

Caracterização de Perda de Pacotes em Redes Wi-Fi com Modelos Ocultos de Markov em Camadas 1 e 2

Carlos Alexandre Gouvea da Silva, Edson Leonardo dos Santos e Carlos Marcelo Pedroso

Resumo— Em redes Wi-Fi, as perdas de pacotes podem ocorrer por diversas circunstâncias e são na maioria das vezes caracterizadas pela ocorrência em rajadas. Para modelar essas perdas são aplicados cadeias de Markov, como o tradicional modelo de Gilbert-Elliot que foi projetado para capturar a dependência temporal observada nas séries de perdas. Os modelos de perdas são desenvolvidos para capturar características de dependência temporal, mas habitualmente considerando os fenômenos observáveis na camada física. Contudo, observa-se que o controle de acesso na camada de enlace também tem grande influência nas perdas de pacotes. Neste artigo é proposto um novo modelo de perdas de pacotes em rajadas para redes IEEE 802.11b/g/n a partir do agrupamento das características de SNR e ocupação do canal usando modelos ocultos de Markov (HMM, *Hidden Markov Model*). Resultados mostram que o uso da ocupação do canal é fundamental para modelagem de perdas de pacotes. Ao considerar este fator, o modelo proposto apresenta melhor aderência à situação real de operação da rede e pode ser utilizado para aprimorar simulações computacionais e também para o desenvolvimento de estratégias de recuperação de erros.

Palavras-Chave— Modelo de Gilbert-Elliot, HMM, IEEE 802.11, Modelos de Perdas de Pacotes.

Abstract— In Wi-Fi networks, packet losses can occur due to different causes and are most often characterized by the occurrence of bursts. Markov chains are applied to model these losses, such as the traditional Gilbert-Elliot model that was designed to capture the temporal dependency of losses. Loss models are developed to capture temporal dependency characteristics, but usually considering the phenomena observable in the physical layer. However, it is observed that the access control in the link layer also has a great influence on packet losses. In this paper, a new packet loss model for IEEE 802.11b/g/n networks is proposed based on the grouping of SNR characteristics and channel occupation using Hidden Markov Model (HMM). Results show that the use of channel occupation is fundamental for modeling packet losses. When considering this factor, the proposed model presents better adherence to the real network operation situation and can be used to improve computational simulations and also to develop error recovery strategies.

Keywords— Gilbert-Elliot Model, HMM, IEEE 802.11, Packet Loss Models.

I. INTRODUÇÃO

Estimativas recentes mostram que até o final de 2023 o número de pontos de acesso à Internet (*hotspots*) conectados por redes Wi-Fi aumentará consideravelmente de 169 milhões em 2018 para 628 milhões. Esse aumento também é impulsionado devido a estimativa de crescimento no número de usuários conectados à Internet de 3,9 bilhões em 2018 para 5,3 bilhões em 2023 [1]. As redes Wi-Fi são baseadas no

padrão IEEE 802.11, a partir das diversas versões atualizadas ao longo dos anos e que ficaram populares devido a facilidade de instalação, utilização, flexibilidade de acesso e mobilidade dos usuários [2].

A qualidade das redes Wi-Fi pode ser avaliada por diversos parâmetros como o atraso fim-a-fim, *jitter*, *throughput* e a taxa perdas de pacotes (PLR, *packet loss rate*). Em redes Wi-Fi, as perdas de pacotes ocorrem quando um ou mais pacote não consegue alcançar o nó destino. Essas perdas ocorrem por diversas razões e podem ser classificadas em três tipos gerais: perdas de camada física devido à problemas no canal de transmissão; por colisões devido a competição pelo acesso ao canal; ou devido a baixa ou excessiva capacidade de memória nos dispositivos, respectivamente chamados de *buffer overflow* e *bufferbloat* [3].

