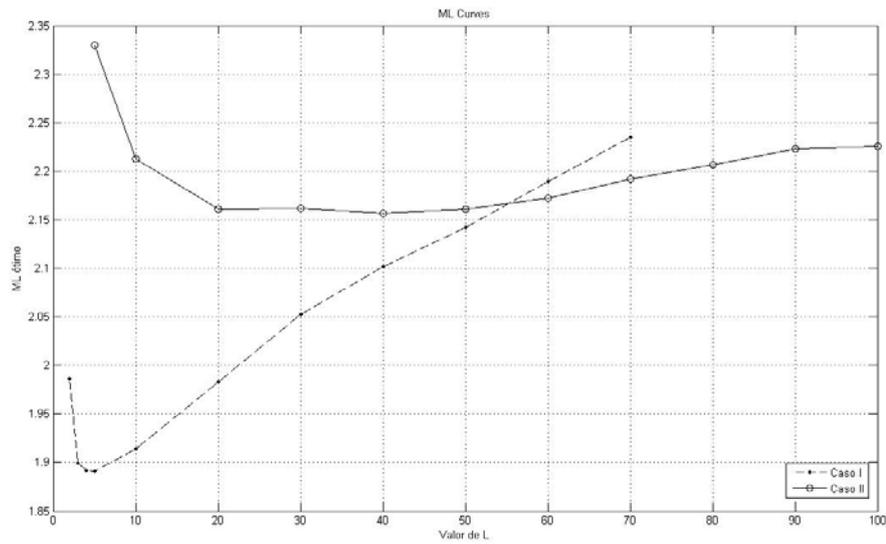


Figura 5.1 – Levantamento do melhor valor do parâmetro L para o modelo proposto para os casos I e II.

Como o algoritmo de otimização empregado busca o ponto de mínimo da função objetivo, assim observamos na Figura 5.1, que o melhor valor de L é 5 e 40, respectivamente para os casos I e II, e na Figura 5.2, que o L ótimo é 30 para o Caso III.

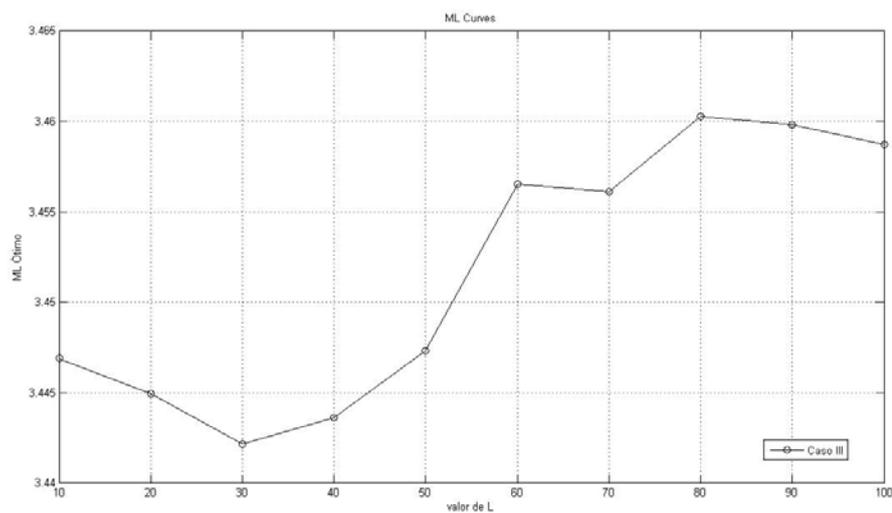


Figura 5.2 – Levantamento do melhor valor parâmetro L para o modelo proposto para o Caso III.

5.2. Validação do método de ajuste do modelo

São apresentados em seguida alguns resultados de validação do método de estimação de parâmetros utilizado, onde se considera como fonte geradora o próprio modelo proposto.

De forma mais específica, o ajuste dos parâmetros deste modelo foi realizado a partir de uma sequência de dados gerada pelo mesmo modelo, arbitrando-se valores para os seus parâmetros. As estatísticas GD, EFRD, EBD, EFBD, $P(m,300)$ e BECF foram estimadas a partir da amostra de erros original e a partir de uma amostra gerada pelo modelo ajustado, sendo ambas as amostras de 2 Mbits. Para levantamento das estatísticas EBD, EFBD e fez-se o valor do limiar *inter-bursts* X igual ao valor do parâmetro L do modelo, igual a 80. Os resultados obtidos estão mostrados na Figura 5.3, que apresenta ajustes excelentes, assim comprovando o correto funcionamento dos algoritmos e técnicas empregadas.

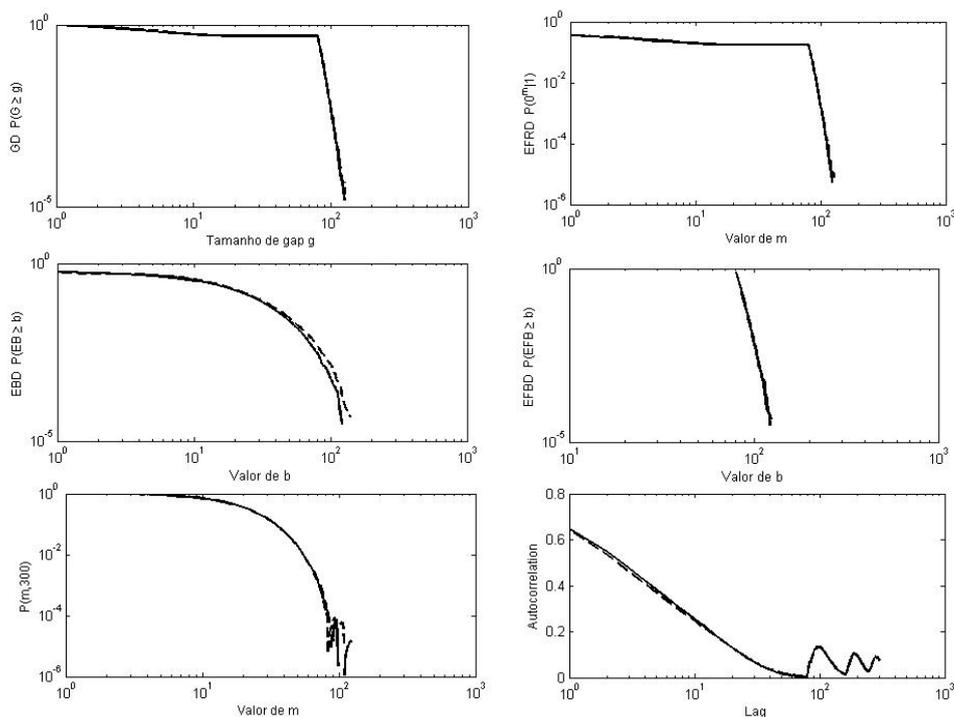


Figura 5.3 – Resultados de validação do método de ajuste dos parâmetros do modelo proposto.

5.3. Equalizador DFE (Caso I)

Neste caso, foi adotado um equalizador DFE com coeficientes calculados partir de uma estimativa de canal ideal, seguindo uma abordagem de mínimo erro médio quadrático. A modulação empregada foi o PSK-2 com relação E_b/N_0 igual a 10 dB; o modelo de canal adotado foi o modelo C do Proakis [13]; o valor do atraso de decisão foi fixado em 10; os números de coeficientes dos filtros avante (“*feedforward*”) e de realimentação (“*feedback*”) foram 21 e 10, respectivamente. O tamanho da sequência de dados original foi de 2 Mbits. O limiar *inter-bursts* X foi fixado em 5, tendo sido este valor escolhido como por ser o melhor valor de L , no sentido ML, com base no gráfico da Figura 5.1.

A Figura 5.4 mostra os resultados obtidos na estimação da distribuição de surtos de erro para este caso. Nota-se que o modelo proposto e o modelo de Fritchman produziram bons ajustes aos dados originais, sendo que o modelo proposto possibilitou uma aproximação um pouco melhor da curva referente aos dados originais.

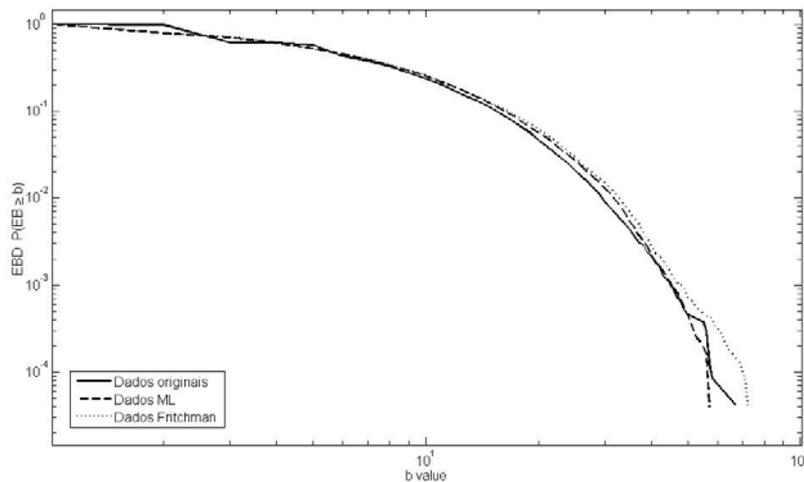


Figura 5.4 – Estimativas da distribuição de surtos de erros para o Caso I.

A Figura 5.5 apresenta a função autocorrelação estimada. Vê-se nela que o modelo proposto mostrou ter uma precisão um pouco melhor para capturar esta estatística, em relação ao modelo de Fritchman.

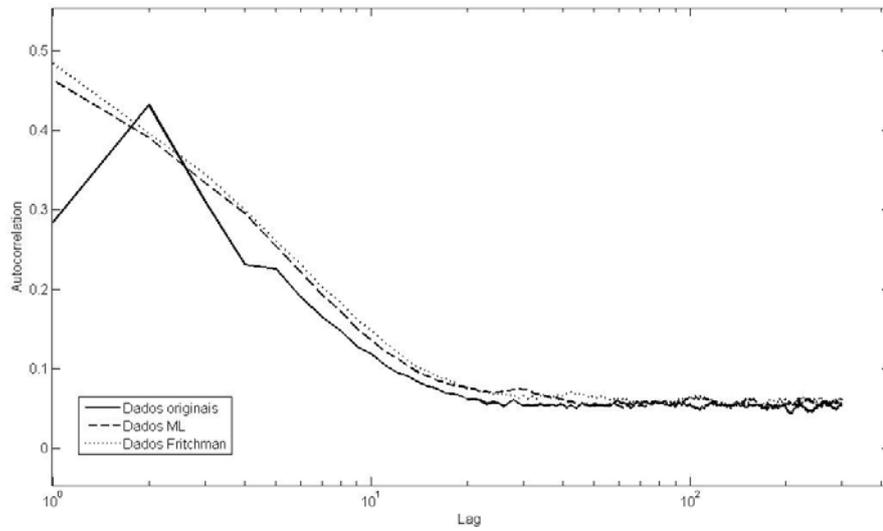


Figura 5.5 – Estimativas da função autocorrelação para o Caso I.

A Figura 5.6, por sua vez, mostra que o modelo proposto obteve excelente desempenho na captura da distribuição de gaps. Aqui nenhuma diferença significativa de ajuste foi observada em relação ao obtido com o modelo de Fritchman.

