

UNIVERSIDADE FEDERAL DO PARANÁ

CARLOS ALEXANDRE GOUVEA DA SILVA

NOVO MODELO DE PERDAS DE PACOTES PARA REDES WIFI

CURITIBA

2020

CARLOS ALEXANDRE GOUVEA DA SILVA

NOVO MODELO DE PERDAS DE PACOTES PARA REDES WIFI

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, Área de Concentração Telecomunicações, Departamento de Engenharia Elétrica, Setor de Tecnologia, Universidade Federal do Paraná, como requisito parcial para obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Prof. Dr. Carlos Marcelo Pedroso

CURITIBA

2020

Catálogo na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

S586n

Silva, Carlos Alexandre Gouvea da
Novo modelo de perdas de pacotes para redes WiFi [recurso eletrônico] /
Carlos Alexandre Gouvea da Silva. – Curitiba, 2020.

Tese - Universidade Federal do Paraná, Setor de Tecnologia, Programa
de Pós-Graduação em Engenharia Elétrica, 2020.

Orientador: Carlos Marcelo Pedroso

1. Redes locais sem fio. 2. Sistemas de comunicação sem fio. 3. Markov,
Modelo de. 4. IEEE 802.11 (Padrão). 5. Sistemas WiFi. I. Universidade
Federal do Paraná. II. Pedroso, Carlos Marcelo. III. Título.

CDD: 004.6

Bibliotecário: Elias Barbosa da Silva CRB-9/1894

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA ELÉTRICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **CARLOS ALEXANDRE GOUVEA DA SILVA** intitulada: **NOVO MODELO DE PERDAS DE PACOTES PARA REDES WIFI**, sob orientação do Prof. Dr. CARLOS MARCELO PEDROSO, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 21 de Dezembro de 2020.

Assinatura Eletrônica
23/12/2020 13:58:37.0
CARLOS MARCELO PEDROSO
Presidente da Banca Examinadora

Assinatura Eletrônica
23/12/2020 15:14:48.0
JULIO CÉSAR NIEVOLA
Avaliador Externo (PONTIFÍCIA UNIVERSIDADE CATÓLICA DO
PARANÁ)

Assinatura Eletrônica
23/12/2020 14:12:23.0
EVELIO MARTÍN GARCÍA FERNÁNDEZ
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
26/12/2020 10:16:12.0
EMILIO CARLOS GOMES WILLE
Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO
PARANÁ)

Dedico este trabalho a minha querida e amada avó Nahir de Castro Gouvea (in memoriam), que em sua saudade deixou muitas lições e ensinamentos para toda vida.

AGRADECIMENTOS

A toda minha família pelo apoio nos momentos difíceis e turbulentos. Em especial a minha querida e amada mãe Célia por ser a base e grande incentivadora dos meus sonhos e anseios. Aos meus queridos e amados sobrinhos João Victor, Mariana Victória e Mateus pela alegria de todos os dias.

Ao professor e orientador Dr. Carlos Marcelo Pedroso pela paciência e imenso suporte para que esta pesquisa pudesse ser desenvolvida. Esse trabalho apenas foi possível pelo seu conhecimento e estimada orientação.

Aos demais professores do grupo de pesquisa, professores Evélio e Eduardo Parente pelo auxílio e suporte em diversas atividades realizadas durante esses últimos anos.

Ao estimado e grande amigo Edson Leonardo pelo incentivo, apoio, ajuda e exemplo à sempre continuar em frente e nunca desistir. Pelos momentos de aflição e conquistas compartilhados, o meu enorme reconhecimento pela amizade.

Aos colegas e amigos dos laboratórios LABSISCOM e LASICO, que compartilharam das mesmas frustrações e alegrias durante esses últimos anos, em especial Rodrigo Negri, Wesley Rodrigues, Carlos Maffini, Cristiano Osinski, Douglas Pelacini, Allan Ferrari e Mariane Bergamini.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Fundação Araucária (FA) pelo apoio financeiro oferecido durante o período de doutoramento no PPGEE da Universidade Federal do Paraná.