As perdas de pacotes ocorridas em redes Wi-Fi podem ser representadas por modelos de perdas. Usualmente, são utilizados modelos baseados em cadeias de Markov para representar esse comportamento, como por exemplo o tradicional modelo de Gilbert-Elliot (GE) [4][5]. O modelo de GE é baseado em uma cadeia de Markov de dois estados com probabilidade de perda, respectivamente dado por $1 - h$ e $1 - k$. A transição entre os estados ocorre com probabilidade p e q [3]. Contudo, o modelo de GE não é o melhor modelo para representação de perdas de pacotes em redes Wi-Fi devido a existência de autocorrelação de longa duração entre as rajadas de perdas [6][7][8].

Os modelos ocultos de Markov (HMM, *Hidden Markov Model*) vem sendo utilizados como uma alternativa para a modelagem de perdas em redes Wi-Fi [9]. O HMM é um modelo estocástico formado por duas partes, no qual a primeira parte é um processo estocástico em que não é possível observar diretamente o atual estado do sistema, sendo que este processo apenas pode ser estimado e caracterizado através da observação de um segundo processo, também estocástico. Devido a natureza dos diversos parâmetros observáveis em redes Wi-Fi, como o SNR (*Signal Noise Ratio*) e a ocupação do canal, o HMM se apresenta como um potencial modelo na representação das perdas de pacotes. Em redes sem fio os modelos de perdas de pacotes disponíveis consideram de forma independente parâmetros de camada PHY como o SNR e parâmetros de camada MAC. Em outros casos os modelos são apenas caracterizados por modelos analíticos de séries temporais desvinculados de informações sobre o estado da rede.

Propomos neste artigo a aplicação de HMM para modelar as perdas de pacotes em redes IEEE 802.11b/g/n considerando como sequência dos dados observáveis o SNR e a ocupação do canal. O SNR é uma das principais características de

camada PHY que afeta a perda de pacotes. A ocupação de canal indica o grau de concorrência na camada MAC que pode ocasionar perdas devido a disputa do canal pelo diversos usuários conectados no *access point* (AP) e possíveis terminais ocultos. A principal contribuição deste trabalho é a utilização desses dois parâmetros de camada PHY e MAC de forma combinada obtendo-se uma representação mais precisa do comportamento real de perdas de pacotes em redes Wi-Fi. Outros parâmetros também foram considerados para uso, porém o SNR e ocupação se mostraram como fatores dominantes nas perdas para o modelo proposto. Essa abordagem é inédita e não foi utilizada nos modelos de perdas de pacotes disponíveis na literatura.

Após esta seção introdutória, a seção II apresenta os trabalhos relacionados. Os procedimentos de coleta das amostras e descrição da metodologia utilizada na identificação e testes do modelo são apresentados na seção III. Os resultados e discussões são apresentados na seção IV, seguida da conclusão na seção V.

II. TRABALHOS RELACIONADOS

Wang e Moayeri [10] propõe o uso de um modelo de estados finitos (FSMC, *Finite State Markov Chain*) para modelagem do canal. Os autores mostram que o modelo de Gilbert-Elliot é um caso particular deste modelo, onde cada estado corresponde a uma faixa de valores do SNR do canal. Em cada estado é usado o desvanecimento Rayleigh parametrizado de acordo com o comportamento do estado. Esse modelo permite analisar o efeito do desvanecimento e a capacidade do canal investigado. E ainda, o efeito de particionamento do SNR tem impacto direto no modelo FSMC e na probabilidade de perdas. Uma solução para esse problema do particionamento do SNR é proposto por [11], onde para cada faixa de SNR é aplicado a um estado diferente onde é analisado a perda de pacotes e o desempenho do atraso em redes sem fio. Resultados mostraram que o número de estados do modelo Markoviano para um FSMC depende do esquema de modulação de cada canal, da média do SNR e de um limite para a taxa de erro de bit.

Em [12] é proposto um modelo chamado 4SMM (*Four-state Markov Model*) no qual apresenta um melhor desempenho em relação à modelos de dois estados como o 2SMM. Resultados de simulações comparados a medições reais coletadas em uma rede 802.11 se mostraram mais próximas nos modelos com quatro estados. O 4SMM utiliza distribuições de probabilidade para caracterizar a duração de estados Bons e Ruins relativos à intervalos de valores do SNR decorrente do desvanecimento do canal de transmissão sem fio.

