

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.