

Um Modelo de Markov Oculto para Representação de Perda de Pacotes em Redes 802.11 em Ambientes Indoor

Kleber V. Cardoso¹, José F. de Rezende¹

¹GTA - PEE - COPPE – Universidade Federal do Rio de Janeiro (UFRJ)
Caixa Postal 68.504 – 21.945-970 – Rio de Janeiro – RJ – Brasil

{kleber, rezende}@gta.ufrj.br

Abstract. *In this paper is made a deep analysis of packet loss in IEEE 802.11 networks under indoor environment restrictions and it is discussed statistical models that have been used to represent this sort of process. This work also shows the failure of Gilbert-Elliot model in representing packet loss in 802.11 wireless networks in indoor environment. It is presented a new model based on Hidden Markov Chains with a birth-death structure. This new model describes with good accuracy the mean error rate, variance, length of loss bursts and process autocorrelation.*

Resumo. *Neste artigo é realizada uma análise detalhada de perda de pacotes em redes IEEE 802.11 em ambientes indoor e são discutidos os modelos estatísticos utilizados para a sua representação. Esse trabalho mostra também que o clássico modelo Gilbert-Elliot, apesar de ser largamente usado na representação de perda de pacotes em redes sem fio, é inadequado para descrição das perdas em redes 802.11. É proposto então um novo modelo baseado em uma cadeia de Markov Oculta com uma estrutura do tipo nascimento-e-morte. O novo modelo é simples, porém consegue capturar com acurácia satisfatória métricas importantes como taxa média de perda, variância, comprimento das rajadas de perda e autocorrelação do processo.*

1. Introdução

Redes sem fio vêm se tornando cada vez mais utilizadas e, até o momento, o padrão IEEE 802.11 é o que apresenta o maior número de dispositivos instalados. Entre as razões para a ampla aceitação da tecnologia 802.11 estão a facilidade de implantação, a capacidade para atender a maior parte das aplicações atuais e o baixo custo dos equipamentos.

Em redes cabeadas, as perdas de pacotes são raras e, geralmente, causadas por transbordo de *buffers*. Em redes sem fio, o meio de transmissão é, em geral, a principal causa de perda. No meio sem fio, um pacote pode ser corrompido devido a problemas como interferência, desvanecimento ou multipercursos [Rappaport 2001]. Tipicamente, há mais perda de pacotes em redes sem fio que em redes cabeadas, tornando-se um fenômeno importante que necessita ser compreendido.

Há vários trabalhos sobre perda de pacotes em redes sem fio e alguns especificamente sobre redes 802.11. Na seção a seguir, é apresentado um resumo dos principais trabalhos sobre o assunto. Além de uma análise estatística da perda em redes sem fio, os trabalhos propõem também modelos estocásticos para representar esse processo. No

entanto, como será mostrado posteriormente, os trabalhos sobre perda de pacotes em redes 802.11 ainda prescindem de aspectos importantes na avaliação desse processo e, sobretudo, não apresentam um modelo estatístico amplamente aceito para a representação das mesmas. De fato, alguns trabalhos defendem o uso do modelo de Gilbert-Elliot, o qual se mostrou inadequado conforme resultados iniciais apresentados no artigo curto [Cardoso and de Rezende 2007].

Dentro desse contexto, esse trabalho estende o anterior, apresentando uma avaliação detalhada da perda em redes sem fio 802.11 e tratando aspectos importantes negligenciados em outros trabalhos. São destacadas as métricas que devem ser usadas na avaliação de um modelo e quais modelos podem representar adequadamente a perda de pacotes em redes 802.11, destacando vantagens e desvantagens dos mesmos. Por fim, um modelo Markoviano Oculto com uma estrutura do tipo nascimento-e-morte é proposto e avaliado, mostrando resultados satisfatórios.

O restante desse documento está organizado da seguinte forma. A Seção 2 apresenta alguns trabalhos relevantes sobre perda de pacotes, discutindo suas principais contribuições, assim como aspectos importantes que não foram abordados. Na Seção 3 é mostrada uma classificação dos tipos de modelos estatísticos utilizados para representação de perdas e são comentadas as vantagens e desvantagens dos mesmos. Nessa seção ainda são apresentados os conceitos de estacionariedade e autocorrelação. A Seção 4 descreve o ambiente de testes utilizado, equipamentos, configuração e metodologia para coleta das informações. A Seção 4 traz também uma análise detalhada sobre a perda de pacotes, usando diferentes métricas de avaliação. Na Seção 5, é feita uma avaliação do modelo de Gilbert-Elliot e de uma cadeia de Markov de ordem k . Um novo modelo, baseado em uma cadeia de Markov Oculta do tipo nascimento-e-morte, é proposto e avaliado na Seção 6. Na Seção 7 são apresentadas as considerações finais do artigo e as perspectivas para trabalhos futuros sobre o assunto.

2. Trabalhos relacionados

Em [Nguyen et al. 1996], é apresentado um estudo sobre perda de pacotes em uma rede sem fio WaveLAN, precursora da tecnologia 802.11. Os autores concluem que as perdas não são independentes, o que torna difícil utilizar como modelo uma cadeia de Markov de dois estados. Os autores apresentam uma contribuição importante ao propor um modelo de dois estados baseado nas distribuições do comprimento das seqüências (ou rajadas) de pacotes com erro e sem erro.

Alguns autores de [Karande et al. 2003] propuseram uma nova métrica para avaliação de modelos de perda em [Khayam and Radha 2003]. A métrica foi chamada de ENK (*Entropy Normalized Kullback Leibler*) e consiste basicamente em medir o excesso de informação produzido por um modelo em relação à fonte original, nesse caso o *trace*. Quanto menor o valor de ENK, melhor é considerado o modelo. Os autores usaram essa métrica para avaliar como uma cadeia de Markov de dois estados representa a perda de pacotes em uma rede 802.11b. Diferentemente do que haviam concluído em [Karande et al. 2003], eles chegaram ao resultado do modelo ser satisfatório. De acordo com os autores, a escolha de quais variáveis aleatórias do processo serão avaliadas é fundamental para que a avaliação tenha sucesso. Entretanto, são apresentadas apenas as duas variáveis escolhidas e argumentado que as mesmas são suficientes. O resultado não é

comparado com outras métricas para verificar se a escolha foi apropriada e não é avaliado o impacto que outras variáveis teriam.

Em [Aráuz and Krishnamurthy 2003], os autores avaliam o uso de modelos Markovianos para representar o processo de perda de quadros¹. Os dados para modelagem são extraídos a partir de testes com uma rede 802.11a e uma 802.11b. Foi observado que um modelo com dois estados, um representando perda e o outro sucesso, quando parametrizado por uma distribuição empírica obtida a partir das medições, apresenta melhor resultado que a cadeia de Markov equivalente. Com base em [Nguyen et al. 1996], esse resultado era esperado, porém, em [Aráuz and Krishnamurthy 2003], as distribuições encontradas podem ser representadas apenas com funções exponenciais. Posteriormente, apresentaremos desvantagens desse tipo de modelo que não são discutidas pelos autores.

O trabalho apresentado em [Konrad et al. 2003] se refere a um serviço de dados em uma rede GSM. Apesar das diferenças entre uma rede GSM e uma 802.11, são apresentadas algumas informações gerais que podem ser utilizadas ao se analisar e modelar outros tipos de rede sem fio. Por exemplo, foi utilizado um teste estatístico não-paramétrico, proposto originalmente em [Bendat and Piersol 1986] para verificar estacionariedade em *traces* de perda de quadros/pacotes. Os autores verificam que o modelo de Gilbert² não é apropriado para representar perda em uma rede GSM e propõem um modelo misto com um algoritmo como solução.

Em [Ji et al. 2004], é feita uma nova avaliação do *trace* GSM coletado por [Konrad et al. 2003] e é proposto um novo modelo baseado na combinação de distribuições geométricas. O trabalho confirma a inadequação do modelo de Gilbert na representação do processo de perda de quadros e utiliza algumas métricas adicionais, como coeficiente de variação e autocorrelação, para incrementar as evidências desse fato.

Outro trabalho importante é apresentado em [Yajnik et al. 1999], o qual não se refere a redes 802.11 e nem mesmo a redes sem fio, mas é um trabalho bem elaborado de perda de pacotes na Internet. Os autores apresentam uma ampla avaliação das perdas, estudando o comportamento das rajadas de pacotes transmitidos com sucesso e de pacotes perdidos, calculando a autocorrelação e correlação cruzada das rajadas, discutindo a distribuição das mesmas e avaliando alguns modelos potenciais para a representação de perda de pacotes: Bernoulli, Gilbert e cadeia de Markov de ordem k . Algumas idéias de [Yajnik et al. 1999] são trazidas para o contexto de redes sem fio 802.11 em nosso artigo.

3. Fundamentos

Nessa seção, é proposta uma classificação dos modelos estocásticos que têm sido usados na representação de perda de pacotes. Em seguida, são apresentados dois conceitos fundamentais na análise e modelagem do processo de perda: estacionariedade e autocorrelação.