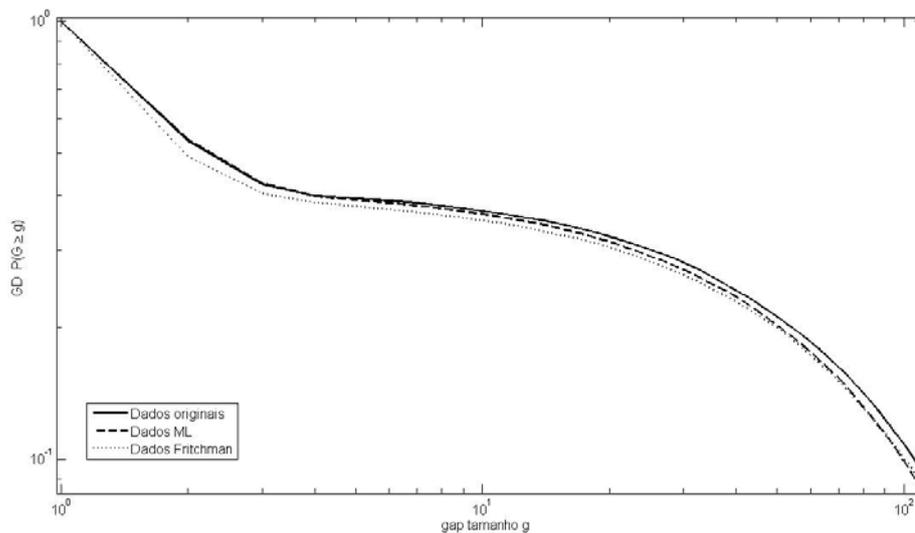


Figura 5.6 – Estimativas da distribuição de gaps para o Caso I.

A Figura 5.7 mostra que, em relação à distribuição de comprimentos de intervalos entre surtos, ambos os modelos obtiveram excelente desempenho,

sendo que o modelo proposto obteve uma sutil vantagem para valores de b menores do que 18, e que para valores maiores o modelo de Fritchman capturou mais precisamente esta estatística.

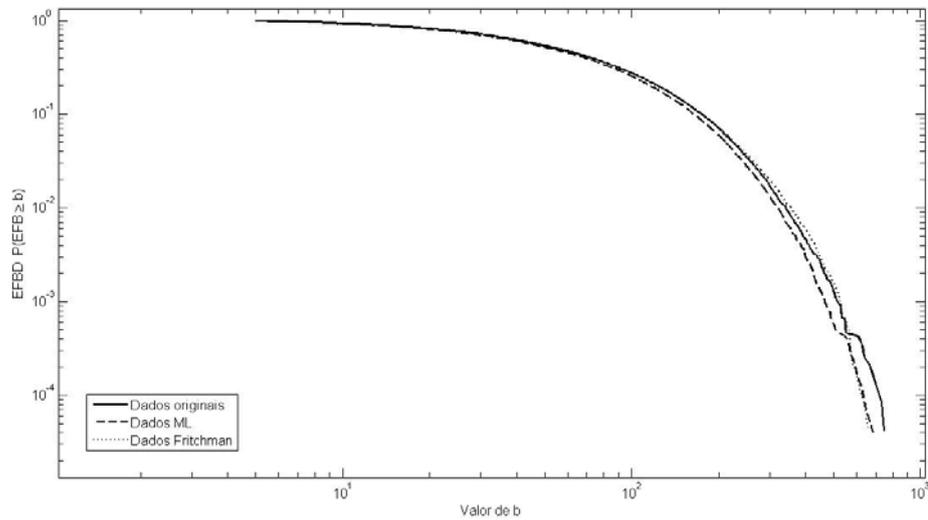


Figura 5.7 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso I.

A Figura 5.8 representa a estimação da função probabilidade de erros por bloco, $P(m,n)$, para um bloco de tamanho 300 bits ($n=300$). Nota-se que o modelo proposto mostrou um desempenho um pouco melhor que o modelo de Fritchman em capturar esta estatística, e que esta vantagem é mais acentuada para valores de m maiores do que 40.

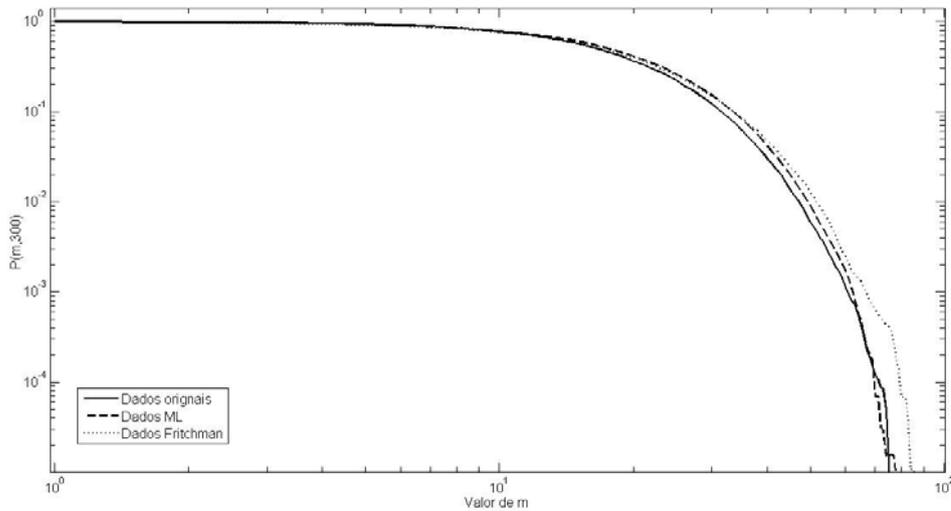


Figura 5.8 – Estimativas da função Probabilidade de erros por bloco para um bloco de tamanho 300 ($P(m,300)$) para o Caso I.

Em suma, os resultados obtidos para Caso I indicam o modelo proposto produziu apenas pequenas melhorias nos ajustes das estatísticas de interesse, em relação ao modelo utilizado como comparação. Porém deve ser lembrada a maior eficiência de ajuste do modelo proposto, que viabilizou um tempo de cálculo de aproximadamente 10 a 30 vezes menor. Deve-se também levar em conta o fato de que o modelo de Fritchman usado para comparação possui oito parâmetros, contra os sete do modelo proposto.

5.4. Canal com desvanecimento Rayleigh (Caso II)

Neste caso foram utilizadas amostras obtidas de três simulações em computador de um canal variante no tempo com desvanecimento Rayleigh e espalhamento Doppler modelado pelo espectro de Jakes. Os valores do desvio Doppler máximo normalizado (pela taxa de símbolos) considerados nestas três simulações foram de 10^{-1} , 10^{-2} e 10^{-3} , e os valores correspondentes da razão E_b/N_0 foram 10 dB, 15 dB e 15 dB. A modulação utilizada foi a PSK-4 com demodulação coerente, sendo suposto um perfeito sincronismo de fase no receptor. O tamanho das sequências de dados (amostras de erros) gerados nas simulações foi de 2 Mbits.

5.4.1. Canal com desvio Doppler normalizado de 10^{-1} e E_b/N_0 de 10 dB

Para o caso em que o desvio Doppler normalizado foi de 10^{-1} , com E_b/N_0 igual a 10 dB, o limiar *interburst* para este caso foi fixado em 5. Este caso será referido daqui por diante como caso II.a.

Diferentemente do caso I.a Figura 5.9 mostra que para este caso houve uma grande diferença entre os ajustes obtidos com os dois modelos, para a estatística distribuição de surtos de erro, sendo observado que o modelo proposto apresentou uma significativa vantagem na captura desta estatística em relação ao modelo de Fritchman.

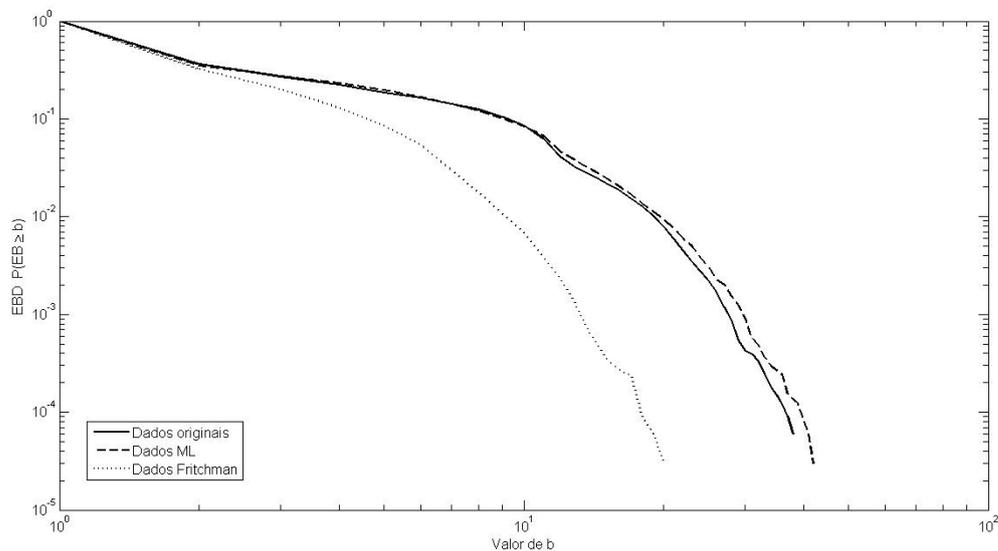


Figura 5.9 – Estimativas distribuição de surtos de erro para o Caso II.a.

A Figura 5.10 mostra que para a distribuição de gaps, as estimativas geradas pelo modelo proposto apresentam um ajuste muito bom àquelas obtidas dos dados originais, superando o modelo de Fritchman, principalmente para valores de comprimentos de gaps acima de 10.

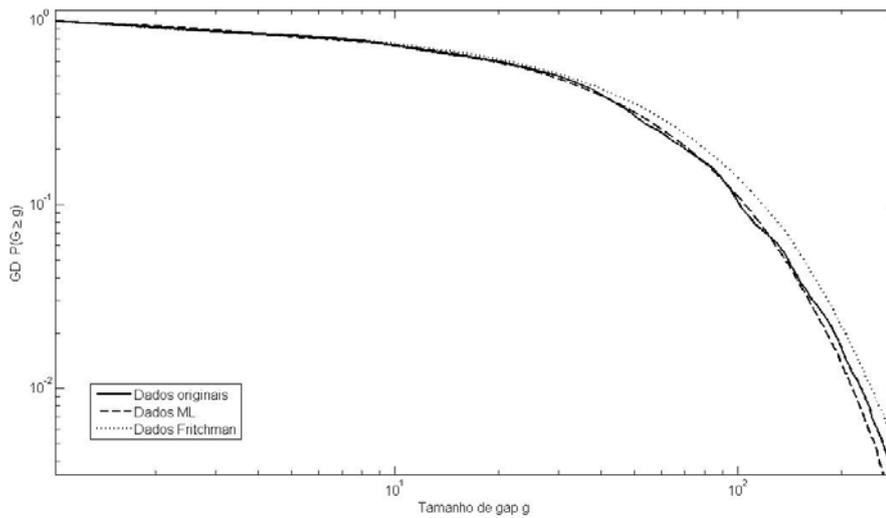


Figura 5.10 – Estimativas da distribuição de gaps para o Caso II.a.

A Figura 5.11 mostra a estimação de comprimentos de intervalos entre surtos para o Caso II.a. Nela se observa mais uma vez que o modelo proposto possibilitou um melhor ajuste aos dados originais do que o modelo de comparação.