Cardoso e Rezende [7] propõem um modelo baseado no modelo de Gilbert-Elliot de três estados com duas estruturas diferentes: primeiro, uma geral com transição entre todos os pares de estados (HMM3g); e por último uma *birth-death* com transições somente entre estados adjacentes (HMM3bd). Resultados da aplicação dessas duas estruturas mostram que o modelo geral HMM3g apresenta pequenas melhorias, enquanto que no modelo HMM3bd a função de autocorrelação (ACF, *Autocorrelation Function*) se aproximou mais da ACF dos dados coletados de um ambiente real. Comparado ao

modelo de Gilbert-Elliot, o HMM3bd também apresentou melhores resultados em relação a ACF e a função de distribuição acumulada complementar (CCDF, *Complementary Cumulative Distribution Function*).

Gomes *et. al.* [13] propõe o uso do HMM na caracterização de perdas de pacotes em redes Wi-Fi. Os autores utilizam 15 estados e um conjunto de observações composto por throughput, taxa de perda de pacotes, taxa de perda de quadros e taxa de erros de bit para inferir o estado atual. A partir dos resultados obtidos, os autores indicam que o HMM é capaz de refletir de maneira confiável a característica de erros e perdas em rajada avaliada em um cenário interno de uma rede sem fio real. Para ambientes *indoor* utilizando o padrão 802.11, em [14] os autores comparam a utilização de um modelo de cinco estados do FSMC. Resultados comprovam que modelos de alta ordem como o HMM de cinco estados (HMM5) promovem melhores aproximações ao comportamento real do que modelos tradicionais de 2 estados como o modelo de Gilbert ou Gilbert-Elliot.

III. NOVO MODELO DE PERDAS

Nesta seção é apresentado um novo modelo de perdas para redes Wi-Fi considerando o SNR e a ocupação do canal. Foram realizadas coletas de dados para possibilitar a parametrização e validação do modelo.

A. Cenário de Coleta das Amostras

Foi utilizado um AP para realização das coletas de dados na rede Wi-Fi em estudo. O AP foi configurado em modo infraestrutura com protocolo IEEE 802.11b/g/n em modo automático. O número do canal utilizado foi configurado de maneira fixa para diferentes coletas.

No cenário utilizado para coleta das amostras haviam diversos usuários conectados ao AP principal e também em outros APs adjacentes e configurados no mesmo canal, como mostrado na Figura 1.

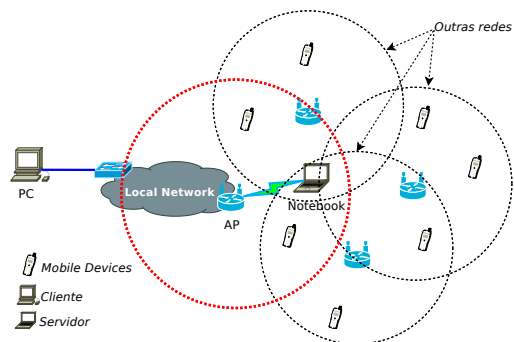


Fig. 1. Cenário utilizado para coleta de dados em rede sem fio IEEE 802.11.

Os testes foram feitos no departamento de engenharia elétrica da UFPR que possui aproximadamente 1200 alunos e 50 professores. A existência de diversos APs que compartilham do mesmo meio de transmissão aumenta a utilização do canal, a probabilidade de colisões e interferências na transmissão, ocasionando perdas de pacotes que são cruciais para análise desse comportamento no modelo proposto.