3.1. Modelos

A modelagem de perda de pacotes em redes sem fio é muitas vezes avaliada por duas óticas diametralmente opostas. Em uma delas, é vista como uma importante ferramenta

¹nesse contexto, são equivalentes a pacotes.

²No modelo de Gilbert há um estado "bom" onde não há perdas, enquanto no modelo de Gilbert-Elliott o estado "bom" também tem uma probabilidade de perda e é, portanto, um modelo mais geral.

para representar o processo de perda em simulações ou análises matemáticas. Em outra, a modelagem é fortemente questionada por representar apenas um conjunto restrito de ambientes ou situações, particulares no tempo e no espaço, muitas vezes de forma excessivamente simples ou imprecisa. Há verdades em ambas as observações e o intuito desse trabalho é mostrar que a modelagem é sim importante, mas que alguns cuidados devem ser tomados para evitar as falhas encontradas em vários trabalhos. Alguns modelos são implementados e avaliados nesse trabalho, além de um novo modelo ser proposto. Os modelos que têm sido usados para descrever a perda de pacotes podem ser agrupados em quatro categorias:

- cadeia de Markov de ordem k ;
- modelo Semi-Markoviano;
- cadeia de Markov Oculta;
- outros modelos baseados em estados.

Dentre os tipos de modelos citados, a cadeia de Markov de ordem k é o mais simples. Quanto maior a ordem da cadeia, maior a sua capacidade de representar processos mais complexos. Porém, o número de estados aumenta exponencialmente, o que pode levar a modelos inviáveis de se implementar. Entre os trabalhos que usam este tipo de modelo estão [Yajnik et al. 1999] e [Khayam 2003].

Um tipo de modelo muito utilizado em perda de pacotes é o Semi-Markoviano, como pode ser visto em [Aráuz and Krishnamurthy 2003], [Carvalho et al. 2005] e [Nguyen et al. 1996]. Para representar perda de pacotes, esse tipo de modelo pode ter apenas dois estados, um no qual os pacotes sempre são perdidos e outro onde sempre são recebidos com sucesso. O tempo de permanência em cada estado é determinado por uma distribuição que precisa ser identificada e parametrizada a partir dos *traces*. Alternativamente, pode se usar com sucesso uma distribuição empírica. Considerando o número de estados, esse modelo é simples, no entanto, identificar a melhor distribuição e sua parametrização pode tornar a tarefa de modelagem complexa. Por um lado, esse tipo de modelagem reflete com acurácia os *traces* capturados, por outro, o modelo obtido está fortemente restrito a esses mesmos *traces*. A distribuição obtida com um conjunto de medidas pode ser bem diferente de outra realizada em ambiente diferente, ou simplesmente em um período de tempo diferente.

Mais recentemente, modelos de Markov Ocultos (*Hidden Markov Models* - HMM) têm sido usados para representar perda de pacotes em redes sem fio, como pode ser visto em [Salamatian and Vaton 2001] e [Lo and Ngai 2004]. Utilizando HMM, é possível representar o processo de perda de forma acurada com um número reduzido de estados. O próprio modelo de Gilbert-Elliot ([Gilbert 1960], [Elliot 1963]) é um modelo de Markov Oculto, porém com apenas dois estados. Inicialmente, o modelo de Gilbert-Elliot foi proposto para tratar a perda de bits em um canal de comunicação, mas o mesmo é comumente usado para tratar também a perda de pacotes. A diferença é que o processo formado por 0s e 1s se refere agora a pacotes recebidos com sucesso ou perdidos, respectivamente, enquanto originalmente o mesmo se referia a bits recebidos ou perdidos. Além disso, um HMM pode ser usado também como um modelo de previsão, ou seja, para antecipar o comportamento futuro da perda com base em seu histórico. Todavia, a identificação do número adequado de estados e da estrutura da cadeia (transições entre os estados) são, geralmente, tarefas não triviais.

Por fim, temos outros tipos de modelos que podem utilizar algum artifício adicional ou uma modelagem totalmente direcionada ao que se quer representar. Como exemplos desses modelos, temos o MTA (*Markov-based Trace Analysis*) que combina um algoritmo com uma cadeia de Markov de ordem k ([Konrad et al. 2003]), *On-Off* Estendido ([Ji et al. 2004]) que usa uma combinação de distribuições geométricas, os modelos discutidos em [Lo and Ngai 2004] que utilizam estados adicionais para representar rajadas de perda mais longas e [Karande et al. 2003] que propõe uma cadeia de Markov hierárquica. Como vantagem, esses modelos apresentam uma representação adequada do processo de perda, mantendo um número tratável de estados. No entanto, apresentam duas desvantagens significativas. A primeira é que esses modelos são altamente associados aos *traces* que representam e uma pequena variação pode levar a uma grande dificuldade de reconfiguração (re-parametrização), podendo exigir até mesmo uma reconstrução do modelo. A segunda desvantagem é a própria criação do modelo, a qual pode ser altamente complexa, exigindo muita experiência e habilidade estatística.

3.2. Estacionariedade e autocorrelação

A não-estacionariedade da perda de pacotes está geralmente associada a um período transitente da rede e, portanto, é importante identificar quando esse fenômeno ocorre a fim de decidir como tratá-lo. Em vários modelos estocásticos, os *traces*, ou partes dos mesmos, que apresentam comportamento não-estacionário são descartados. De acordo com [Box et al. 1994], uma definição intuitiva de um processo estacionário é dada por um processo que permanece em equilíbrio em torno de um nível médio constante. Geralmente, se um processo (ou série de tempo) é **estacionário**, o mesmo pode ser descrito através de sua **média**, **variância** e **função de autocorrelação**. Essas propriedades serão usadas posteriormente na avaliação de modelos de perda. Usando uma definição mais rigorosa, como a apresentada em [Trivedi 2002], temos dois tipos de estacionariedade: Estrita (ou Forte) e Ampla (ou Fraca ou de Segunda ordem).

Um processo aleatório $\{X(t)|t \in T\}$ é estritamente estacionário se sua função distribuição de probabilidade³ conjunta de n -ésima ordem satisfizer, para $n \geq 1$, a seguinte condição:

$$F(x; t) = F(x; t + \tau) \quad (1)$$

para todos vetores $x \in \mathbb{R}^n$ e $t \in T^n$, e todos escalares τ , tais que $t_i + \tau \in T$. Onde $t + \tau$ significa que o escalar τ (geralmente chamado de atraso ou defasagem) é adicionado para todos os componentes do vetor t .

Para ser considerado estacionário de segunda ordem, um processo estocástico deve satisfazer as seguintes condições:

1. $\mu(t) = E[X(t)]$ é independente de t ,
2. $\gamma(t, t + \tau) = Cov(X(t), X(t + \tau)) = \gamma(\tau)$ é independente de t para cada τ ,
3. $\gamma(0) = E[X^2(t)] < \infty$, ou seja, existe segundo momento finito.

Para utilizar a estacionariedade forte é necessário determinar a CDF conjunta de um processo aleatório. Na prática, esse é um procedimento muito difícil, e na maior parte

³CDF - *Cumulative Distribution Function* - Função Distribuição de Probabilidade

das vezes até inviável. Por essa razão, operacionalmente se utiliza a estacionariedade fraca. Ainda assim, verificar a estacionariedade de dados coletados de ambientes reais é uma tarefa complexa. Essa é, provavelmente, uma das razões para que vários trabalhos negligenciem a análise dessa propriedade. Para determinados tipos de processos e dependendo do tipo de modelagem que se pretende utilizar, esse tipo de negligência pode levar a erros graves nos resultados obtidos.

Nesse trabalho, utilizamos dois testes estatísticos não-paramétricos para verificação de estacionariedade: Teste de Iteração (*Run Test*) e Teste de Arranjos Reversos (*Reverse Arrangements Test*). Ambos são testes de hipótese, ou seja, se baseiam em uma hipótese nula (h_0) e em uma hipótese alternativa (h_1). Nesse caso, h_0 diz que o processo sendo avaliado não possui tendência, ou seja, é estacionário e h_1 estabelece que há alguma tendência e, portanto, é não-estacionário. Como qualquer outro teste de hipótese, os testes de estacionariedade estão sujeitos a erros do Tipo I (falso negativo) e do Tipo II (falso positivo). Realizamos todos os testes com um nível de significância de 0,02. Maiores detalhes sobre os mesmos podem ser obtidos em [Bendat and Piersol 1986].

Em geral, é útil ter uma medida que seja independente de escala, a qual é obtida com o **coeficiente de correlação** definido da seguinte forma:

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var[X] \cdot Var[Y]}} = \frac{\gamma(X, Y)}{\sigma_X \cdot \sigma_Y} \quad (2)$$

O coeficiente de correlação $\rho(X, Y)$ está definido dentro do intervalo $[-1 : 1]$, onde mais próximo de zero indica menor correlação (linear). O coeficiente de correlação pode ser utilizado para variáveis aleatórias de um mesmo processo estocástico e, nesse caso, é obtida a correlação entre $X(t)$ e $X(t + h)$, onde h é o atraso (*lag*) ou defasagem. Assim, tem-se o coeficiente de **autocorrelação**.