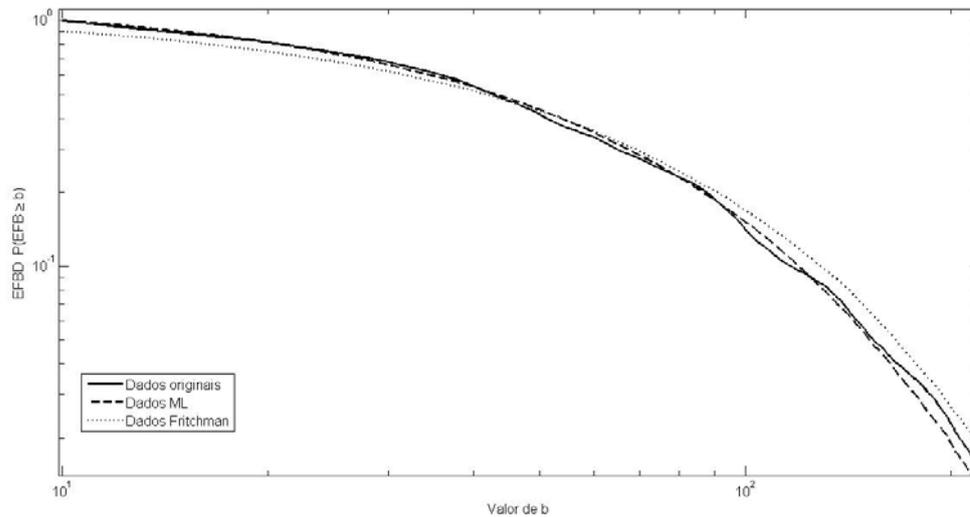


Figura 5.11 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso II.a.

A Figura 5.12 apresenta a função autocorrelação (BCEF) estimada para o Caso II.a. Apesar de neste caso, os resultados produzidos por ambos os modelos terem deixado a desejar, ainda se observa que o modelo proposto produziu nesta estatística um ajuste um pouco melhor do que o do modelo de comparação.

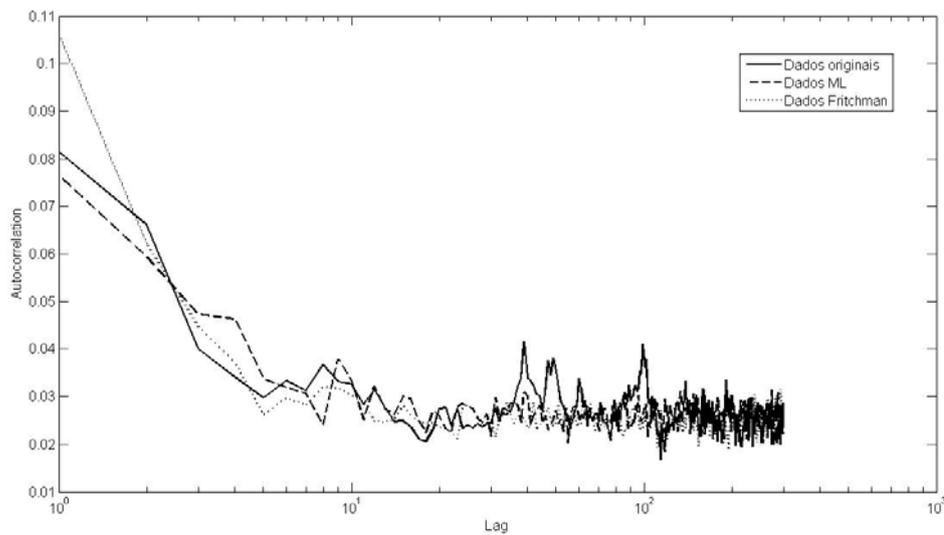


Figura 5.12 – Estimativas da função autocorrelação (BCEF) para o Caso II.a.

A Figura 5.13 apresenta as estimativas da função probabilidade de erros por bloco, $P(m,n)$, para o Caso II.a, para blocos de tamanho 300 bits ($n=300$). O modelo proposto mostrou um desempenho discretamente melhor do que o modelo de Fritchman para capturar esta estatística, com pequena vantagem para valores de m entre 8 e 20. Para os demais valores de m os dois modelos produziram ajustes aproximadamente iguais.

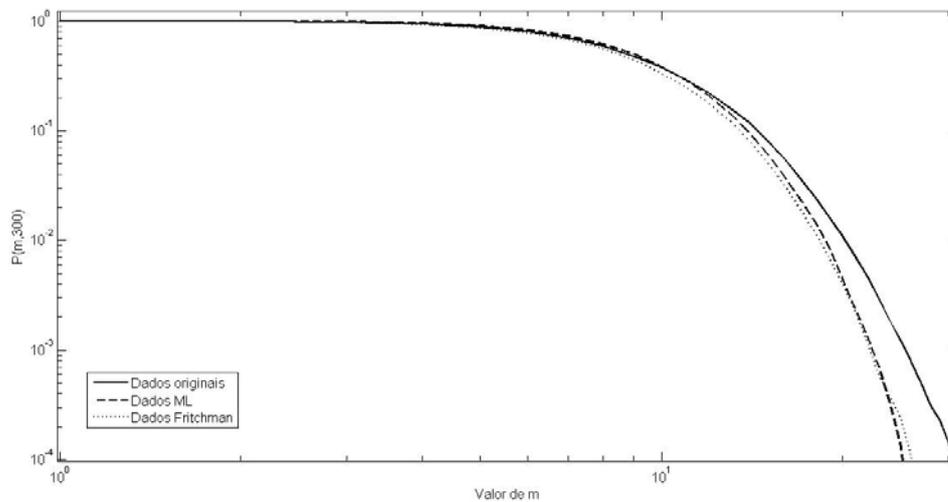


Figura 5.13 – Estimativas da função Probabilidade de erros por bloco de tamanho 300, $P(m,300)$, para o Caso II.a.

5.4.2.

Canal com desvio Doppler normalizado de 10^{-2} e E_b/N_0 de 15 dB

Analogamente à seção anterior, as Figuras 5.14 a 5.18 apresentam as estatísticas de interesse levantadas a partir da sequência de dados obtida com desvio Doppler máximo normalizado de 10^{-2} e razão E_b/N_0 de 15 dB. O limiar *interburst* para este caso foi fixado em 40 para a obtenção destas curvas. Este valor foi baseado na busca do melhor valor de L mostrado no gráfico da Figura 5.1. Este caso será referido como Caso II.b.

A Figura 5.14 indica que também para este caso houve um melhor ajuste do modelo proposto em relação ao modelo de Fritchman, no que diz respeito à distribuição de surtos de erro, para todo os valores de b considerados.

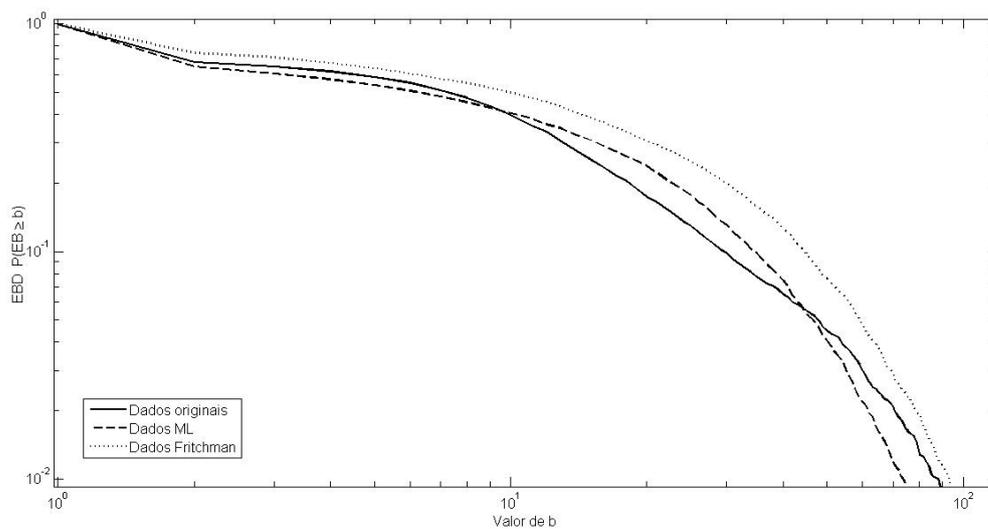


Figura 5.14 – Estimativas da distribuição de surtos de erro para o Caso II.b.

A Figura 5.15 mostra que para a distribuição de gaps obteve-se um ajuste muito bom aos dados originais com o modelo proposto, superando neste aspecto o modelo de comparação, principalmente para valores de comprimentos de gap acima de 6.

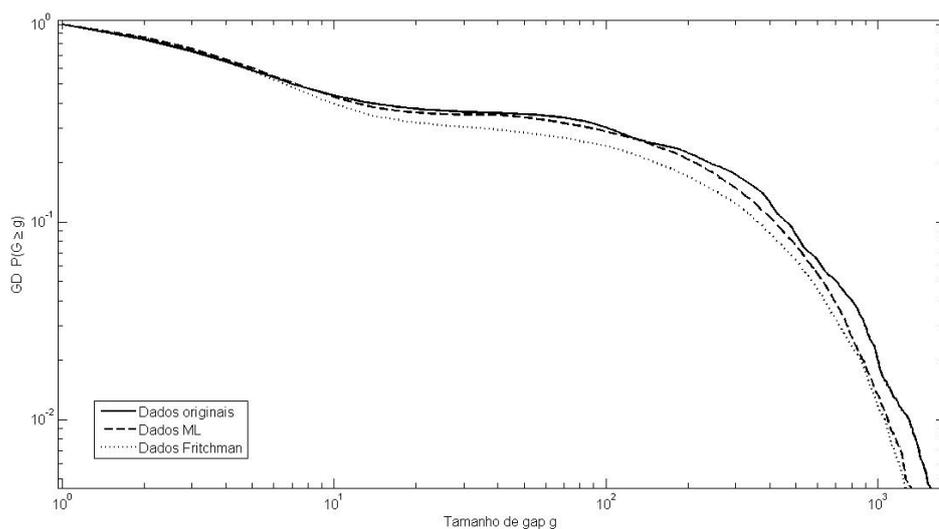


Figura 5.15 – Estimativas da distribuição de gaps para o Caso II.b.

A Figura 5.16 mostra os resultado de estimação da estatística distribuição de comprimentos de intervalos entre surtos para o Caso II.b. Nela pode ser observado o fato interessante em que o modelo proposto apresentou ajuste com precisão aproximadamente idêntica à obtida com o modelo de Fritchman.

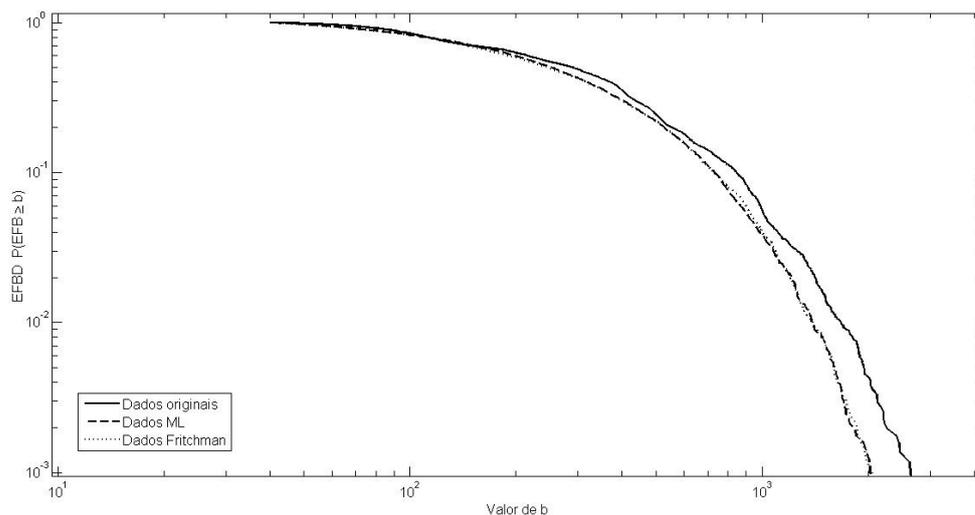


Figura 5.16 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso II.b.