RESUMO

Modelos de perdas de pacotes para redes WiFi são utilizados para representar o comportamento das perdas que são causadas nessas redes por diferentes razões, como interferências, problemas de coexistência com outras redes ou dispositivos, degradação do canal, colisões ou espaço de memória insuficiente ou excessiva. O modelo de Gilbert-Elliot, baseado em uma cadeia de Markov de dois estados, é o modelo mais usado para representar perdas em redes de comunicação. Contudo, diversas pesquisas e trabalhos publicados mostram que o modelo de Gilbert-Elliot é inadequado para representar o real comportamento das perdas de pacotes em redes WiFi. Ainda, os modelos atuais de perdas de pacotes não consideram características observáveis conjuntamente nas camadas física e de enlace. Incluir no modelo de perda características de camada de enlace permite que a acurácia dos modelos de perda seja melhorada. Assim, esta tese apresenta o desenvolvimento de um novo modelo de perdas de pacotes para redes IEEE 802.11b/g/n, baseado em um Modelo Oculto de Markov (HMM, *Hidden Markov Model*) que considera simultaneamente características de camada física e enlace. Na camada física é utilizada a razão sinal-ruído (SNR, *signal noise ratio*) e na camada de enlace é considerado a ocupação do canal de transmissão. Com a aplicação do HMM foi possível identificar o estado ideal da rede baseado no nível da SNR e na ocupação do canal. Resultados mostraram que cada estado obtido do HMM pode ser usado para modelar o comprimento das rajadas de perdas a partir de uma distribuição de cauda pesada e/ou outra exponencial. Ainda, o modelo mostra que considerar apenas características de camada física reduz a acurácia dos modelos porque a ocupação do canal influencia o comportamento das perdas nestas redes.

Palavras-chave: IEEE 802.11. Modelo de Gilbert-Elliot. Modelo Oculto de Markov. Modelo de Perda de Pacotes. WiFi.

ABSTRACT

Packet loss models for WiFi networks are used to represent the behavior of losses caused by several reasons such as interference, problems of coexistence with other networks or devices, channel degradation, competition for media access, and insufficient or excessive memory space. The Gilbert-Elliot model, based on a two-state Markov chain, is the most used model to represent losses in communication networks. However, researches and published works show that the Gilbert-Elliot model is inadequate to represent real behavior of packet loss in WiFi networks. Additionally, there are no available model that jointly consider physical and link layer variables. New packet loss models that consider link layer characteristics increase the accuracy of current models. Thus, this thesis presents the development of a new packet loss model for IEEE 802.11b/g/n networks that simultaneously considers physical and link layer characteristics, based on a Hidden Markov Model (HMM). In the physical layer the signal-to-noise ratio (SNR) is used, and in the link layer the channel occupation is considered in the proposed new model. With the application of HMM was possible to identify the loss state of the WiFi network based on SNR and channel occupation. Results showed that the burst loss length in each state can be modeled using a heavy tail and/or other exponential distribution. Furthermore, the model shows that considering only physical layer characteristics reduces the accuracy of the models, since the channel occupation also had a direct influence on the behavior of losses in these networks.

Key words: Gilbert-Elliot Model. Hidden Markov Model. IEEE 802.11. Packet Loss Model. WiFi.