[Brockwell and Davis 2002] apresenta uma estimativa para a função de autocorrelação, também conhecida como **função de autocorrelação amostral** que é útil quando se precisa obter o coeficiente de autocorrelação a partir dos dados coletados de um processo que não se conhece. A função de autocorrelação amostral é dada por:

$$\hat{\rho}(\tau) = \frac{\hat{\gamma}(\tau)}{\hat{\gamma}(0)}, -n < \tau < n \quad (3)$$

Onde a **função de autocovariância amostral** é dada por:

$$\hat{\gamma}(\tau) = n^{-1} \sum_{t=1}^{n-|\tau|} (x_{t+|\tau|} - \bar{x})(x_t - \bar{x}), -n < \tau < n \quad (4)$$

Ao se usar autocorrelação amostral é definido também um intervalo de confiança para identificar a partir de qual limite um valor pode ser considerado zero, ou seja, sem autocorrelação. Esse intervalo é dado por $\pm \frac{Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}$.

Conforme [Box et al. 1994], para se obter, na prática, uma estimativa útil da função de autocorrelação é necessário ter no mínimo 50 amostras. Além disso, $\hat{\rho}(\tau)$

seria calculado para $\tau = 0, 1, 2, \dots, T$, onde T não é maior que $n/4$. Nesse trabalho, utilizamos a função de autocorrelação amostral como uma métrica de avaliação de modelos para perda de pacotes.

4. Medições e análise da perda de pacotes

Nessa seção são apresentados os detalhes sobre os equipamentos, softwares e metodologia utilizados na geração e coleta dos *traces* usados para a análise e modelagem do processo de perda de pacotes.

4.1. Topologia e configuração

Para realizar a coleta do tráfego que foi usado na análise e modelagem da perda de pacotes, foi criada um ambiente de testes que representa uma rede 802.11b *indoor* típica. De fato, é difícil definir um ambiente físico que possa ser considerado representativo para todas as redes locais sem fio. No entanto, é razoável considerar que a maior parte das redes sem fio em um ambiente *indoor* deve lidar com paredes, portas, pessoas e vários outros objetos. Em nosso ambiente de testes tentou-se representar duas situações comuns, sendo uma onde há menor atenuação e outra onde o número de obstáculos é relativamente maior.

A Figura 1 mostra uma representação esquemática do ambiente usado para os testes, sendo que à esquerda é apresentada um vista lateral e à direita uma vista superior. O enlace formado pelos equipamentos indicados com os números 1 e 2 apresentam menor atenuação do sinal, pois estes equipamentos estão separados apenas por um piso de madeira com armação metálica e uma parede fina (do tipo divisória). Os enlaces formados pelos equipamentos 1-3 e 2-3 sofrem maior atenuação do sinal, pois o mesmo precisa atravessar duas paredes duplas de tijolos, embora as condições de multipercurso possam levar o sinal por caminhos com atenuação um pouco menor, uma vez que as características físicas (sobretudo portas e janelas) do ambiente utilizado permitem esse fato.

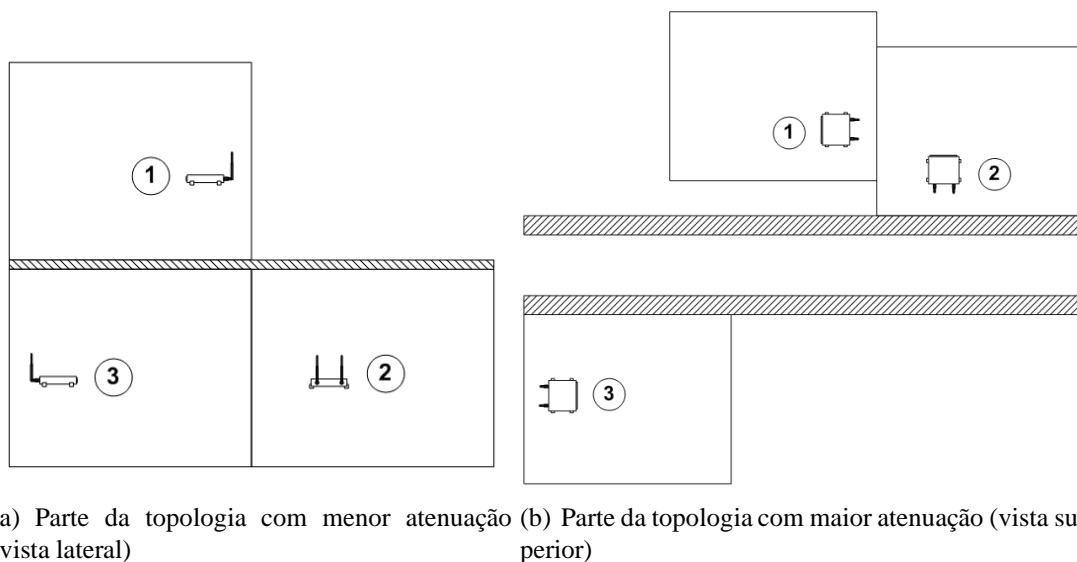


Figura 1. Topologia utilizada durante os testes.

Os testes foram realizados com equipamentos LinkSys WRT54G, tendo o *firmware* original sido substituído pela distribuição GNU/Linux OpenWrt na versão WhiteRussian RC4 [OpenWrt 2006]. Foi usada uma potência de transmissão de 100 mW e a

rede foi configurada em modo *ad-hoc* para permitir que os pares se comunicassem diretamente e se reduzisse a quantidade de informações de controle. Os pacotes foram gerados em *multicast* para evitar que o controle de erro da camada de enlace fosse ativado e, então, ocorressem retransmissões de pacotes perdidos. Além disso, os testes foram feitos com um transmissor e dois receptores simultâneos. Foi utilizada uma versão modificada e portada para a plataforma MIPS da ferramenta *Tangram-II Traffic Generator* [LAND 2006] para geração dos pacotes.

Inicialmente, foi realizada uma pesquisa de campo (*site survey*) para identificar quais canais estavam disponíveis. O canal 1 (um) foi escolhido por não apresentar outros equipamentos no mesmo, no entanto, não é possível garantir que esse canal não sofra interferências eventuais. Os equipamentos possuem também portas Ethernet, sendo que uma delas foi usada para conectar o equipamento que realizava a coleta dos dados. Durante os testes, os equipamentos não foram movidos, no entanto, era esperado um grande número de variações de curta duração no sinal devido ao movimento de pessoas, sobretudo em alguns períodos do dia.

A avaliação da perda de pacotes em redes 802.11b foi feita utilizando pacotes de 500 e 1400 *bytes* e com as interfaces configuradas para transmitir com taxa fixa de 11Mbps. Essas escolhas se baseiam em experimentos realizados, os quais não serão detalhados por restrições de espaço do documento. Um breve resumo desses experimentos e as principais constatações dos mesmos seguem.

Primeiro, foram realizados testes variando o tamanho dos pacotes de 100 a 1400 *bytes*. Constatamos que existe influência do tamanho do pacote na taxa de perda de uma rede 802.11b, no entanto, essa influência não segue uma função linear com crescimento estritamente monotônico, mas uma espécie de função degrau, na qual é possível distinguir com clareza dois patamares. O patamar mais baixo ocorre até um tamanho de pacote entre 700 e 900 *bytes*. Em testes posteriores, se verificou que embora exista um aumento na taxa média de perda para diferentes tamanhos de pacotes, outras estatísticas importantes não são influenciadas, por exemplo a função distribuição de probabilidade do tamanho das rajadas de perda.

Segundo, a influência das taxas de transmissão de uma interface 802.11b foi avaliada. Os testes consistiram em usar as 4 velocidades disponíveis: 1, 2, 5,5 e 11Mbps, verificando como as mesmas afetam a taxa média de perda. Não foi observada diferença significativa da taxa média de perda de pacotes para as diferentes velocidades da interface. Resultado semelhante foi obtido em [Aguayo et al. 2004] para ambientes *outdoor* e em [Bicket 2005] para ambientes *indoor*. A taxa de 11 Mbps foi escolhida por ser, operacionalmente, capaz de atender o maior número de aplicações e, portanto, de maior interesse para uma rede 802.11b.

4.2. Taxa de geração de pacotes

Essa seção apresenta uma avaliação da influência da taxa de geração de pacotes da fonte sobre o processo de perda de pacotes. O intuito é identificar qual taxa de geração de pacotes é adequada para uma rede 802.11b, de forma a extrair apropriadamente o processo de perda. Essa verificação é importante porque a geração escolhida foi a periódica e, portanto, se faz necessário encontrar uma taxa que seja suficiente alta para capturar o processo que se deseja observar [Tariq et al. 2005]. Por outro lado, a taxa de geração

de pacotes não pode ser arbitrariamente alta devido à limitação da vazão efetiva da interface, a qual é mais baixa que a velocidade nominal informada, e devido à restrição da capacidade de processamento dos equipamentos utilizados.