A Figura 5.17 apresenta a função autocorrelação estimada para o Caso II.b. Apesar de neste caso, ambos os modelos terem deixado a desejar, observa-se que o modelo proposto produziu um desempenho um pouco melhor do que o do modelo de comparação no que diz respeito a capturar esta estatística.

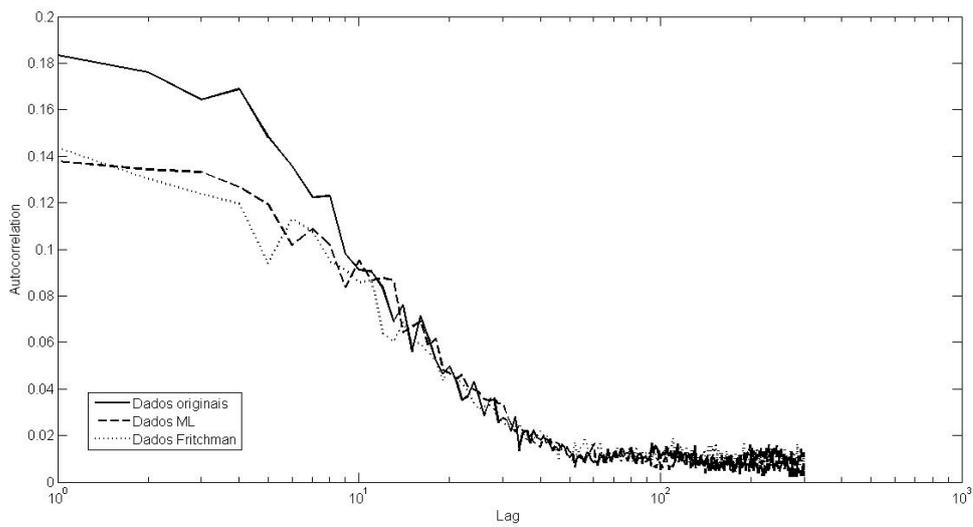


Figura 5.17 – Estimativas da função autocorrelação (BECE) para o Caso II.b.

A Figura 5.18 apresenta os resultados de estimação da função probabilidade de erros por bloco, $P(m,n)$, para o Caso II.b, tendo sido adotado tamanho de bloco igual a 300 ($n=300$). Observa-se que, de maneira geral, o modelo proposto mostrou ajuste desta estatística um pouco melhor do que o modelo de Fritchman.

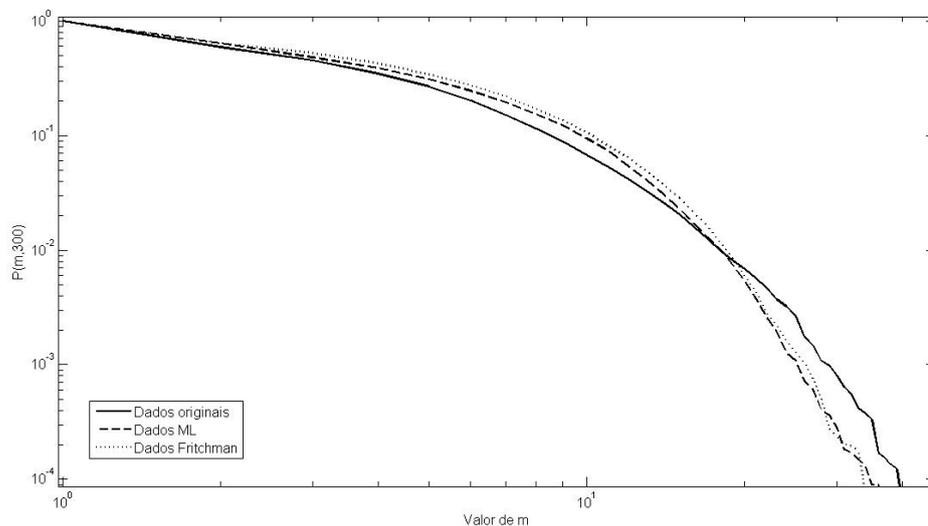


Figura 5.18 – Estimativas da função Probabilidade de erros por bloco de tamanho 300 ($P(m,300)$), para o Caso II.b.

5.4.3. Canal com desvio Doppler máximo de 10^{-3} e E_b/N_0 de 15 dB

Analogamente às duas seções anteriores, as Figuras 5.19 a 5.22 apresentam as estatísticas de interesse levantadas a partir da sequência de dados obtida com desvio Doppler máximo normalizado de 10^{-3} , e razão E_b/N_0 igual de 15 dB. O limiar *interburst* para este caso foi fixado em 504 para a obtenção destas curvas. Este valor foi o que levou a métrica ML ao seu máximo na pesquisa de L, sendo igualado ao limiar *interburst*, como nos outros casos. Este caso será referido como Caso II.c.

A Figura 5.19 demonstra que também para este caso obteve-se muito melhor ajuste com o modelo proposto em relação ao do modelo de Fritchman, para a estatística distribuição de surtos de erro.

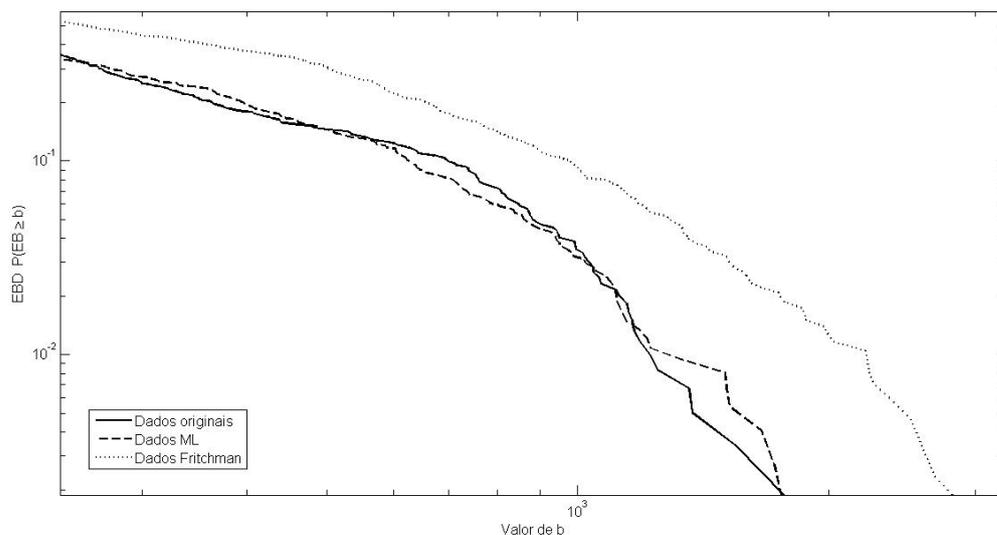


Figura 5.19 – Estimativas da distribuição de surtos de erro para o Caso II.c.

A Figura 5.20 mostra que para a distribuição de gaps, a estimativa gerada pelo modelo proposto apresenta um ajuste muito bom à dos dados originais, superando neste aspecto o modelo de comparação.

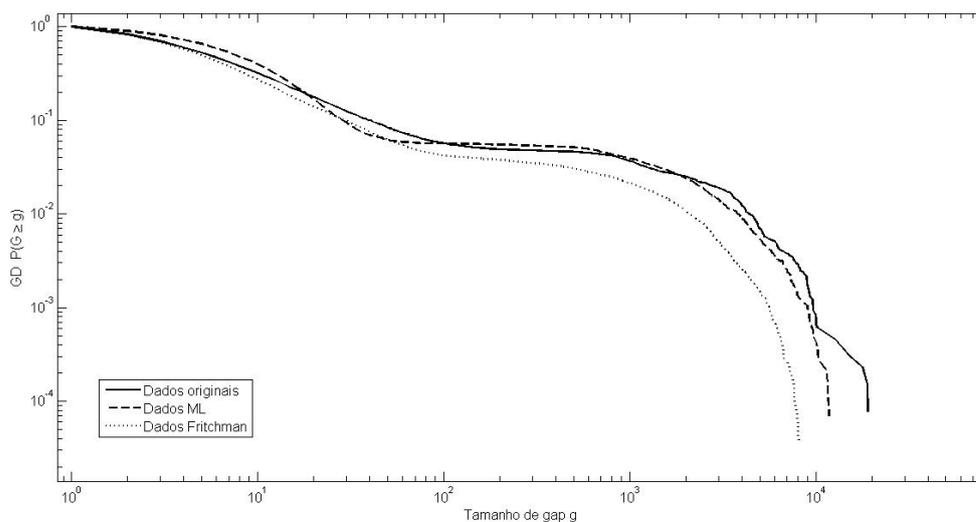


Figura 5.20 – Estimativas da distribuição de gaps para o Caso II.c.

A Figura 5.21 mostra os resultados de estimação da distribuição de comprimentos de intervalos entre surtos para o Caso II.c, e nela se observa que o modelo proposto possibilitou ajuste com precisão bem superior ao do modelo de Fritchman.

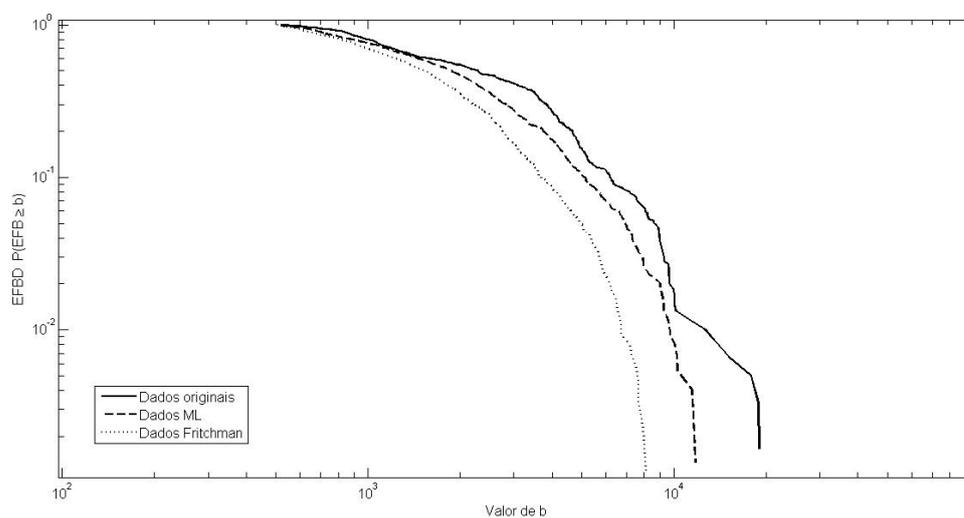


Figura 5.21 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso II.c.

A Figura 5.22 apresenta os resultados de estimação da função probabilidade de erros por bloco, $P(m,n)$, para o Caso II.c, tendo sido adotado tamanho bloco 300 ($n=300$). Observa-se que de uma maneira geral o modelo proposto possibilitou um ajuste desta estatística bem melhor do que o modelo de Fritchman.