Foi criada uma aplicação Cliente/Servidor sobre o protocolo UDP, no qual um computador conectado à rede Ethernet enviou sequencialmente pacotes de 162 bytes a cada 1 milissegundo para um Notebook conectado ao AP. Durante a coleta dos parâmetros de qualidade da rede e dos pacotes transmitidos, o receptor foi movido para diferentes distâncias em relação ao AP. Foram realizadas 6 horas e 50 minutos de transmissões no qual foram coletados simultaneamente as sequências de chegadas de pacotes ocorridas, além do SNR (dB) e ocupação do canal (%). O SNR foi medido no Notebook e a ocupação do canal por um analisador de redes Fluke Networks AirCheck Wi-Fi Tester. As coletas foram realizadas em dias e horários variados de maneira a capturar diferentes situações de comportamento da rede em estudo. As redes onde foram realizadas as coletas apresentavam diversos níveis de ocupação, apresentando sempre uma quantidade variada de usuários conectados no AP ou em outros APs de mesmo canal de transmissão.

B. Metodologia de Análise

As amostras coletadas foram separadas em conjuntos com a sequência de pacotes recebidos e perdidos, representados respectivamente por 0's e 1's, sendo que para cada intervalo de 10 segundos de amostragem foram registrados os valores médios do SNR e da ocupação do canal. Supomos que os diversos estados de perda não são fenômenos observáveis diretamente. Desta forma, é observado o SNR e a utilização do canal, e a partir destes fenômenos estima-se o estado de perda. O primeiro problema é determinar o número ideal de estados.

A definição do número de estados ideal para treinamento do HMM foi realizada utilizando algoritmos de clusterização de dados. O objetivo da clusterização é agrupar um conjunto de objetos de dados em vários grupos (clusters), para que os objetos dentro do mesmo cluster sejam semelhantes entre si e que tenham grande dissimilaridade com os objetos de outros clusters [15]. Neste artigo foi aplicado o método da distância Euclidiana já implementado em uma biblioteca disponível no software R[16]. Esse método é amplamente utilizado para clusterização de dados, no qual a partir de um ponto fixo é possível medir a distância para os demais pontos disponíveis na série de dados. A apresentação desses agrupamentos é realizada através de gráficos do tipo dendrogramas.

A Figura 2 apresenta o dendrograma resultante da clusterização do SNR e ocupação do canal. Visualmente é possível verificar a existência de 4 grupos de similaridade dos dados de ocupação e SNR. Neste artigo assumimos que o número ideal de 4 estados utilizados no treinamento do HMM foi baseado no agrupamento realizado pela clusterização (ou seja, HMM4). Contudo, também foram realizados testes para 3 e 5 estados, que apresentaram resultado inferiores em relação ao HMM4.

A partir dos dados de SNR e ocupação do canal foi realizado procedimento treinamento de um HMM de forma a determinar o estado mais provável a partir do SNR e ocupação, bem como as probabilidades de transição de estado. Para o treinamento foi utilizado um conjunto de bibliotecas (*depmixS4*, *quantmod* e *qcc*) disponíveis no software *open source* R

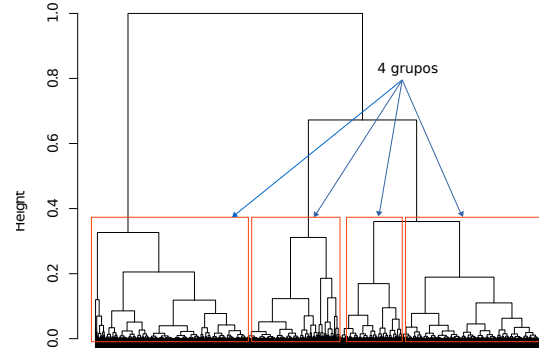


Fig. 2. Dendrograma da clusterização com distância Euclidiana.

[16]. Durante o treinamento, o HMM ajusta os parâmetros do modelo $\varphi = (\mathbf{A}, \mathbf{B}, \pi)$ baseado na sequência de treinamento utilizando o método iterativo de Baum-Welch [17], no qual a matriz \mathbf{A} é a distribuição de probabilidade da transição de estados, a matriz \mathbf{B} é a distribuição de probabilidade da sequência de treinamento, e o vetor π é a distribuição do estado inicial do modelo. Esse processo de ajuste dos parâmetros do modelo φ é realizado de forma a satisfazer um certo critério de otimização. O critério utilizado para definir qual o melhor treinamento considerou o maior valor de máxima verossimilhança. Como resultado do treinamento, o HMM4 apresentou uma sequência provável do número do estado relacionado aos dados de SNR e ocupação. A dispersão desses pontos de SNR e ocupação em relação aos estados definidos pelo HMM4 é apresentada na Figura 3.