Para se avaliar a influência da taxa de geração de pacotes, foram feitos testes com intervalos entre pacotes de 5, 10, 20 e 50 ms. Para cada intervalo foram realizados 3 testes, iniciando às 08:00, 14:00 e 20:00, com 360.000 amostras cada. Não houve indicações de relação entre a taxa de geração e o processo formado pela taxa média de perda. Essa conclusão se baseou no comportamento variável da taxa média de perda, a qual não apresentou diferença significativa em nenhum dos intervalos de geração. No estudo de perda de pacotes, o comprimento das rajadas de perda é uma propriedade muito importante, uma vez que seu comportamento tende a permanecer similar, independentemente da taxa média de perda observada [Sanneck and Carle 2000]. Muitos modelos se baseiam nessa invariante para representar a perda de pacotes. Logo, a função densidade de probabilidade⁴ e a CDF do tamanho das rajadas foram calculadas, conforme pode ser visto na Figura 2. Na figura são apresentados os resultados para as amostras agregadas dos testes nos 3 horários diferentes. As amostras de cada horário foram analisadas separadamente e apresentaram um comportamento semelhante ao observado na agregação no que se refere ao comprimento das rajadas.

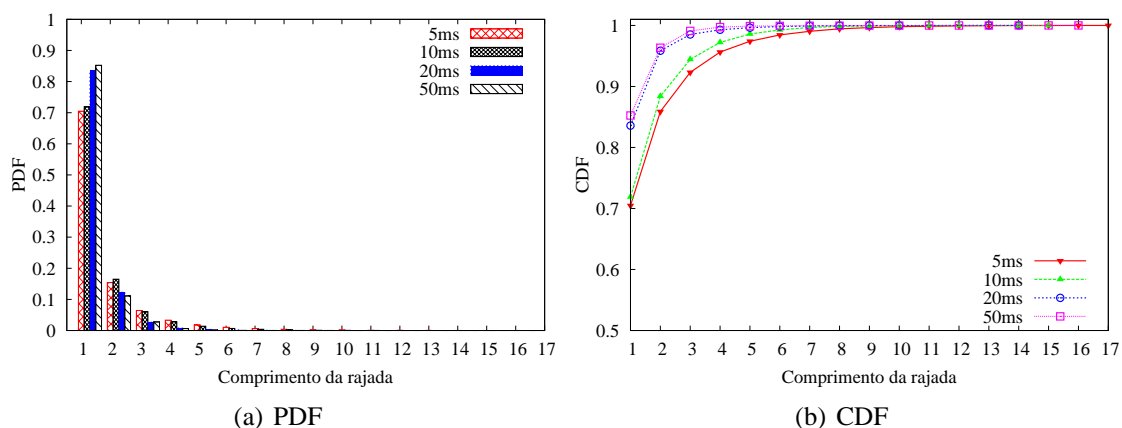


Figura 2. Tamanho das rajadas de perda para diferentes intervalos entre pacotes.

Na Figura 2, é possível observar que o comprimento das rajadas de perda é afetado pelo intervalo de geração dos pacotes. Quanto menor a frequência de geração, maior é o número de perdas isoladas, ou seja, rajadas de tamanho um. Além disso, há uma diferença significativa na quantidade de rajadas de tamanho 2 até 7, quando o intervalo de geração é maior que 10 ms. De fato, pode-se dizer que a diferença das funções de densidade entre 5 e 10 ms é negligenciável e, portanto, qualquer uma das duas taxas de geração seria adequada para avaliação das perdas. No entanto, foi observada uma variância muito maior na taxa de geração de 5 ms em relação à de 10 ms. A média do intervalo de geração de 5 ms das três amostras apresentadas foi de 5,027 ms, com um desvio padrão médio de 3,596 ms. Para o intervalo de 10 ms, o valor médio foi de 10,002 ms e o desvio padrão médio de 0,999 ms. Logo, o intervalo de 10 ms foi escolhido para a geração dos pacotes. Essa variância é justificada pelo *hardware* e *software* utilizados, os quais não apresentam

⁴PDF - Probability Density Function

suporte a tempo-real e, portanto, não conseguem oferecer garantias rígidas de tempo em escalas na ordem de poucas dezenas de milissegundos e inferiores.

4.3. Estatísticas básicas

Nesta seção, são analisados os resultados de 15 *traces* de 24 horas consecutivas cada, ou seja, 360 horas no total. Inicialmente, foi observado o comportamento das perdas ao longo do tempo, com o intuito de identificar algum padrão que influenciasse a taxa média de perda. Previamente, era esperado que duas características distintas do ambiente criassem alguma distinção na taxa média de perda. A primeira é responsável pelas variações de longo prazo nas perdas e está relacionada com elementos como paredes, *layout* da mobília, posição dos equipamentos de comunicação e similares, para os quais a mobilidade é mínima ou inexistente. A segunda característica diz respeito às condições que mudam em pequenos intervalos de tempo como movimento de pessoas, abertura/fechamento de portas e janelas, alteração da disposição de objetos, uso de equipamentos que podem gerar interferência, etc. Esse último tipo de evento pode causar flutuações significativas na taxa média de perda em diferentes escalas de tempo [Rappaport 2001].

Dadas as características do ambiente onde foram realizados os testes, seria esperado que durante a maior parte do tempo houvesse predominantemente perdas provocadas pelas características do ambiente e durante os períodos de expediente ocorresse uma variância maior provocada por eventos semelhantes aos citados anteriormente. Como é comentado a seguir, oito *traces* foram bem comportados nesse sentido, no entanto, os outros sete exibiram algumas peculiaridades.

Na Figura 3, é mostrado um *trace* de 24 horas com a taxa média de perda suavizada dentro de uma janela de 60.000⁵ pacotes, ou seja, 10 minutos. Pela figura, é possível observar que entre 09:00 e 18:00 ocorrem as maiores variações na taxa média de perda. O restante do período exibe uma taxa média apenas com pequenas variações.

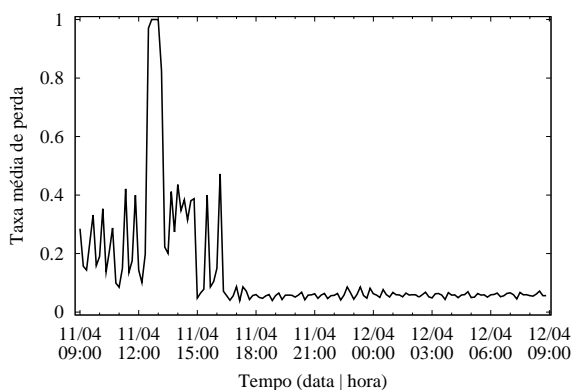


Figura 3. Taxa média de perda com um comportamento típico.

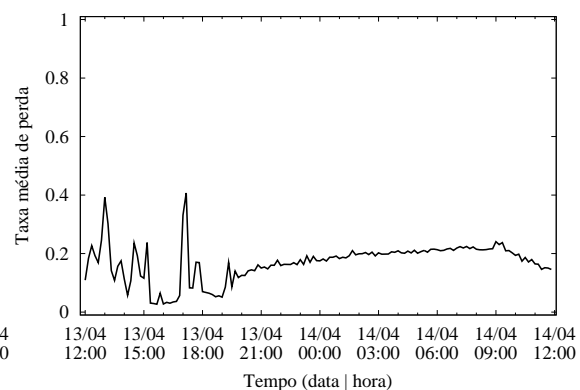


Figura 4. Taxa média de perda com um comportamento peculiar.

Na Figura 4, é apresentado outro *trace* de 24 horas, novamente com a taxa média de perda suavizada dentro de uma janela de 60.000 pacotes. No gráfico ainda é possível observar as maiores variações no período de expediente do laboratório. No entanto, é

⁵A escolha de uma janela de 60.000 pacotes se baseou na avaliação de janelas de outros tamanhos, a qual demonstrou que o valor escolhido seria adequado.

vista uma elevação gradual da taxa média de perda a partir das 18:00 até as 09:00 do dia seguinte. De fato, essa imprevisibilidade das perdas é descrita em outros trabalhos sobre redes sem fio como [Bai and Atiquzzaman 2003] e [Gaertner and Cahill 2004], tanto da tecnologia 802.11 quanto em outras. Logo, se faz necessário avaliar outras estatísticas das perdas que sejam mais favoráveis à modelagem.

Como foi visto, é difícil identificar características na taxa média de perda ao longo do tempo em redes 802.11. Mesmo em diferentes tamanhos de janela, o processo que representa a taxa média de perda se mostra de difícil identificação e modelagem. Uma propriedade da perda que tem sido muito usada em sua modelagem é o tamanho das rajadas. O modelo de Gilbert-Elliot e outros modelos baseados em estados são construídos usando os comprimentos das rajadas de perda e sucesso, e como elas se alternam. Na Figura 5, são mostradas as rajadas de perda identificadas ao longo das 360 horas. Pode-se observar que menos de 0.1% das rajadas é superior a 23 pacotes e, portanto, haveria pouco ganho na acurácia do modelo a representação de rajadas maiores que esse valor.