LISTA DE FIGURAS

2.1	Padrão MAC e PHY para IEEE 802.11 (STALLINGS, 2004).	27
2.2	Arquitetura de conjunto de serviço estendido (SILVA; PEDROSO, 2019).	27
2.3	Mapa de sobreposição de canais em tecnologia sem fio WiFi (SILVA; PEDROSO, 2019).	36
2.4	<i>Handoff</i> vertical e horizontal (SILVA; PEDROSO, 2019).	39
2.5	Cenário de colisão em redes WiFi.	41
3.1	Modelo de Gilbert.	46
3.2	Modelo de Gilbert-Elliot.	46
3.3	Modelo de Gilbert Estendido (YOUNESIAN et al., 2014).	50
3.4	Modelo LRL com número ilimitados de estados ($m \rightarrow \infty$)(SANNECK; CARLE, 1999).	51
3.5	Modelo LRL com limitado número de estados [($m+1$) estados](SANNECK; CARLE, 1999).	52
3.6	Adaptação do modelo de Gilbert Estendido (YOUNESIAN et al., 2014).	53
3.7	Modelo de Gilbert-Elliot com sub-estados (FENG; LIU; JI, 2014).	56
3.8	Modelo hMM de duas camadas (TAO; LU; CHUANG, 2001).	57
3.9	Modelo de cadeias de Markov com estados finitos para um canal de desvanecimento Rayleigh (WANG; MOAYERI, 1995).	58
3.10	Modelo DEPHMM.	60
3.11	Modelo de quatro estados 4SMM (YU; MILLER, 2007).	61
4.1	Sequência de atividades para realização do trabalho.	67
4.2	Cenário utilizado para coleta de dados em rede sem fio IEEE 802.11.	69
4.3	Cenário de coleta das amostras e arquivos de saída.	71
4.4	Dados coletados na interface WiFi do Notebook: (A) SNR, (B) qualidade do sinal, (C) potência do sinal e (D) taxa de transmissão.	73
4.5	Dados coletados no analisador de redes: (A) ocupação do canal e (B) número de usuários conectados no mesmo canal de transmissão.	74
4.6	Dados observados: (A) sequência das chegadas e perdas de pacotes e (B) taxa de perdas por janela de amostragem.	74
4.7	Amostra para teste do atraso e jitter: (A) taxa de perdas, (B) SNR e ocupação do canal, (C) atraso fim-a-fim, e (D) jitter.	80

4.8	Dados observados: (A) ocupação em relação ao número de usuários conectados no canal e (B) taxa de perdas por janela de amostragem. .	84
4.9	Diagrama de transição de estado de um HMM (TOLEDO; KATZ, 2009).	86
4.10	Exemplo de dendrograma com clusterização usando método distância Euclidiana.	91
4.11	Representação linhas de corte em dendrograma gerado a partir da distância Euclidiana.	92
4.12	Séries de dados principais de SNR e ocupação do canal.	92
5.1	Dendrograma da clusterização com distância Euclidiana.	99
5.2	Estados treinados a partir do HMM4.	99
5.3	Série de comprimento das rajadas de perdas por estado.	102
5.4	Gráficos de QQ-Plot para a distribuição de Pareto Tipo II por estado. . .	103
5.5	Gráficos de QQ-Plot para as distribuições de Pareto tipo II e Exponencial dos estados 1 e 2.	104
5.6	QQ-Plot dos quantis das séries obtidas do HMM4 e os simulados por estado.	106
5.7	QQ-Plot da simulação da matriz de transição A com a amostra observada.	106
5.8	LLCD das amostras observada, simulada e do modelo de GE.	108
5.9	ACF do modelo de GE.	109
5.10	QQ-Plot trabalhos relacionados com a amostra real observada.	115

LISTA DE TABELAS

2.1	Especificação histórica dos padrões IEEE 802.11 (ZHU et al., 2004)(AB-DELRAHMAN; MUSTAFA; OSMAN, 2015)	30
2.2	Trabalhos relacionados de interferência e coexistência entre WiFi e outras tecnologias	37
3.1	Comparação dos parâmetros dos modelos de Gilbert simples, Gilbert e Gilbert-Elliot (HASSLINGER; HOHLFELD, 2008).	47
3.2	Comparação modelos de perdas de pacotes em redes sem fio (SILVA; PEDROSO, 2019)	64
4.1	Características de configuração do AP.	70
4.2	Teste de correlação utilizando os métodos Pearson, Kendall e Spearman.	78
4.3	Taxas de transmissão IEEE 802.11n.	81
4.4	Teste de correlação da ocupação do canal com o número de usuários.	83
4.5	Exemplo típico da sequência de observação baseado no modelo proposto.	89
5.1	Informação da média de SNR e ocupação do canal por estado.	100
5.2	Informação da média e desvio padrão do percentual de perdas por estado.	101
5.3	Valores dos parâmetros de escala, forma e média para distribuições dos estados 1 e 2 (SILVA; SANTOS; PEDROSO, 2020).	104
5.4	Comparativo da média e desvio padrão do comprimento das rajadas de perdas obtidas no modelo do HMM4 e do modelo simulado.	105
5.5	Comparativo da média e do desvio padrão dos percentuais de perdas de pacotes obtidos das amostras de validação.	107
5.6	Comparativo entre o modelo proposto e trabalhos relacionados.	112