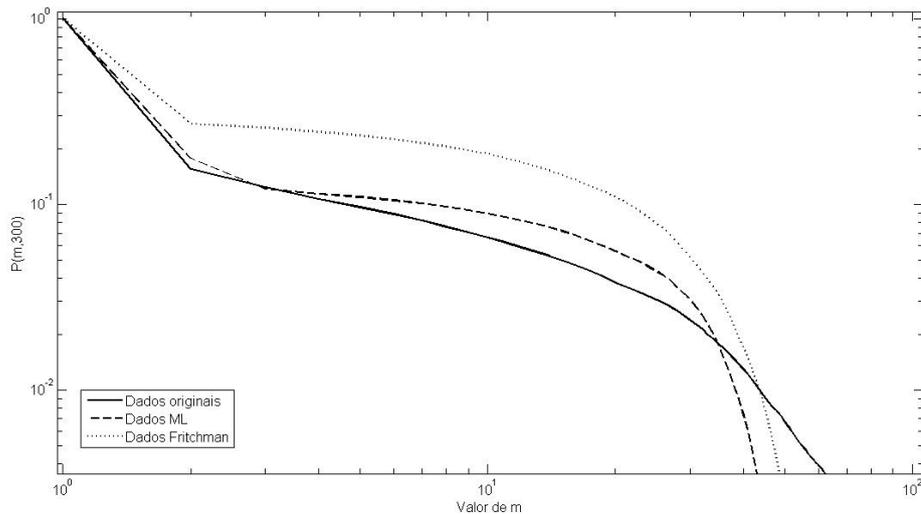


Figura 5.22 – Estimativas da função Probabilidade de erros por bloco de tamanho 300 ($P(m,300)$), para o Caso II.c.

Analisando o conjunto de estatísticas estimadas para o Caso II.c observa-se que, apesar de o modelo proposto possibilitar uma melhor ajuste em relação ao modelo de comparação, de uma maneira geral não as estimativas obtidas não foram muito precisas.

5.5. Decodificador de Viterbi (Caso III)

Foram realizadas duas simulações em computador da decodificação de Viterbi com decisão abrupta aplicada a um código convolucional de taxa $\frac{1}{2}$, comprimento de restrição (“constraint length”) 7, e polinômios geradores (133-171), sendo a probabilidade de erro na sua entrada de 0,1 e 0,05 respectivamente. O tamanho das sequências de dados foi de 2 Mbits.

5.5.1. Decodificador com probabilidade de erro de entrada de 0,1

As Figuras 5.23 a 5.27 mostram os resultados obtidos para o caso em que a probabilidade de erro na entrada do decodificador de Viterbi é de 0,1. O limiar *interburst* para este caso foi fixado em com base na busca do melhor valor de L mostrado no gráfico da Figura 5.2. Este caso será referido como Caso III.a.

A Figura 5.23 mostra que houve uma diferença significativa entre os ajustes dos dois modelos para a estatística distribuição de comprimentos de surtos de erro. Observa-se também que o modelo proposto apresentou uma sensível vantagem na captura desta estatística em relação ao modelo de Fritchman.

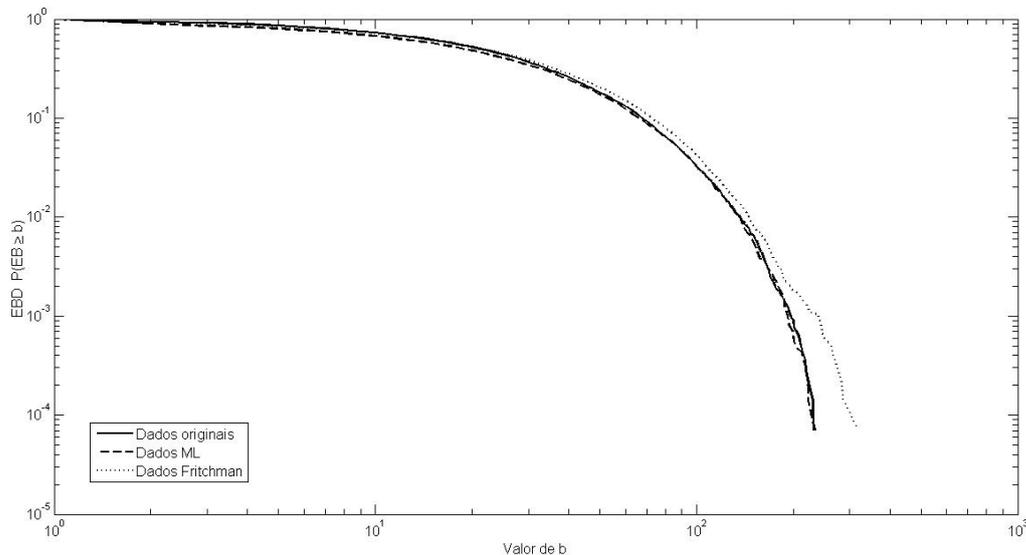


Figura 5.23 – Estimativas da distribuição de comprimentos de surtos de erro para o Caso III.a.

A Figura 5.24 mostra que para a distribuição de gaps as estimativas geradas pelo modelo proposto mostram um excelente ajuste às obtidas com os dados originais. Observamos ainda que o modelo de comparação também possibilitou grande precisão na captura desta estatística.

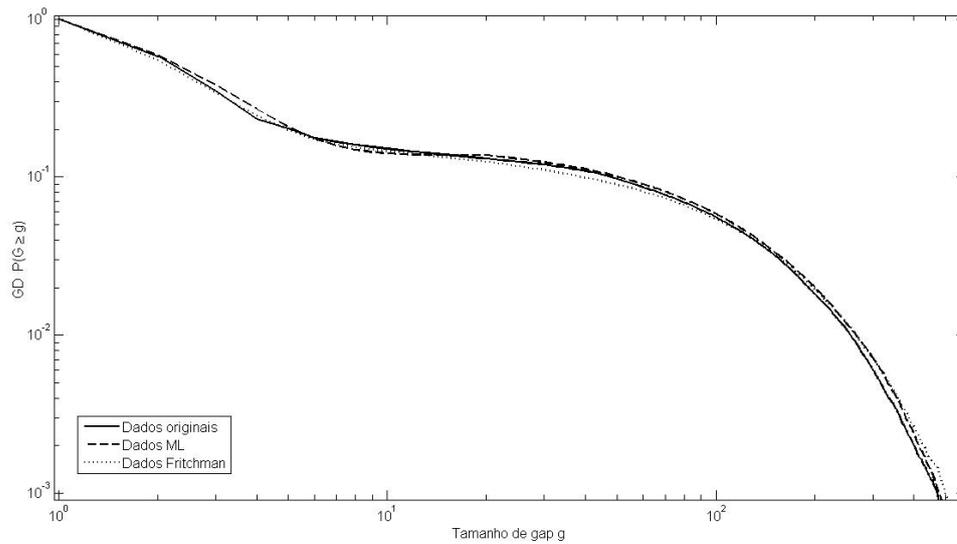


Figura 5.24 – Estimativas da distribuição de gaps para o Caso III.a.

A Figura 5.25 mostra os resultados de estimação da estatística distribuição de comprimentos de intervalos entre surtos para o Caso III.a. Observa-se mais uma vez que o modelo proposto possibilitou um melhor ajuste aos dados originais do que o modelo de comparação.

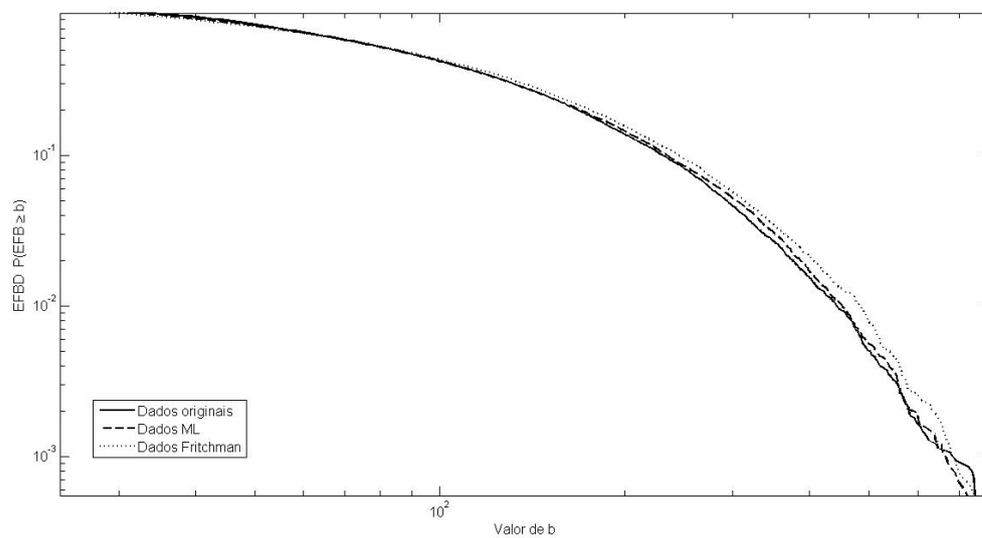


Figura 5.25 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso III.a.

A Figura 5.26 apresenta a função autocorrelação (BCEF) estimada para o Caso III.a. Observa-se que o modelo proposto produziu um bom ajuste na captura desta estatística, superando neste aspecto o modelo de comparação,

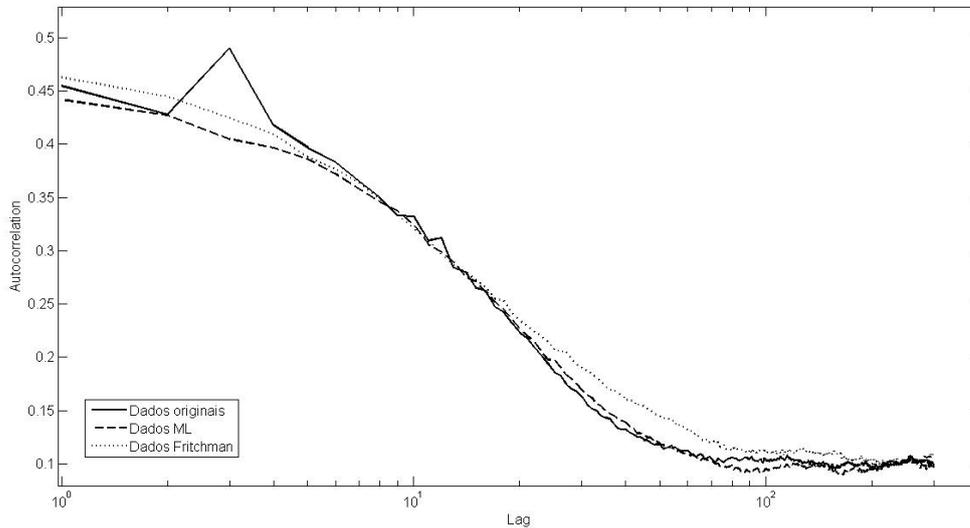


Figura 5.26 – Estimativas da função autocorrelação (BCEF) para o Caso III.a.

A Figura 5.27 representa a estimação da função probabilidade de erros por bloco, $P(m,n)$, para o Caso III.a com tamanho de bloco 300 ($n=300$). Tanto o modelo proposto quanto o modelo de Fritchman apresentaram resultados excelentes na captura desta estatística, ao ponto de praticamente não se distinguir uma curva da outra na Figura 5.27.