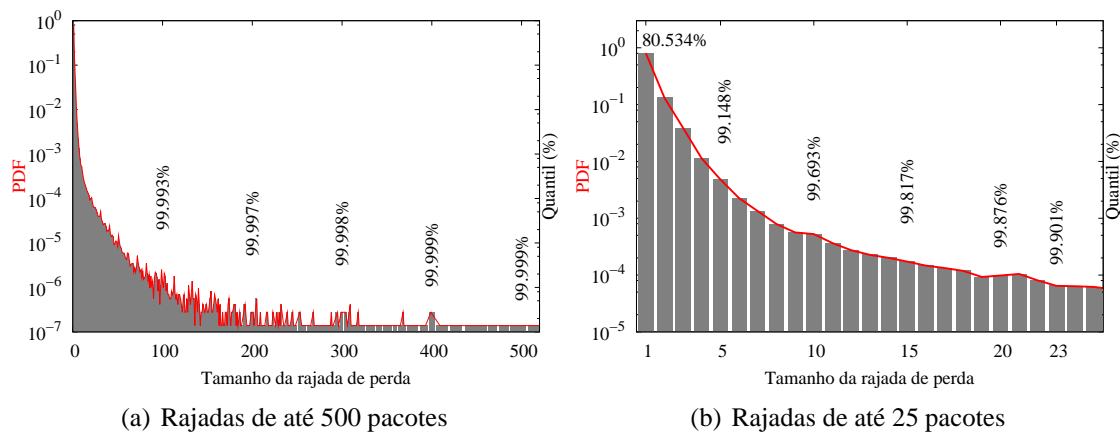


Figura 5. Tamanho das rajadas de perda.

Também é importante identificar o tamanho das rajadas de sucesso, ou seja, seqüências de pacotes sem perda e a proporção de tempo em que o processo permanece em cada "estado" de perda ou sucesso. No entanto, essas duas propriedades, assim como a taxa média de perda, são específicas de um determinado ambiente. Embora tenho sido feito um grande número de experimentos e o resultado tenha sido comparado com outros trabalhos, é difícil obter valores para essas duas propriedades que sejam representativos. Ao se caracterizar e/ou modelar uma rede sem fio 802.11, é importante estar atento a esse fato: algumas propriedades fundamentais são específicas de um ambiente ou conjunto de ambientes e, portanto, uma modelagem específica deve ser feita em cada contexto.

4.4. Análise da estacionariedade

Antes de aplicar os testes aos *traces*, esses foram divididos em *subtraces* de 1 hora cada. Os *traces* menores facilitam a análise das propriedades estatísticas (inclusive da estacionariedade) e o tamanho escolhido é grande o suficiente para capturar as características importantes do processo aleatório associado à perda de pacotes. Portanto, os testes são realizados em 360 *traces* de 1 hora cada. Os resultados são apresentados a seguir.

De forma semelhante a [Ji et al. 2004], para realizar os testes de estacionariedade assumimos que a perda de pacotes em uma rede 802.11 é descrita por um processo aleatório $\{X_i\}_{i=1}^n$, onde n é o número total de amostras na sequência e cada amostra X_i pode assumir um valor "1" quando ocorre uma perda ou "0" quando o pacote é recebido com sucesso. É importante aplicar os dois testes a cada *trace* porque o teste de iteração é mais poderoso para detectar tendências flutuantes, enquanto que o teste de arranjos reversos é mais poderoso na identificação de tendências monotônicas.

A Figura 6 mostra a aplicação dos dois testes de estacionariedade a seis *traces* diferentes. Pelo gráfico, a estacionariedade é verificada através da posição em relação aos limites inferior e superior, os quais são definidos pelo nível de significância escolhido para os testes. Ser estacionário significa estar entre os limites inferior e superior. Para cada teste, foram escolhidos *traces* que ilustram os três resultados possíveis: estacionário, não-estacionário e estacionário a partir de um determinado número de amostras. Esse último resultado é equivalente ao primeiro, mas é interessante destacá-lo porque o *trace* não pode ser considerado estacionário em todas as escalas de tempo. Por exemplo, o *subtrace* 1 não é estacionário para escalas de tempo inferiores a 50 segundos, sendo que a escala de tempo é obtida pelo produto entre o número de amostras (indicado no gráfico) e o intervalo de geração entre pacotes (10 ms). Podemos observar que os *subtraces* 3 e 6 não são estacionários. A estacionariedade dos 360 *subtraces* foi testada, sendo que 14 foram considerados não-estacionários pelo teste de iteração e 19 foram identificados como não-estacionários pelo teste de arranjos reversos. Ou seja, mais de 90% dos *traces* apresentaram um comportamento estacionário, sendo utilizados na análise e modelagem da perda de pacotes.

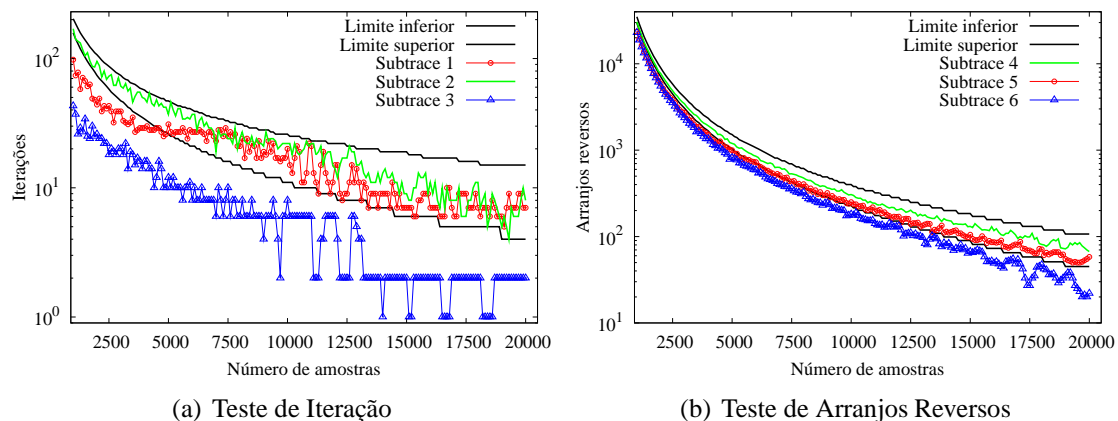


Figura 6. Testes de estacionariedade.

5. Avaliação do modelo de Gilbert-Elliot

Muitos usuários necessitam de um modelo que seja fácil de parametrizar de forma que o mesmo possa capturar um novo processo de forma simples e rápida. Essa propriedade é muitas vezes tão importante que até a acurácia do modelo é negligenciada. Essa é uma das razões para modelos simples, como Gilbert-Elliot, serem tão utilizados. De fato, a menos das cadeias de Markov de ordem k , todos os demais modelos de perda de pacotes demandam um esforço significativo para realizar uma nova parametrização. Entretanto,

até mesmo usando uma cadeia de ordem k , o usuário precisa saber qual a ordem mínima da cadeia que atenda as suas necessidades, exigindo uma análise estatística do processo.

Nessa seção, foram desenvolvidos dois modelos Markovianos usando os *traces* coletados: o modelo de Gilbert-Elliot e uma cadeia de Markov de ordem 4. De fato, dada escala de tempo utilizada e a autocorrelação observada nos *traces* seria necessário o uso de uma cadeia de ordem > 10 . No entanto, isso implicaria em um modelo com mais de 1000 estados, o que não é desejável em um modelo simples. Ainda assim, a cadeia de Markov de ordem 4 cumpre sua função de servir como referência para comparação com o modelo de Gilbert-Elliot, sobretudo com relação ao comprimento das rajadas de perda. Para realizar o treinamento da cadeia, ou seja, a estimação dos parâmetros do modelo de Gilbert-Elliot, foi utilizado o algoritmo Baum-Welch. Há várias implementações desse algoritmo, sendo que nesse trabalho foi usada a ferramenta Jahmm[François 2007]. Todos os *traces* utilizados na criação e avaliação dos modelos foram identificados como estacionários de acordo com os testes de estacionariedade apresentados na Seção 3.2.

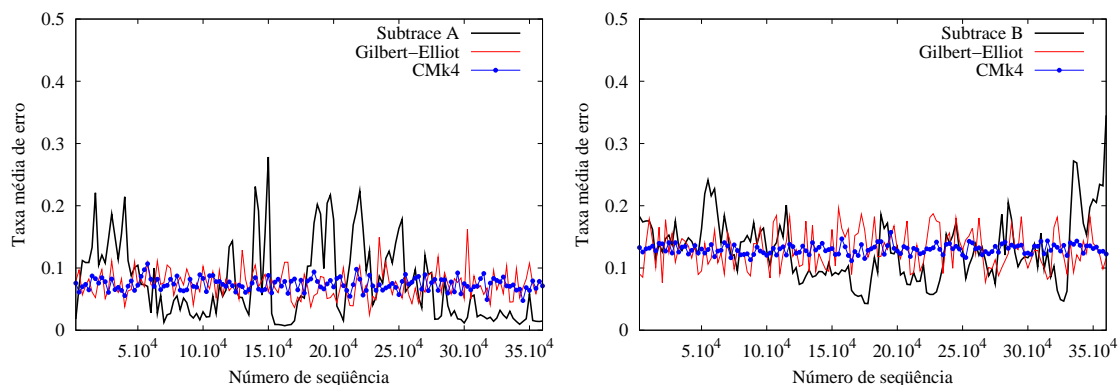


Figura 7. Avaliação da taxa média de perda de pacotes.