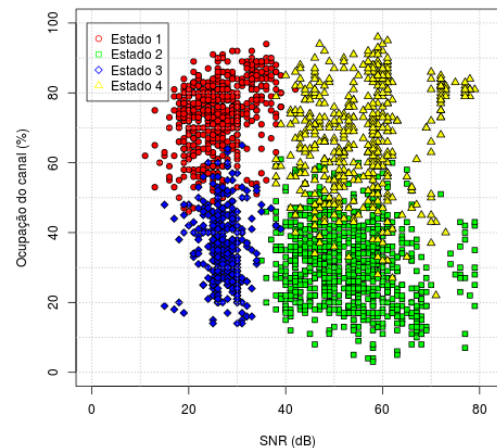


Fig. 3. Estados treinados a partir do HMM4.

Observa-se que o treinamento foi capaz de identificar os quatro estados a partir dos parâmetros observáveis. É possível verificar que o Estado 4 possui valores altos de SNR, com a perda de pacotes dominada pelo alto nível de utilização. No estado 3 a perda é dominada pelo SNR, com baixo nível de ocupação do canal. O Estado 2 é o melhor estado de operação, com maior SNR e menor utilização. O Estado 1 é o pior estado, com maior utilização e menor SNR, no qual neste estado espera-se os maiores níveis de perda de pacotes.

As sequências de 0's e 1's correspondentes a cada um dos 4 estados foram separadas para calcular o tamanho das rajadas.

O tamanho das rajadas foi calculado pela quantidade de perdas ocorridas de forma sequencial até que o próximo pacote tenha sido recebido. Observou-se que o tamanho das rajadas de perda apresentou comportamento de cauda pesada [8]. Cauda pesada é um comportamento em que a forma assintótica da distribuição é hiperbólica.

Para cada estado o tamanho das rajadas foram submetidos ao teste de aderência para diferentes distribuições de probabilidades comumente utilizadas em séries de dados que apresentam comportamento de cauda pesada: Weibull, Pareto tipo II, Log-Normal, Cauchy e Log-Cauchy. A estimação dos parâmetros de cada distribuição foi realizada utilizando algoritmos disponíveis no software R [16].

Testes de aderência mostraram que a distribuição acumulada das rajadas de perdas de cada estado é aproximadamente melhor representada por uma distribuição Pareto Tipo II (distribuição de Lomax). Essa distribuição utiliza a abreviação $X \sim \text{Lomax}(\alpha, \lambda)$ para indicar que a variável aleatória X tem a distribuição Lomax com os parâmetros λ e α . Uma variável aleatória Lomax X com parâmetro de escala λ e parâmetro de forma α tem função de densidade probabilidade:

$$f(x) = \frac{\lambda\alpha}{(1 - \lambda x)^{\alpha+1}} \quad \text{com } x > 0, \quad (1)$$

onde, $\alpha > 0$ e $\lambda > 0$.

Contudo, nas amostras do estado 1 e 2 a distribuição de Pareto Tipo II não conseguiu representar corretamente todo esse comportamento presente no tamanho das rajadas, sendo necessário uma combinação entre esta e uma distribuição exponencial com função de densidade probabilidade:

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \quad \text{com } x > 0, \quad (2)$$

onde, μ é a média.

A Tabela I apresenta os valores estimados dos parâmetros α , λ e μ das distribuições de probabilidade para cada estado. Como os estados 1 e 2 foram modelados com duas distribuições de probabilidade, o tamanho das rajadas de perdas igual ou superior a 400 e 40 foram modeladas com uma distribuição exponencial, respectivamente. Os estados 3 e 4 foram modeladas com apenas a distribuição de Pareto Tipo II.

TABELA I
CARACTERÍSTICAS DOS 4 ESTADOS MAPEADOS.