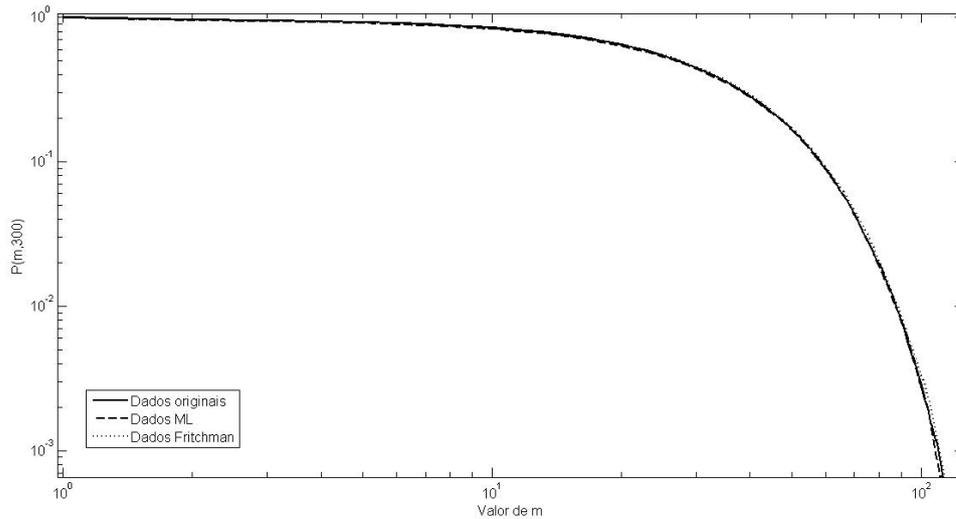


Figura 5.27 – Estimativas da função Probabilidade de erros por bloco de tamanho 300 ($P(m,300)$), para o Caso III.a.

5.5.2. Decodificador com probabilidade de erro de entrada de 0,05

As Figuras 5.28 a 5.32 mostram os resultados obtidos com o decodificador de Viterbi quando a probabilidade de erro de entrada foi de 0,05. O limiar *interburst* para este caso foi fixado em 30 com base na busca do melhor valor de L usando a Figura 5.2. Este caso será referido como Caso III.a.

A Figura 5.28 mostra que houve uma diferença significativa entre os ajustes produzidos pelos dois modelos para a estatística distribuição de comprimentos de surtos de erro, sendo observado que o modelo proposto, de uma maneira geral, viabilizou uma melhor captura desta estatística do que o modelo de Fritchman.

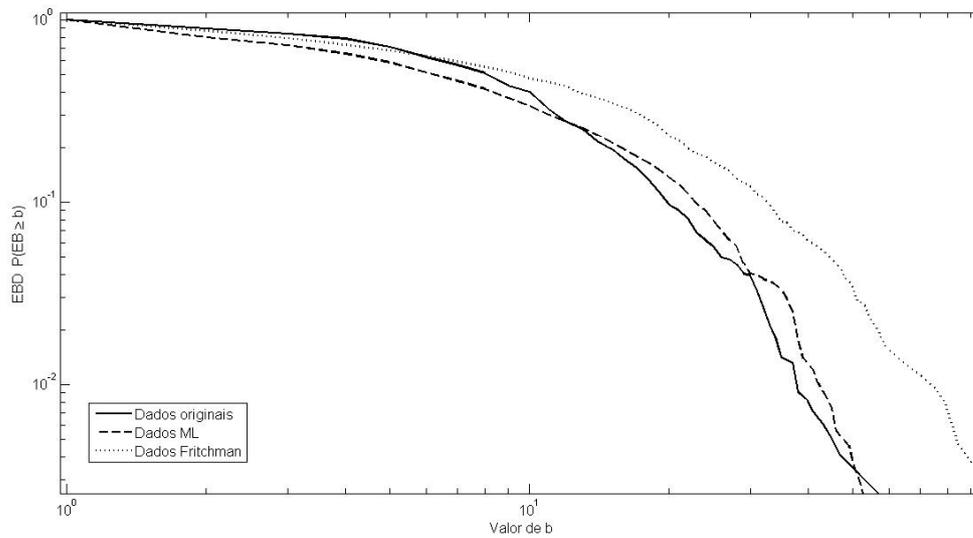


Figura 5.28 – Estimativas da distribuição de comprimentos de surtos de erro para o Caso III.b.

A Figura 5.29 mostra que para a distribuição de gaps a estimativas gerada pelo modelo proposto mostrou um ajuste muito bom à dos dados originais, superando o modelo de comparação, que só obteve ajuste satisfatório para pequenos valores de g .

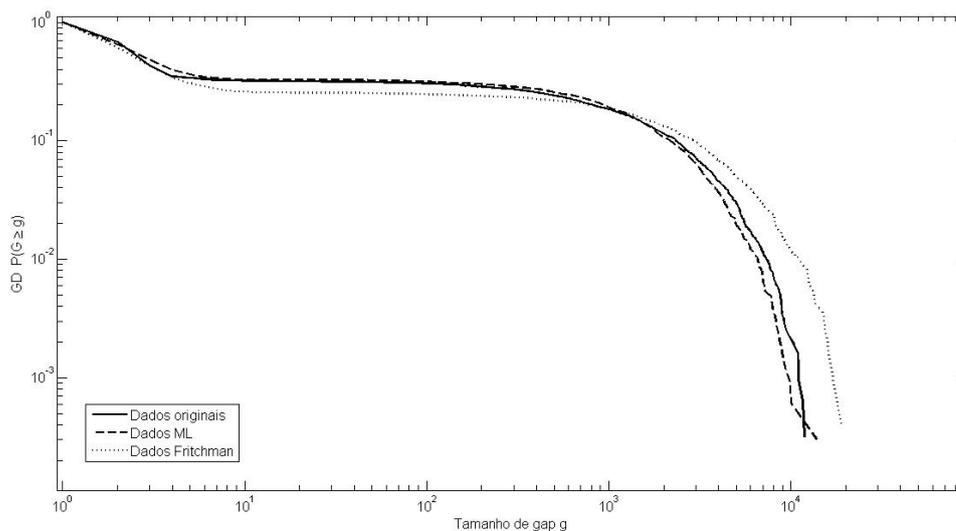


Figura 5.29 – Estimativas da distribuição de gaps para o Caso III.b.

A Figura 5.30 mostra os resultado de estimação da estatística da distribuição de comprimentos de intervalos entre surtos para o Caso III.b. Observa-se nesta figura que mais uma vez o modelo proposto possibilitou ajuste com precisão melhor do que o modelo de Fritchman.

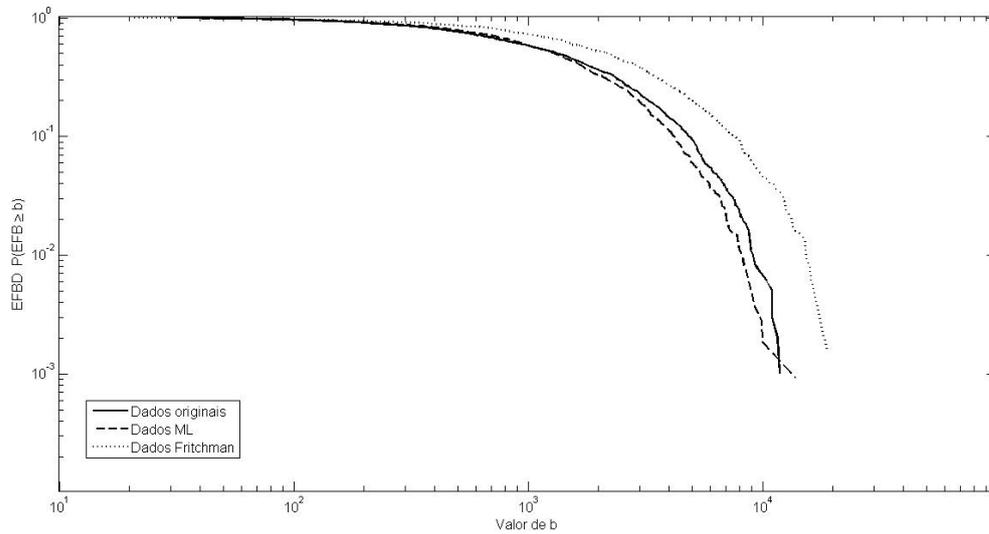


Figura 5.30 – Estimativas da distribuição de comprimentos de intervalos entre surtos para o Caso III.b.

A Figura 5.31 apresenta a função autocorrelação estimada para o Caso III.b. Observa-se que o modelo proposto teve um comportamento um pouco melhor do que o do modelo de comparação, no que diz respeito a capturar esta estatística.

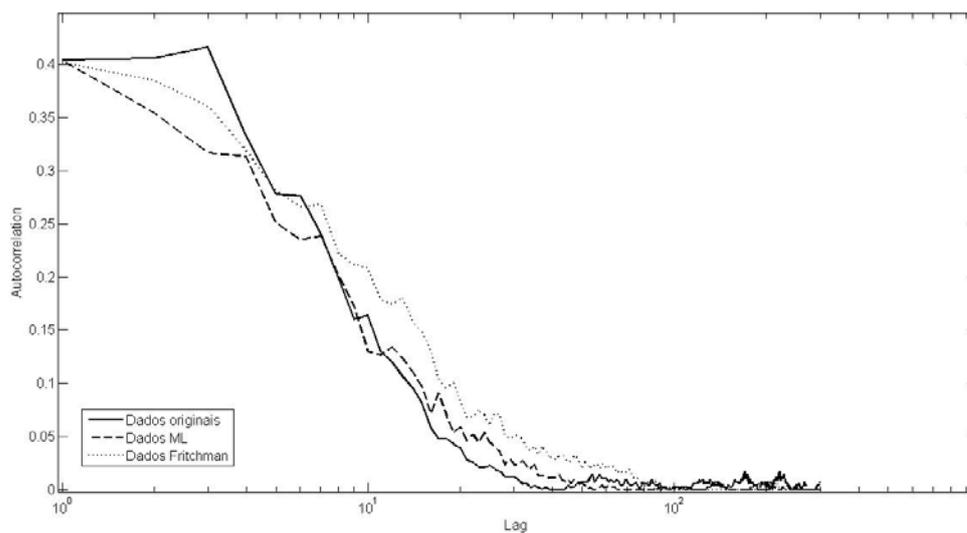


Figura 5.31 – Estimativas da função autocorrelação (BECE) para o Caso III.b.

A Figura 5.32 representa a estimação da probabilidade de erros por bloco $P(m,n)$ para o Caso III.b, com tamanho de bloco de tamanho 300 bits ($n=300$). De maneira geral o modelo proposto produziu um ajuste desta estatística bem melhor do que o obtido com o modelo de Fritchman.

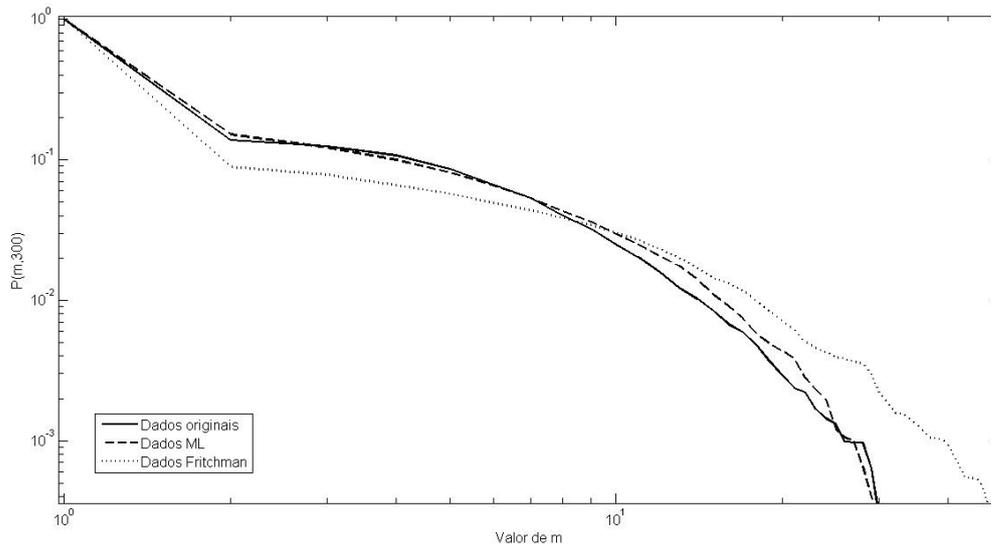


Figura 5.32 – Estimativas da função Probabilidade de erros por bloco de tamanho 300, $P(m,300)$, para o Caso III.b.