Nos gráficos da Figura 7, é mostrada a taxa média de perda em uma janela de 2.500 pacotes (ou 25 segundos). Em cada gráfico, as curvas correspondem a um *subtrace* original, a um modelo de Gilbert-Elliot e uma cadeia de Markov de ordem 4 (CMk4). É possível notar que tanto o modelo de Gilbert-Elliot como a cadeia de ordem 4 conseguem capturar a taxa média do *subtrace* completo. Embora a figura não exiba de forma nítida, ambos são incapazes de representar adequadamente a média e a variância em intervalos mais curtos. Essa informação é apresentada pela Tabela 1, a qual apresenta a variância da taxa média de perda para os *subtraces* originais e seus respectivos modelos de Gilbert-Elliot e cadeia de Markov de ordem 4. Pode ser observada uma acurácia maior do modelo de Gilbert-Elliot em relação à cadeia de ordem 4, ainda que a variância seja significativamente menor que a do processo de perda original.

	Variância
<i>Subtrace A</i>	0.003800
Gilbert-Elliot A	0.000499
CMk4 A	0.000104

	Variância
<i>Subtrace B</i>	0.002624
Gilbert-Elliot B	0.000911
CMk4 B	0.000058

Tabela 1. Variância da taxa média de perda

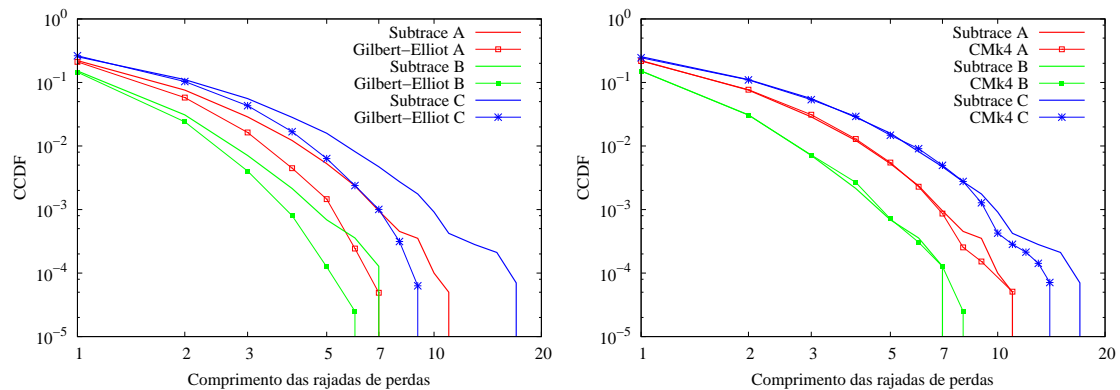


Figura 8. Avaliação do comprimento das rajadas de perda.

A Figura 8 apresenta a CDF complementar (CCDF⁶) do tamanho das rajadas de perda. Os *subtraces* escolhidos mostram três distribuições diferentes de tamanho de rajadas a fim de ilustrar como os modelos as representam. Como pode ser visto, o modelo de Gilbert-Elliot é incapaz de representar adequadamente a distribuição do tamanho das rajadas. Por outro lado, a cadeia de ordem 4 consegue criar uma distribuição de perdas semelhante à observada pelo processo original. Esse resultado ilustra como uma modelagem inadequada pode levar a resultados distorcidos, pois alguns mecanismos de controle automático de taxa e o controle de erro da camada de enlace das redes 802.11 são altamente sensíveis ao comprimento das rajadas de erro. Por exemplo, enquanto o *subtrace* B exibe rajadas com comprimento de até 7 pacotes, seu modelo de Gilbert-Elliot equivalente mostra apenas rajadas mais curtas. Nessa situação, o mecanismo de controle de erro da camada de enlace que notifica a camada superior quando há 7 perdas consecutivas não produziria o resultado esperado.

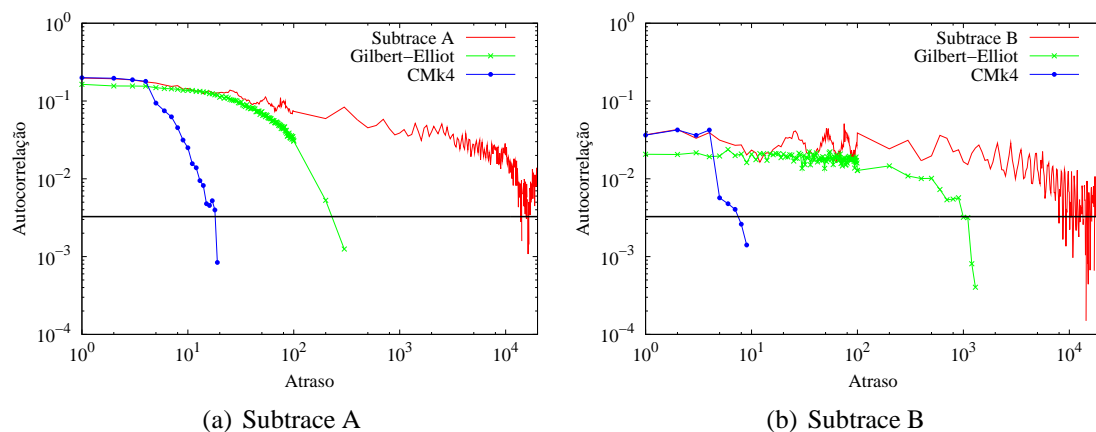


Figura 9. Avaliação da autocorrelação.

Na Figura 9, é mostrada a autocorrelação de dois *subtraces* e de seus respectivos modelos. A reta paralela ao eixo das abscissas indica o limite (superior) de confiança do nível de significância de 0,05 utilizado na medição da autocorrelação. Valores abaixo do limite superior de confiança indicam que a autocorrelação nesses pontos pode ser

⁶CCDF complementar é igual a $1 - CDF$ e, nesse caso, proporciona melhor visualização dos dados

considerada nula. Mais uma vez o modelo de Gilbert-Elliot se mostra inadequado na representação do comportamento do *subtrace* original. O modelo de Gilbert-Elliot para o *subtrace* A mostra autocorrelação não nula até uma defasagem próxima de 10^2 , enquanto que o mesmo tipo de modelo para o *subtrace* B exibe atraso em torno de 10^3 . Para ambos os *subtraces*, a cadeia de Markov de ordem 4 alcança um atraso da ordem de 10^1 . Os dois *subtraces* originais apresentam autocorrelação não nula até uma defasagem próxima de 10^4 , ou seja, significativamente mais alta que a exibida pelos dois modelos.

A partir desses resultados, podemos concluir que o modelo de Gilbert-Elliot pode ser usado quando se deseja capturar apenas a taxa média de perda, com pequena variância. Em alguns casos esse pode ser o comportamento mais comum da perda em uma rede, no entanto, em nossos experimentos observamos alta variância em torno da taxa média durante longos períodos de tempo. A cadeia de Markov de ordem 4 conseguiu representar com acurácia satisfatória o comprimento das rajadas de perda, no entanto, não foi capaz de modelar com precisão a variância e a autocorrelação do processo. Aumentar a ordem da cadeia melhora seu desempenho, mas elevaria exponencialmente o número de estados do modelo. Em [Khayam 2003], é apresentada uma técnica para manter tratável esse número de estados, porém, a cadeia de Markov de ordem k perde sua principal vantagem: a baixa complexidade de modelagem. O problema do aumento da complexidade ocorre também ao se utilizar um modelo de Markov Oculto com muitos estados, sobretudo com relação às transições entre os estados.

6. Modelo de Markov Oculto com estrutura do tipo nascimento-e-morte

Conforme foi apresentado anteriormente, um modelo Markoviano Oculto torna-se mais complexo à medida que seu número de estados aumenta. De fato, não é apenas o número de estados de um modelo HMM que o torna complexo, mas também a estrutura a ser escolhida para a cadeia. A estrutura está associada com o processo do mundo real que se deseja representar, ou seja, o número de estados e as transições entre os mesmos são escolhidos de forma a descrever o fenômeno em estudo, sob o ponto de vista de quem cria o modelo. Esse procedimento existe em qualquer tipo de modelo, porém modelos de Markov Ocultos apresentam a dificuldade adicional do processo (que se deseja modelar) não ser observado diretamente. O modelo de Gilbert-Elliot é um exemplo de HMM com uma estrutura simples, mas eficiente em um grande número de representações. Isto significa que o modelo de Gilbert-Elliot pode ser usado como ponto de partida para a criação de um modelo mais sofisticado, sendo essa a abordagem desse trabalho.

Basicamente, o modelo de Gilbert-Elliot descreve um canal de comunicação que alterna entre um estado "bom" e um estado "ruim". Uma probabilidade de perda maior é associada ao estado ruim e, em condições normais, o canal tende a ficar mais tempo no estado bom. Partindo dessa idéia, poderíamos descrever um canal de comunicação com mais níveis de qualidade, por exemplo: "bom", "regular" e "ruim". Inicialmente, poderíamos estabelecer que é possível mudar entre quaisquer níveis. A Figura 10 ilustra um modelo geral (HMMg) com 3 estados do canal e as transições entre os mesmos. Uma variante desse modelo é apresentada na Figura 11, a qual tem a estrutura de um processo Markoviano de nascimento-e-morte (HMMnm), ou seja, há probabilidade de transição (não nula) apenas entre vizinhos adjacentes e para o próprio estado. Essa variação é motivada pelo comportamento de um canal sem fio, no qual as mudanças entre níveis de qualidade geralmente não ocorrem de maneira instantânea, embora possam ser muito

rápidas conforme foi discutido na Seção 4.3.