Estado	Função de Probabilidade	Parâmetros estimados
1	Pareto Tipo II	$\alpha = 3,21$ e $\lambda = 12,32$
	Exponencial	$\mu = 1682,61$
2	Pareto Tipo II	$\alpha = 3,19$ e $\lambda = 6,33$
	Exponencial	$\mu = 44,36$
3	Pareto Tipo II	$\alpha = 3,42$ e $\lambda = 7,23$
4	Pareto Tipo II	$\alpha = 2,07$ e $\lambda = 4,94$

A validação do modelo foi realizada a partir de um conjunto de dados e da utilização do algoritmo de Viterbi [18] disponível na biblioteca *depmixS4* no software R [16]. Esse conjunto de dados foi coletado em pouco mais de uma hora de coleta. O algoritmo de Viterbi verifica qual o estado mais provável a partir dos dados observados, confrontado com o HMM4 já treinado. Os tamanhos médios e desvios padrão das rajadas em

cada estado das amostras treinadas no HMM4 são comparadas com o tamanho das rajadas do conjunto de validação e também do simulado. Por fim, a distribuição acumulada do tamanho da rajada de perdas da amostra geral é comparada dados simulados utilizando o modelo proposto.

IV. ANÁLISE DE DESEMPENHO

A Tabela II apresenta as características dos estados analisados após o treinamento. Nestes quatro estados é possível identificar que o pior (Ruim) estado é representado pelo estado 1, pois ele apresenta o maior percentual médio de perdas de 60,97%. Esse percentual alto de perdas é justificado ao observar a Figura 3 no qual os pontos referentes ao estado 1 estão concentrados nos menores níveis de SNR e maior taxa de ocupação. O estado 2 é considerado o melhor (Bom) estado no qual concentra-se o comportamento da rede quando os níveis de SNR são altos e a ocupação é baixa. Neste caso, o estado Bom possui a menor taxa média de perdas. Os demais estados são considerados intermediários. Observa-se que o estado 4 apresenta considerável taxa média de perdas mesmo com níveis de SNR considerados ideais para uma rede Wi-Fi, porém essa taxa é influenciada pela ocupação do canal. A coluna amostra de validação apresenta as taxas de perda de um novo conjunto de dados que é utilizado para comparar com as taxas de perdas da amostra de treinamento. Essa comparação é feita através da aplicação do algoritmo de Viterbi, no qual ele estima a partir da amostra de validação qual é o estado provável usando o HMM4 previamente treinado. Neste teste, o algoritmo de Viterbi não conseguiu identificar o estado 3 (N/A) devido ao fato de que na amostra de validação não haviam níveis de SNR e ocupação do canal que pudessem ser classificados como estado 3. Para os estados 2 e 4 observa-se uma proximidade dessas taxas do treinamento e da validação. O estado 1 apresentou alta taxa de perdas da amostra de validação, contudo bem abaixo do identificado na amostra de treinamento, sendo sugerido a coleta de novos conjuntos de validação para uma análise mais profunda.

TABELA II
CARACTERÍSTICAS DE PERDAS DOS 4 ESTADOS MAPEADOS.

Estado	Condição	Treinamento HMM	Amostra Validação
1	Ruim	60,97%	14,81%
2	Bom	0,55%	0,59%
3	Intermediário1	2,02%	N/A
4	Intermediário2	12,78%	9,10%

O treinamento do HMM4 resultou na matriz \mathbf{A} de transição entre os quatro estados a partir da sequência de treinamento, como mostrado a seguir.

$$\mathbf{A} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 0,955 & 0,014 & 0,017 & 0,014 \\ 0,002 & 0,959 & 0,013 & 0,027 \\ 0,051 & 0,017 & 0,932 & 0,000 \\ 0,014 & 0,061 & 0,000 & 0,925 \end{pmatrix} \end{matrix}$$

A matriz mostrou que a probabilidade de permanecer no mesmo estado é superior a 95% para os estados 1 e 2, sendo que nestes estados existe a possibilidade de transição para

todos os demais. Ao verificar os estados intermediários 3 e 4, a probabilidade de permanecer no mesmo estado é superior a 90%, porém não há possibilidade de transição entre estes dois estados, apenas para 1 e 2. O vetor do estado inicial é igual a $\pi = [0 \ 0 \ 0 \ 1]$, no qual para a sequência de treinamento utilizada o estado inicial proposto pelo HMM4 é o estado 4.