Observando-se os resultados apresentados pelas Figuras 5.4 a 5.32, conclui-se parcialmente, que na grande maioria dos casos o modelo proposto produz melhores ajustes do que os obtidos com o modelo de Fritchman. Numa pequena fração dos casos avaliados os dois modelos apresentam precisão aproximadamente igual.

5.6. Influência do emprego da técnica PSO

Como já mencionado anteriormente, a técnica PSO foi usada para gerar as estimativas iniciais dos parâmetros do modelo proposto, a serem utilizadas no algoritmo de otimização clássica.

Para ilustrar, de uma forma qualitativa, o efeito do emprego desta técnica, as duas Figuras seguintes apresentam os resultados de estimações feitas com uma sequência de erro de 2 Mbits, originada da simulação do decodificador de Viterbi acima mencionado, com probabilidade de erro de bit em sua entrada igual a 0,1.

A Figura 5.33 mostra as estimativas obtidas com inicialização arbitrária do algoritmo de otimização. Neste caso observa-se claramente que a convergência não aconteceu ou se deu para um ponto de máximo local, uma vez que foram obtidos péssimos ajustes para as seis estatísticas escolhidas, a saber: distribuição de gaps, distribuição de surtos de erro, função probabilidade de erros por bloco, distribuição de $P(0^m|1)$, distribuição de comprimentos de intervalos entre surtos e função autocorrelação.

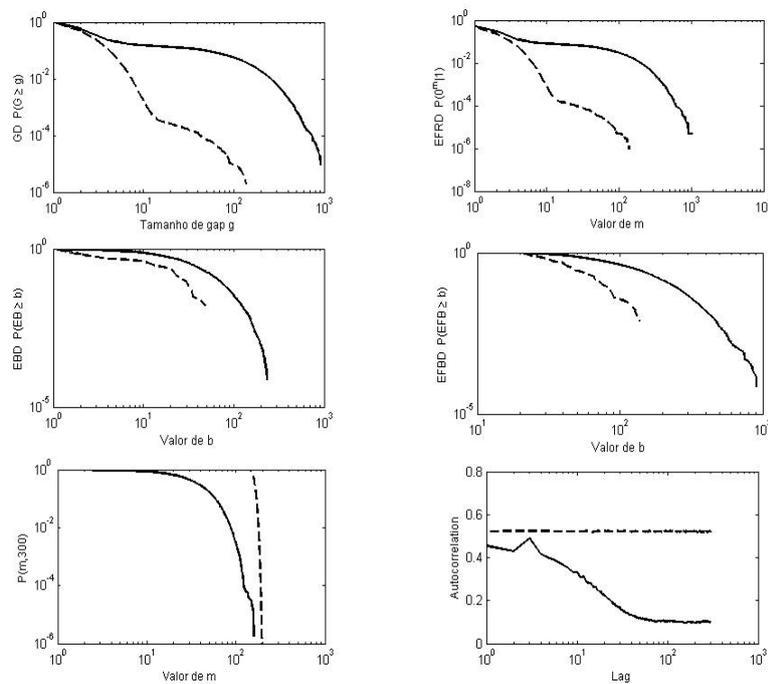


Figura 5.33 – Estimativas obtidas sem o emprego do PSO no ajuste do modelo proposto.

A Figura 5.34 apresenta resultados obtidos para o mesmo experimento de simulação (com a mesma amostra de erros), com a diferença de que a inicialização dos parâmetros foi feita empregando-se o PSO. Observam-se agora excelentes ajustes, o que comprova a eficácia na estimação dos parâmetros, conseguida com o uso combinado do PSO e da otimização simples.

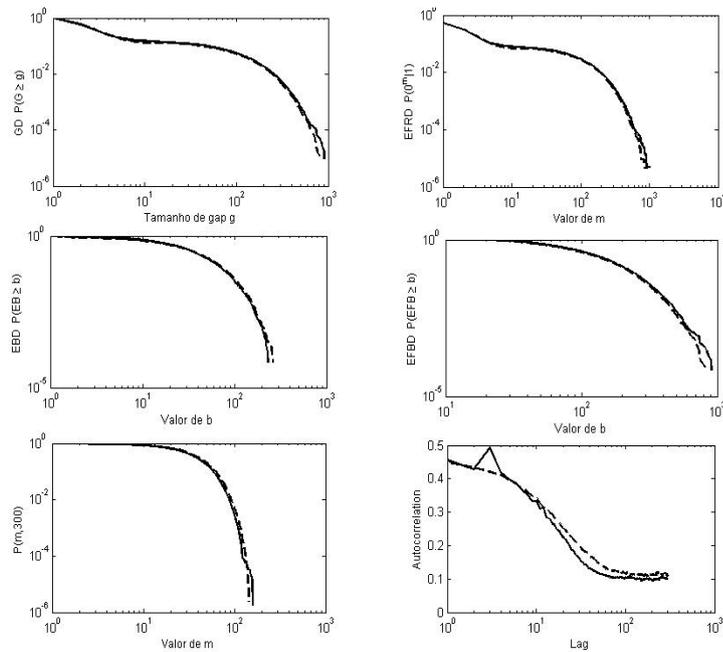


Figura 5.34 – Estimativas obtidas com o emprego do PSO no ajuste do modelo proposto.

Em suma, as Figuras 5.33 e 5.34 mostram a importância de uma escolha adequada para os valores iniciais dos parâmetros a serem usados pelo algoritmo de otimização clássica e ainda comprovam a eficácia da técnica PSO como solução para este tipo de problema. Cabe lembrar ainda que este efeito foi obtido de forma eficiente, uma vez que o número de partículas empregado na técnica PSO foi pequeno ao ponto de fazer com que todo o processo de otimização ainda pudesse ser feito com tempo de processamento de 10 a 30 vezes menor do que o gasto para ajuste de parâmetros do modelo de Fritchman usado como referência.

6 Conclusão

A crescente evolução tecnológica dos sistemas de comunicações digitais tem sido em grande parte direcionada às comunicações sem fio, que por suas características, originam os denominados canais com memória. Este problema, por sua importância, motiva o estudo e análise dos erros em surtos; o impacto gerado pelos mesmos nos sistemas de comunicações e o desenvolvimento de modelos capazes de representar satisfatoriamente o comportamento estatístico destes canais.

A literatura atual apresenta diversos trabalhos de pesquisa e desenvolvimento de modelos capazes de representar o comportamento estatístico de canais com memória. Porém observa-se que de um modo geral cada modelo ou família de modelos é direcionada a um determinado conjunto de casos como se pode verificar em [2], [4], [29], [31], que atendem satisfatoriamente a alguns casos, porém, não sendo aplicável a outros, os quais necessitam outros modelos mais adequados para representá-los.

Foram revistos os modelos dos tipos Gilbert-Elliott e Fritchman, amplamente utilizados na modelagem de canais sujeitos a erros em surtos. O modelo de Gilbert-Elliott, sendo um modelo de estrutura simplificada, com quatro parâmetros, possibilita menor complexidade no seu tratamento analítico do que os modelos de maior número de estados e parâmetros, porém, sendo uma simples aproximação torna-se inadequada a sua aplicação para alguns casos de interesse.

Os modelos de Fritchman são freqüentemente utilizados para modelagem de alguns tipos de canais com memória não atendidos pelo de Gilbert-Elliott [4], porém ainda é restrito a um número limitado de casos. O número de parâmetros deste modelo depende do seu número de estados, sendo que o aumento de estados para compensar uma deficiência em algum ajuste pode levar a um número de parâmetros que inviabilize o tratamento computacional [2].

No presente trabalho foi apresentada a proposta de um novo modelo HMM, com uma estrutura especial direcionada para a geração de em seqüências de erros em surtos.

Assim sendo, o modelo proposto é especialmente estruturado para reproduzir a geração dos padrões típicos dos processos de erros em surtos. Este modelo tem apenas sete parâmetros a serem ajustados, independentemente do seu número de estados, sendo um deles inteiro não negativo.

O modelo proposto possibilitou a dedução e aplicação de expressões simples para estimação ML desses parâmetros, assim como o emprego do PSO para inicialização da estimação clássica.

Os resultados numéricos de uma comparação com um modelo Fritchman usual mostraram que a proposta de modelo é superior no que diz respeito à reprodução as propriedades estatísticas dos dados empíricos. Uma redução significativa no esforço computacional para a estimação dos parâmetros foi também obtida, permitindo ajustes compatíveis com os de modelos de maior complexidade, porém com um tempo de processamento bem inferior.

Os resultados obtidos do ajuste dos dados simulados em computador de um equalizador DFE, de um canal com desvanecimento Rayleigh e de um decodificador de Viterbi, todos geradores de erros em surtos, mostraram-se promissores, indicando ser este modelo bastante flexível e capaz de capturar estatísticas de interesse de diversas fontes de surtos de erro.

Como propostas para trabalhos futuros, pode-se citar as seguintes:

- a pesquisa de métodos mais eficientes para a estimação do parâmetro inteiro L do modelo proposto;
- a investigação de melhorias na estimação ML dos parâmetros contínuos do modelo proposto;
- uso de critérios alternativos de ajuste pelas estatísticas de interesse.

7

Referências bibliográficas

- 1 CCITT “Blue book”. **Terms and definitions**. Vol. 1, fasc. I.3, rec. M.60 n.34, rec. Q.9 n.0222.
- 2 CHENG, XIANG WANG e PATZOLD MATTHIAS. **A new deterministic process based generative model for characterizing bursty error sequences**. Faculty of Engineering and Science, Agder University College, Grimstad Norway. IEEE 2004
- 3 CHOCKALINGAM A., ZORZI, MICHELE e RAO, RAMESH R. **Performance of TCP on wireless fading links with memory**. ICC’ 98. Atlanta, GA: June 7-11, 1998.
- 4 CHU, V. Y. Y. e SWEENEY, P. **Channel modelling and error control strategies for the LEO satellite channel**. Dept. of Electrical Engineering, The University of Surrey. Disponível: [http://www.ee.surrey.ac.uk/CCSR/Software/OPNET /chu-sweeney.pdf](http://www.ee.surrey.ac.uk/CCSR/Software/OPNET/chu-sweeney.pdf) [capturado em 28 Set 2009].
- 5 FRANCHI, A., HARRIS, R. A. **On the error burst properties of Viterbi decoding**. IEEE International Communications Conference, 1993. pag. 1086 a 1091.
- 6 FRANCHI, A., HARRIS, ROBERT A. **On the error burst properties of the “standard” K=7, rate-1/2 convolutional code with soft-decision viterbi decoding**. Telecommunication Systems. May 1995. Vol. 6, nº 3, pag 337 a 351.
- 7 FRITCHMAN, B. D. **A binary channel characterization using partitioned Markov chains**. IEEE Transactions in Information Theory, April 1967. Vol. IT-13, pp. 221 – 227.
- 8 HEISLER, JEFFREY R., BARSOUM, YOSRY A. e Richard Condello. **An analysis of the Viterbi decoder error statistics for ATM and TCP/IP over satellite communication**. IEEE, 1999.
- 9 JOHNSON, ERIC E. **HF SchEMe: a skywave channel error model**. IEEE Military Communications Conference (MILCOM), 1994.