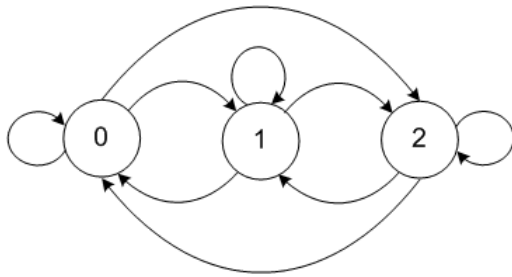


Figura 10. Cadeia de Markov com uma estrutura geral.

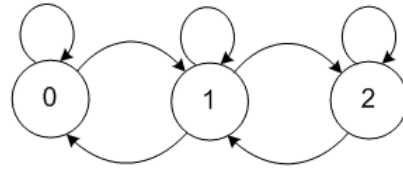


Figura 11. Cadeia de Markov com uma estrutura do tipo nascimento-e-morte.

Todos os modelos HMM implementados foram treinados com o algoritmo Baum-Welch, usando 1000 iterações que foi o mesmo número usado para otimizar o modelo de Gilbert-Elliot. Para avaliação dos modelos foi utilizada a mesma metodologia apresentada na seção anterior, ou seja, foram examinadas a taxa média, variância, comprimento das rajadas de erro e autocorrelação. O modelo HMM geral (Figura 10) apresentou um desempenho superior ao do Gilbert-Elliot em algumas métricas de alguns *traces*. No entanto, com determinados *traces*, o modelo HMM geral teve um comportamento apenas semelhante ao do modelo de Gilbert-Elliot. O modelo HMM do tipo nascimento-e-morte (Figura 11) apresentou um desempenho superior ao do Gilbert-Elliot em todas as métricas de todos os *traces* avaliados. A Figura 12 ilustra os resultados da autocorrelação dos *sub-traces* utilizados anteriormente, incluindo agora os valores obtidos com os dois novos modelos HMM de 3 estados.

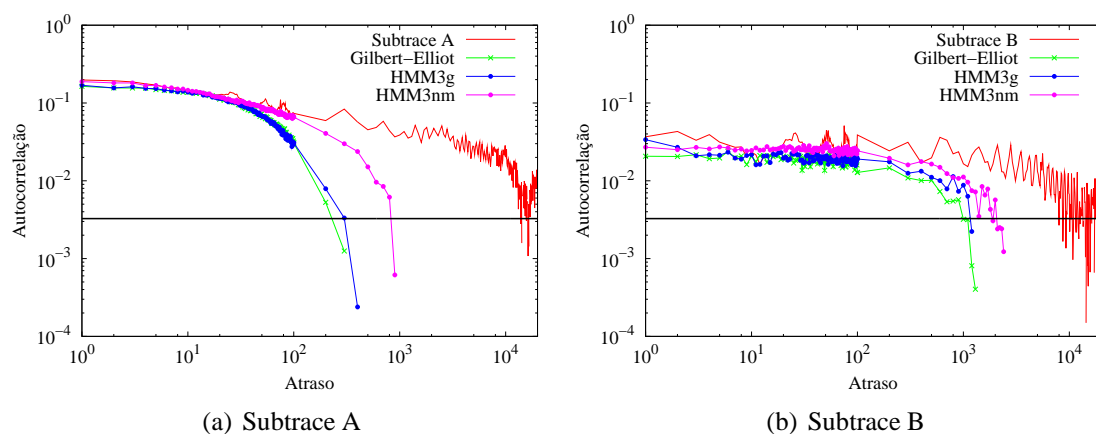


Figura 12. Avaliação da autocorrelação.

Os gráficos da Figura 12 ilustram o comportamento comumente observado com os dois novos modelos, isto é, o modelo HMM com 3 estados e estrutura geral (HMM3g) representa melhor a perda de pacotes que o modelo Gilbert-Elliot em alguns *traces*, mas em outros não há melhoria. Por outro lado, o modelo HMM com 3 estados e estrutura do tipo nascimento-e-morte (HMM3nm) sempre representa melhor a perda de pacotes, tanto em relação ao modelo Gilbert-Elliot, quanto em relação ao próprio modelo HMM geral.

Entretanto, é importante observar que os valores de defasagem da autocorrelação de todos os modelos ainda são inferiores ao do processo original de perda de pacotes.

Retornando à idéia original dos níveis de qualidade de um canal de comunicação, podemos generalizá-la e descrever o canal com quantos níveis se façam necessários para capturar as propriedades estatísticas do processo de perda de pacotes. Para verificar se o aumento do número de estados melhoraria a capacidade de representação do modelo HMM do tipo nascimento-e-morte, foram criados modelos com número crescente de estados. Foi observado que a medida que o número de estados aumenta, a precisão do modelo é incrementada, no entanto, isso ocorre até um determinado número de estados para cada *trace*. Ou seja, cada *trace* demanda um modelo com um determinado número máximo de estados, a partir do qual se começa a ter oscilações no desempenho dos modelos. No entanto, isso não é considerado uma desvantagem do modelo, uma vez que a sua estrutura não se modifica. Além disso, o modelo se mostrou robusto em relação aos valores de inicialização da cadeia de Markov.

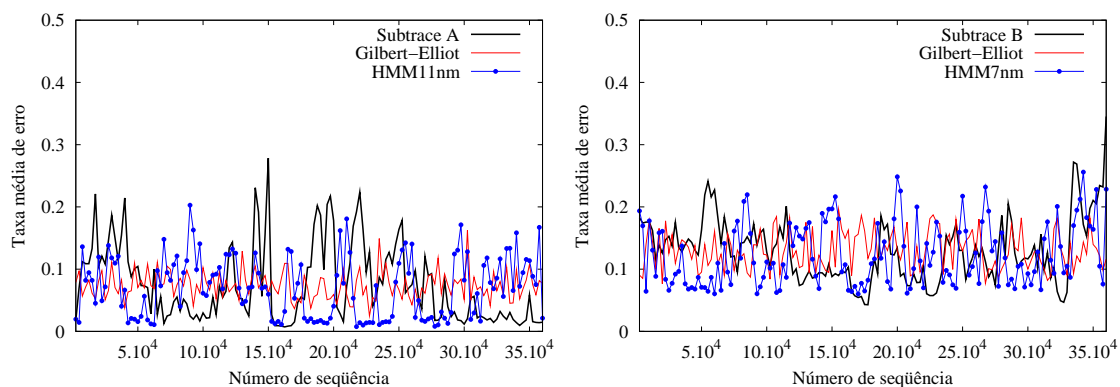


Figura 13. Avaliação da taxa média de perda de pacotes.

As Figuras 13, 14 e 15 mostram como o novo modelo se comportou ao representar os *traces* avaliados na seção anterior. Como pode ser visto, houve melhoria significativa em todas as métricas analisadas. A Tabela 2 confirma os resultados dos gráficos para a métrica variância. Nenhum dos *traces* avaliados necessitou mais que 11 estados para ser representado. Isso significa que um modelo HMM com 11 estados e estrutura do tipo nascimento-e-morte seria capaz de representar adequadamente todos os *traces* analisados. Para evitar as oscilações do modelo, o número de estados efetivamente visitados deveria ser alterado para cada *trace*. Foi observado também que a métrica que demanda o maior número de estados para sua correta representação é a autocorrelação. O comprimento das rajadas e a variância foram capturados por modelos com no máximo 7 estados.

	Variância
<i>Subtrace A</i>	0.003800
Gilbert-Elliott A	0.000499
HMM11nm A	0.002354

	Variância
<i>Subtrace B</i>	0.002624
Gilbert-Elliott B	0.000911
HMM7nm B	0.002570

Tabela 2. Variância da taxa média de perda

O novo modelo HMMnm proposto apresentou algumas propriedades muito úteis para a representação estocástica do processo de perda de pacotes, a saber:

- O modelo foi avaliado sob estatísticas de 1ª e 2ª ordem, atendendo a todas.
- A estrutura do modelo é simples e não precisa ser alterada, facilitando o seu emprego sem a necessidade de uma avaliação estatística muito detalhada.
- Foram obtidos resultados satisfatórios com poucos estados (≤ 11).
- O modelo se mostrou robusto aos valores de inicialização da cadeia, desde que o número de iterações do algoritmo Baum-Welch seja suficientemente alto.