A Figura 4 apresenta o *quantile-quantile* Plot (qqPlot) do tamanho das rajadas simuladas em relação aos dados de tamanho das rajadas coletados na rede. O primeiro qqPlot mostra que os dados utilizados na validação do modelo apresenta aderência à distribuição de probabilidade (vide Tabela I) utilizada no modelo proposto para as rajadas inferiores a 500. O segundo qqPlot indica que a amostra coletada para treinamento do HMM4 apresenta boa aderência aos dados simulados baseado na distribuição proposta neste modelo.

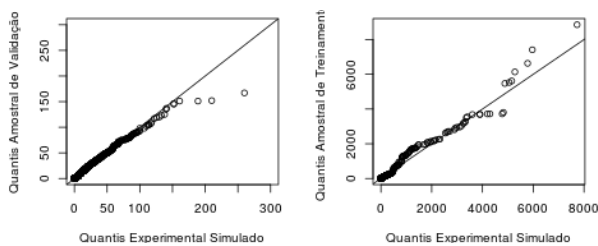


Fig. 4. qqPlot: simulado vs amostra de validação e simulado vs amostra coletada considerando o tamanho da rajada de perda.

A Figura 5 apresenta um comparativo entre a distribuição acumulada do tamanho da rajada de perdas das amostras coletadas com outros modelos testados. Foram testados dois modelos, o GE parametrizado com a sequência de perdas observadas na amostra original no qual $p = 0,0393$ e $q = 0,1862$, e o FSMC com 4 estados baseado na matriz de transição obtido do treinamento do HMM4. Como observado a CDF dos modelos de GE e 4FSMC não conseguiram representar corretamente o comportamento do tamanho das rajadas, em especial a curva assintótica (cauda pesada) presente nessa amostra. Dados simulados da distribuição proposta no modelo HMM4 apresentou melhor aderência à CDF.

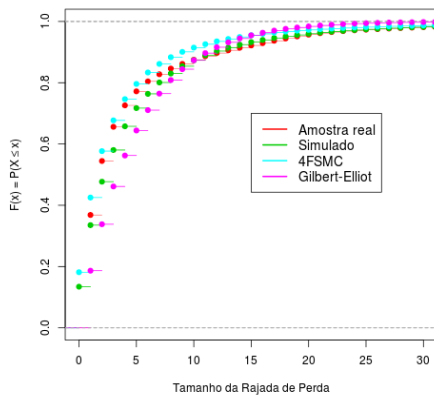


Fig. 5. CDF do tamanho das rajadas de perdas.

V. CONCLUSÃO

Foi proposto neste trabalho um novo modelo de perdas de pacotes em redes Wi-Fi baseado em HMM, no qual os

dados observáveis de SNR e ocupação do canal são utilizados simultaneamente. A partir do treinamento do HMM de 4 estados foi possível identificar o melhor, o pior e os estados intermediários. As rajadas de perdas dos estados intermediários 3 e 4 foram modelados unicamente com uma distribuição de Pareto Tipo II, com diferentes valores dos parâmetros λ e α . O melhor e pior estados, respectivamente 2 e 1, foram modelados com uma combinação de distribuições, no qual foi usado a distribuição de Pareto Tipo II e a distribuição Exponencial. Os modelos disponíveis atualmente não são capazes de capturar corretamente o comportamento das perdas, pois elas dependem fortemente do SNR e da ocupação do canal de transmissão. O modelo proposto é capaz de utilizar estas duas variáveis observáveis para prever de forma mais precisa o comportamento de perdas de pacotes em redes Wi-Fi.