- 10 KOPKE, ANDREAS, WILLIG ANDREAS e KARL, HOLGER. **Chaotic maps as parsimonious bit error models of wireless channels.** Telecommunication Networks Group, Technische Universitat Berlin. IEEE, 2003.
- 11 MELO, ALEXANDRE GUEDES de. **Equalização cega em canais WSS-US.** Dissertação de Mestrado (Mestrado em Engenharia Elétrica) - IME, jul. 2000.
- 12 PAPOULIS, ATHANASIOS. **Probability, random variables, and stochastic processes.** 3. ed. Editora McGraw-Hill International Editions, 1991.
- 13 PROAKIS, J. G. **DIGITAL COMMUNICATIONS.** Mc Graw Hill, New York: 1995.
- 14 RABINER, L. R. e JUANG, B. H. **Fundamentals of speech recognition.** Editora Prentice Hall, 1993.
- 15 RAMSEIER, STEFAN e KALTENSCHNEE, THOMAS. **ATM over satellite: analysis of ATM QOS parameters.** Ascom Tech Ltd. Gewerbepark, CH-5506 Magenwill, Switzerland: 1995. pag. 1562 a 1565.
- 16 SHEN, JIA-PEI e GILL, JOHN. **Analysis on a hidden Markov channel model.** Global Telecommunications Conference, Globecom' 99, 1999.
- 17 TURIN, WILLIAN e NOBELEN, ROBERT VAN. **Hidden Markov modeling of flat fading channels.** IEEE Journal on selected areas in communications. dez. 1998. v. 16, n. 9.
- 18 TURIN, WILLIAN. **Digital Transmission Systems – Performance Analysis and Modeling.** Editora Mc Graw-Hill Telecommunications,
- 19 ZORZI, MICHELE, RAO, RAMESH R. e MILSTEIN, LAURENCE B. **On the accuracy of a first-order Markov model for data transmission on fading channels.** ICUPC'95. Tokyo, Japão: nov. 1995.
- 20 ZORZI, MICHELE, RAO, RAMESH R. e MILSTEIN, LAURENCE B. **A Markov model for block errors on fading channels.** PIMRC'96. Taiwan: out.1996.
- 21 ZORZI, MICHELE E RAO, RAMESH R. **ARQ error control for delay-constrained communications on short-range burst-error channels.** VTC' 97. Phoenix, AZ: mai. 1997.

- 22 ZORZI, MICHELE E RAO, RAMESH R. **The effect of correlated errors on the performance of TCP**. IEEE Communications Letters, vol. 1, no. 5, pp. 127-129, September, 1997.
- 23 ZORZI, MICHELE. **Outage and error events in bursty channels**. IEEE Transactions on Communications. mar. 1998. Vol. 46, n^o 3, pag. 349 a 356.
- 24 ZORZI, MICHELE. **Performance of FEC and ARQ error control in bursty channels under delay constraints**. VTC'98. Ottawa, Canadá: mai. 1998.
- 25 ZORZI, MICHELE E RAO, RAMESH R. **Impact of burst errors on framing**. PIMRC' 98. Boston: set. 1998.
- 26 ZORZI, MICHELE E RAO, RAMESH R. **Perspectives on the impact of error statistics on protocols for wireless networks**. IEEE Personal Communications Magazine. Oct. 1999.
- 27 ZORZI, MICHELE, CHOCKALINGAM, A. E RAO, RAMESH R. **Throughput analysis of TCP on channels with memory**. IEEE Journal on Selected Areas in Communications. Jul. 2000. Vol. 18 n^o 7, pag. 1289 a 1300.
- 28 ZORZI, MICHELE E RAO, RAMESH R. **On the statistics of block errors in bursty channels**. IEEE Transactions on Communications vol. 45, pp. 660–667, Jun. 1997.
- 29 WANG, CHENG-XIANG e PATZOLD, MATTHIAS. **A generative deterministic model for digital mobile fading channels**. IEEE Communications Letters, Vol. 8, No. 4, April 2004.
- 30 LASKARI, E. C., PARSOPOULOS, K. E. e VRAHATIS, M. N. **Particle Swarm Optimization for Integer Programming**. IEEE Computer Society, Proceedings of the Evolutionary Computation on 2002. CEC '02. Proceedings of the 2002 Congress - Vol 02, Pag 1582-1587.
- 31 WANG, C. X. e XU, W. **A new class of generative models for burst error characterization in digital wireless channels**. IEEE Transactions on Communications, 55(3): 453-462, 2007.
- 32 Wang, C. X. e Patzold, M. **Deterministic modeling and simulation of error sequences in digital mobile fading channels**. IEEE International Conference on Communications, Paris, Jun 2004.

Apêndice A

Notação matricial

A fim de compactar a as expressões para cálculo recursivo da função de verossimilhança, adotaremos uma notação matricial. Define-se inicialmente uma matriz de probabilidades de transição especial, cujos elementos são dados por:

$$P_{ij}(m_t, n_t) = P\left[S_t = e_i; (1^{m_t} 0^{n_t}) \mid S_{t-1} = e_j\right], \quad e_i, e_j \in \{e_0, e_3\} \quad (\text{A.1})$$

Calculando-se os elementos, em função do número de uns m , e dos parâmetros $\alpha, \beta, \delta, \lambda, \mu, \rho$ e L do modelo proposto, para todas as combinações possíveis de i e j , que envolvem os dois estados de erro e_0 e e_3 , obtem-se as seguintes expressões:

$$P_{ij}(m, n) = \begin{cases} \mu \cdot \delta^{n-1} \cdot \lambda \cdot (1-\mu)^{m-1} & i=0, j=0 \\ \mu \cdot \delta^{n-1} \cdot (1-\delta-\lambda) \cdot \alpha^{m-1} & i=3, j=0 \\ (1-\alpha-\beta) \cdot \delta^{n-L} \cdot \lambda \cdot (1-\mu)^{m-1} & i=0, j=3, n \geq L \\ 0 & i=0, j=3, n < L \\ (1-\alpha-\beta) \cdot \delta^{n-L} \cdot (1-\delta-\lambda) \cdot \alpha^{m-1} & i=3, j=3, n \geq L \\ \beta \cdot \rho^{n-1} \cdot \alpha^{m-1} & i=3, j=3, n = L-1 \\ \beta \cdot \rho^{n-1} \cdot (1-\rho) \cdot \alpha^{m-1} & i=3, j=3, n < L-1 \end{cases} \quad (\text{A.2})$$

para a partida do processo de cálculos recursivos, necessita-se calcular o valor inicial de $a_t(i)$, como se segue:

$$a_0(i) = P[S_0 = e_i; (1^{m_0} 0^{n_0}) | 1] =$$

$$\left(\frac{P(S_{-1} = e_0)}{P(S_{-1} = e_0) + P(S_{-1} = e_3)} \right) P_{i_0}(m_0, n_0) + \left(\frac{P(S_{-1} = e_3)}{P(S_{-1} = e_0) + P(S_{-1} = e_3)} \right) P_{i_3}(m_0, n_0) =$$

$$\left(\frac{x_0}{x_0 + x_3} \right) P_{i_0}(m_0, n_0) + \left(\frac{x_3}{x_0 + x_3} \right) P_{i_3}(m_0, n_0) \quad (\text{A.3})$$

Logo:

$$a_0(i) = \left(\frac{x_0}{x_0 + x_3} \right) P_{i_0}(m_0, n_0) + \left(1 - \frac{x_0}{x_0 + x_3} \right) P_{i_3}(m_0, n_0) \quad (\text{A.4})$$

$$a_t(i) = \sum_{j \in \{0,3\}} P_{ij}(m_t, n_t) \cdot a_{t-1}(j) \quad (\text{A.5})$$

Como:

$$\bar{P} = \frac{x_0}{x_0 + x_3} = \frac{1}{1 + \frac{x_3}{x_0}} = \frac{1}{1 + \left(\frac{\mu}{\lambda} \right) \cdot \left(\frac{1 - \delta - \lambda}{1 - \alpha - \beta} \right)} \quad (\text{A.6})$$

Então,

$$\bar{P} = \frac{1}{1 + p} \quad (\text{A.7})$$

Define-se:

$$\bullet P(m, n) = \begin{bmatrix} P_{00}(m, n) & P_{03}(m, n) \\ P_{30}(m, n) & P_{33}(m, n) \end{bmatrix} \quad (\text{A.8})$$

$$\bullet \underline{a}_{-1} = \begin{bmatrix} \bar{P} \\ 1 - \bar{P} \end{bmatrix} \quad (\text{A.9})$$

$$\bullet \underline{a}_t = \begin{bmatrix} a_t(0) \\ a_t(3) \end{bmatrix} \text{ para } t \in \{0, 1, 2, \dots, K\} \quad (\text{A.10})$$

tem-se: $\underline{a}_t = P(m_t, n_t) \cdot \underline{a}_{t-1}$

Assim, finalmente obtemos a expressão da função verossimilhança na forma matricial:

$$V(\underline{m}_K, \underline{n}_K) = [1 \quad 1] \cdot \left\{ \prod_{k=0}^K P(m_k, n_k) \right\} \cdot \begin{bmatrix} \bar{P} \\ 1 - \bar{P} \end{bmatrix} \quad (\text{A.11})$$

Apêndice B

Técnicas de otimização

As técnicas de otimização são usadas para se encontrar um conjunto de parâmetros $x = \{x_1, x_2, \dots, x_n\}$, que pode em algum sentido, ser definido como ótimo. O seu caso mais simples visa à minimização ou maximização de determinada função dependente de x , normalmente chamada de função objetivo. Em uma abordagem mais avançada, a função objetivo $f(x)$ a ser minimizada ou maximizada, está sujeita a restrições das seguintes formas:

- Restrições de igualdade: $G_i(x) = 0, (i = 1, \dots, m_e);$
- Restrições de desigualdade: $G_i(x) \leq 0, (i = m_e + 1, \dots, m);$
- Limites dos parâmetros: $x_l, x_u.$

Um problema geral pode ser descrito da seguinte forma:

$$\underset{x}{\text{Min}} f(x) \tag{B.1}$$

Sujeito às restrições:

$$\begin{aligned} G_i(x) &= 0, (i = 1, \dots, m_e) \\ G_i(x) &\leq 0, (i = m_e + 1, \dots, m) \\ x_l &\leq x \leq x_u \end{aligned} \tag{B.2}$$

Onde x é um vetor cujos elementos são os n parâmetros de interesse, limitados pelos respectivos elementos dos vetores x_l e x_u ; $f(x)$ é a função objetivo, que resulta um valor escalar, e as funções vetoriais $G_i(x)$ resultam os valores correspondentes às restrições de igualdade e de desigualdade, calculadas em x .

A precisão e eficiência da solução deste tipo de problema dependem não somente do número de parâmetros e de restrições, mas também das características da função objetivo e das restrições, as quais sendo funções lineares dos

parâmetros, o problema é conhecido como de Programação Linear (LP). Quando a minimização ou maximização é realizada em uma função objetivo quadrática sujeita restrições lineares, o problema é conhecido como de Programação Quadrática (QP). Tanto para os problemas de LP como de QP, existem procedimentos que resultam em confiáveis soluções. Porém, a maior dificuldade encontra-se na solução dos problemas conhecidos como de Programação Não-linear (NP) onde a função objetivo e as restrições são funções não lineares dos parâmetros. A solução destes problemas geralmente requer procedimentos iterativos mais complexos.

O problema de otimização empregado pelo modelo proposto faz parte do universo de problemas de otimização não-linear, já que definimos como função objetivo a função de verossimilhança de seis variáveis, com algumas restrições impostas pelas propriedades das probabilidades e pela estrutura do modelo proposto.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)