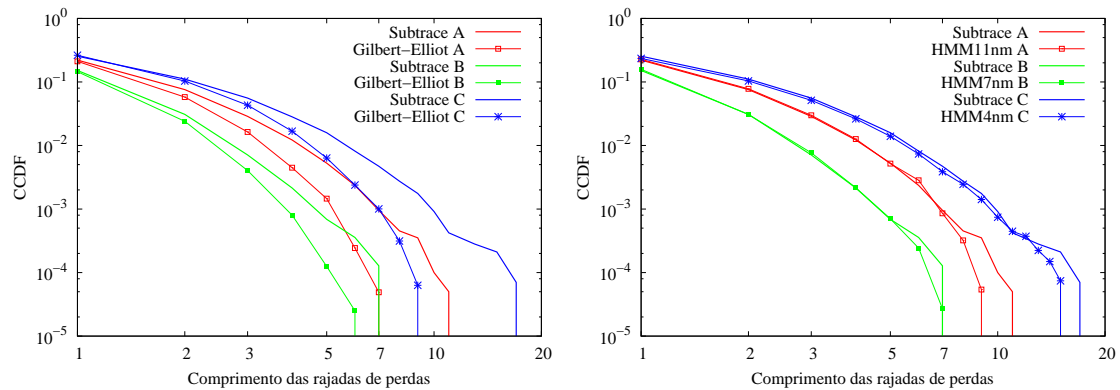


Figura 14. Avaliação do comprimento das rajadas de perda.

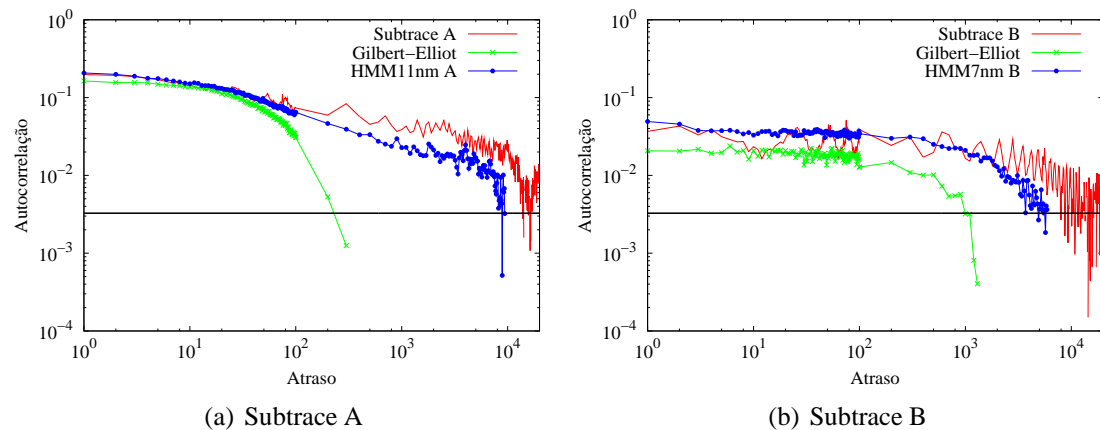


Figura 15. Avaliação da autocorrelação.

7. Conclusão e trabalhos futuros

Nesse trabalho, é realizada uma avaliação detalhada da perda de pacotes em redes 802.11. A perda de pacotes é tratada como um processo aleatório, nos quais as principais propriedades estatísticas são destacadas, a saber: taxa média de perda, variância, distribuição do tamanho das rajadas de perda e a autocorrelação do processo. Embora a taxa média de perda seja a característica mais avaliada, é recomendado que estas três propriedades adicionais não sejam negligenciadas, sob pena da obtenção de resultados errôneos.

São mostrados os principais tipos de modelos usados na representação de perda de pacotes em redes sem fio, com suas vantagens e desvantagens. Em especial é mostrada a inadequação do clássico modelo de Gilbert-Elliott na representação de perda em redes 802.11, obtendo alguns resultados comparativos com uma cadeia de Markov de ordem 4.

O modelo de Gilbert-Elliot captura a taxa média de perda de pacotes, mas não descreve adequadamente a variância, a distribuição do comprimento das rajadas e a autocorrelação.

Um novo modelo HMM com estrutura do tipo nascimento-e-morte foi proposto e avaliado, apresentando resultado muito superiores ao modelo Gilbert-Elliot e à cadeia de Markov de ordem 4. O modelo possui pequeno número de estados, uma estrutura simples e é robusto em relação aos parâmetros de inicialização da cadeia. O modelo HMMnm representou satisfatoriamente as estatísticas de 1ª e 2ª ordem.

Como trabalho futuro se pretende avaliar a possibilidade de uso do modelo como parte de um novo algoritmo para controle automático de taxa para redes 802.11. Planeja-se avaliar também outras estruturas especializadas com potencial para representação de perda de pacotes, por exemplo, a estrutura *Branch-Erlangian*, a qual poderia ser apropriada para lidar com mudanças abruptas na taxa de perda.

Referências

- Aguayo, D., Bicket, J., Biswas, S., Judd, G., and Morris, R. (2004). Link-level Measurements from an 802.11b Mesh Network. In *ACM Special Interest Group on Data Communication (SIGCOMM)*.
- Aráuz, J. and Krishnamurthy, P. (2003). Markov Modeling of 802.11 Channels. In *IEEE Vehicular Technology Conference (VTC)*.
- Bai, H. and Atiquzzaman, M. (2003). Error Modeling Schemes for Fading Channels in Wireless Communications: A Survey. *IEEE Communications Surveys and Tutorials*.
- Bendat, J. S. and Piersol, A. G. (1986). *Random Data: Analysis and Measurement Procedures*. John Wiley and Sons, Inc.
- Bicket, J. C. (2005). Bit-rate Selection in Wireless Networks. Master's thesis, Dept. of Electrical Engineering and Computer Science - Massachusetts Institute of Technology.
- Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time Series Analysis - Forecasting and Control*. Prentice Hall, Inc.
- Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer.
- Cardoso, K. V. and de Rezende, J. F. (2007). Análise e Modelagem de Perda de Pacotes em Redes 802.11 em Ambientes Indoor. In *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*.
- Carvalho, L., Angeja, J., and Navarro, A. (2005). A New Packet Loss Model of the IEEE 802.11g Wireless Network for Multimedia Communications. *IEEE Transactions on Consumer Electronics*.
- Elliot, E. O. (1963). Estimates of Error Rates for Codes on Burst Noise Channels. *Bell System Technical Journal*.
- François, J.-M. (2007). Java implementation of Hidden Markov Model. <http://www.run.montefiore.ulg.ac.be/~francois/software/jahmm/>. [Último acesso: 19-Fevereiro-2007].
- Gaertner, G. and Cahill, V. (2004). Understanding Link Quality in 802.11 Mobile Ad Hoc Networks. *IEEE Internet Computing*.

- Gilbert, E. N. (1960). Capacity of a Burst-Noise Channel. *Bell System Technical Journal*.
- Ji, P., Liu, B., Towsley, D., Ge, Z., and Kurose, J. (2004). Modeling Frame-level Errors in GSM Wireless Channels. *Performance Evaluation*.
- Karande, S., Khayam, S. A., Krappel, M., and Radha, H. (2003). Analysis and Modeling of Errors at the 802.11b Link Layer. In *IEEE International Conference on Multimedia & Expo (ICME)*.
- Khayam, S. A. (2003). Analysis and Modeling of Errors and Losses Over 802.11b Networks. Master's thesis, Department of Electrical and Computer Engineering - Michigan State University.
- Khayam, S. A. and Radha, H. (2003). Markov-based Modeling of Wireless Local Area Networks. In *ACM International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*.
- Konrad, A., Zhao, B. Y., Joseph, A. D., and Ludwig, R. (2003). A Markov-based Channel Model Algorithm for Wireless Networks. *Wireless Networks*.
- LAND (2006). Tangram-II Traffic Generator. <http://www.land.ufrj.br/tools/tangram2/traffic/traffic.html>. [Último acesso: 04-Novembro-2006].
- Lo, P. and Ngai, S. (2004). Characterising Errors in Wireless LANs. Technical report, School of Electrical Engineering & Telecommunications - University of New South Wales.
- Nguyen, G. T., Katz, R. H., Noble, B., and Satyanarayanan, M. (1996). A Trace-Based Approach for Modeling Wireless Channel Behavior. In *Winter Simulation Conference (WSC)*.
- OpenWrt (2006). OpenWrt. <http://openwrt.org>. [Último acesso: 20-Dezembro-2006].
- Rappaport, T. S. (2001). *Wireless Communications: Principles and Practice*. Prentice Hall, Inc.
- Salamatian, K. and Vaton, S. (2001). Hidden Markov Modeling for Network Communication Channels. In *ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*.
- Sanneck, H. and Carle, G. (2000). A Framework Model for Packet Loss Metrics Based on Loss Runlengths. In *SPIE/ACM SIGMM Multimedia Computing and Networking Conference (MMCN)*.
- Tariq, M. M. B., Dhamdhare, A., Dovrolis, C., and Ammar, M. (2005). Poisson versus periodic path probing (or, does PASTA matter?). In *USENIX/ACM Internet Measurement Conference (IMC)*.
- Trivedi, K. S. (2002). *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*. John Wiley and Sons, Inc.
- Yajnik, M., Moon, S., Kurose, J., and Towsley, D. (1999). Measurement and Modelling of the Temporal Dependence in Packet Loss. In *IEEE International Conference on Computer Communications (INFOCOM)*.