REFERÊNCIAS

- [1] Cisco, “Cisco Annual Internet Report (2018–2023),” *White Paper*, 2020.
- [2] M. S. Afaqui, E. Garcia-Villegas, and E. Lopez-Aguilera, “IEEE 802.11ax: Challenges and Requirements for Future High Efficiency WiFi,” *IEEE Wirel. Commun.*, vol. 24, no. 3, pp. 130–137, Jun. 2017.
- [3] C. A. G. Silva and C. M. Pedroso, “MAC-Layer Packet Loss Models for Wi-Fi Networks: A Survey,” *IEEE Access*, vol. 7, pp. 180 512–180 531, 2019.
- [4] E. N. Gilbert, “Capacity of a Burst-noise Channel,” *The Bell System Techn. Journal*, vol. 39, no. 5, pp. 1253–1265, Sep. 1960.
- [5] E. O. Elliott, “Estimates of Error Rates for Codes on Burst-noise Channels,” *The Bell System Techn. Jour.*, vol. 42, no. 5, pp. 1977–1997, Sep. 1963.
- [6] L. Carvalho, J. Angeja, and A. Navarro, “A New Packet Loss Model of the IEEE 802.11g Wireless Network for Multimedia Communications,” *IEEE Trans. on Cons. Elect.*, vol. 51, no. 3, pp. 809–814, Aug. 2005.
- [7] K. V. Cardoso and J. F. D. Rezende, “Accurate hidden markov modeling of packet losses in indoor 802.11 networks,” *IEEE Commun. Lett.*, vol. 13, no. 6, pp. 417–419, Jun. 2009.
- [8] L. J. Rohling, C. A. G. Silva, E. M. G. Fernandez, and C. M. Pedroso, “Evidências da Falha do Modelo de Gilbert-Elliott em Redes sem Fio,” in *XXXIV Simpósio Brasileiro de Telecomunicações - SBt2016*, Sep. 2016, pp. 488–492.
- [9] A. Allahdadi and R. Morla, “Anomaly Detection and Modeling in 802.11 Wireless Networks,” *Journal of Network and Systems Management*, vol. 27, no. 1, pp. 3–38, 2019.
- [10] H. S. Wang and N. Moayeri, “Finite-state Markov Channel—a Useful Model for Radio Communication Channels,” *IEEE Transactions on Vehicular Technology*, vol. 44, no. 1, pp. 163–171, Feb. 1995.
- [11] M. Hassan, M. M. Krunz, and I. Matta, “Markov-based Channel Characterization for Tractable Performance analysis in Wireless Packet Networks,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 3, pp. 821–831, May 2004.
- [12] Y. Yu and S. L. Miller, “A Four-State Markov Frame Error Model for the Wireless Physical Layer,” in *2007 IEEE Wireless Communications and Networking Conference*, Mar. 2007, pp. 2053–2057.
- [13] D. Gómez, R. Agüero, M. García-Arranz, and L. Muñoz, “On The Modeling of a Realistic Wireless Channel By Means of a Hidden Markov Process,” in *8th Intern. Confer. on Wirel. and Mobile Comput., Networ. and Communic.*, Oct. 2012, pp. 397–402.
- [14] J. A. Hartwell and A. O. Fapojuwo, “Modeling and Characterization of Frame Loss Process in IEEE 802.11 Wireless Local Area Networks,” in *Proc. IEEE 60th Vehic. Techn. Conf.*, vol. 6, Sep. 2004, pp. 4481–4485.
- [15] Y. Li, J. Cai, H. Yang, J. Zhang, and X. Zhao, “A Novel Algorithm for Initial Cluster Center Selection,” *IEEE Access*, vol. 7, pp. 74 683–74 693, Jun. 2019.
- [16] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2019. [Online]. Available: <https://www.R-project.org/>
- [17] S. Adams, P. A. Beling, and R. Cogill, “Feature Selection for Hidden Markov Models and Hidden Semi-Markov Models,” *IEEE Access*, vol. 4, pp. 1642–1657, 2016.
- [18] H.-L. Lou, “Implementing the Viterbi Algorithm,” *IEEE Signal Processing Magazine*, vol. 12, no. 5, pp. 42–52, 1995.