LISTA DE SIGLAS

4SMM	Four-State Markov Model
2SMM	Two-State Markov Model
ACF	Autocorrelation Function
ACK	Acknowledgement
AP	Access Point
AV	Algoritmo de Viterbi
B	Estado Bom
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
BSA	Basic Service Area
BSS	Basic Service Set
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CFP	Contention-Free Period
CP	Content Period
CS	Carrier Sense
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
CTS	Clear to Send
DCF	Distributed Coordination Function
DEPHMM	Double Embedded Process HMM
DIFS	Distribution Interframe Space
DSSS	Direct Sequence Spread Spectrum
DPBGM	Deterministic Process Based Generative Model
DPSK	Differential Phase Shift Keying
EM	Expectation-Maximization method

FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FHSS	Frequency Hopping Spread Spectrum
FSK	Frequency Shift Keying
FSM	Full-state Markov
FSMC	Finite-State Markov Channel
GE	Gilbert-Elliot
hMM	Hierarchical Markov Model
HMM	Hidden Markov Model
HSMM	Hidden Semi-Markov Model
ICMP	Internet Control Message Protocol
IEEE	Institute of Electrical and Electronics Engineers
IID	Independent and Identically Distributed
IoT	Internet of Things
IP	Internet Protocol
ISM	Industrial Scientific and Medical
ITU-T	Telecommunication Standardization Sector
LAN	Local Area Network
LLC	Link Layer Control
LLCD	Log-log Complementary Distribution
LOS	Line-of-Sight
LRL	Loss Run-Length
MAC	Medium Access Control
MIMO	Multiple-Input, Multiple-Output
MSE	Mean Square Error
NAV	Network Allocation Vector
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiplexing Access
PCF	Point Coordination Function

PDU	Protocol Data Unit
PER	Packet-error Rate
PH	Phase Distribution
PHY	Physical Layer
PL	Path loss
PLS	Plasma Lighting System
PMF	Probability Mass Function
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QQ-Plot	Quantile-Quantile Plot
R	Estado Ruim
RF	Radio Frequency
RRL	Reception Run-Length
RTS	Request to Send
RTT	Round Trip Time
SFM	Simplified Fritchman Model
SIFS	Short Interframe Space
SISO	Single Input, Single Output
SNR	Signal Noise Ratio
STA	Station
TCP	Transmission Control Protocolo
UDP	User Datagram Protocol
VA	Variável Aleatória
VoIP	Voice over IP
WiFi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network

LISTA DE SÍMBOLOS

X_i	Variável aleatória igual a 0 ou 1 do pacote i
r	Probabilidade de $X_i = 1$
\hat{r}	Taxa média de perda de pacotes
n_1	Número de ocorrências de perda em uma série observada
n	Número de pacotes em um série observada
$f_c(j)$	Distribuição de tempo de duração das chegadas consecutivas de pacotes
$f_p(j)$	Distribuição de tempo de duração das perdas consecutivas de pacotes
h	Complemento da probabilidade de perda de pacotes no estado Ruim
k	Complemento da probabilidade de perda de pacotes no estado Bom
p	Probabilidade de transição do estado Bom para Ruim
q	Probabilidade de transição do estado Ruim para Bom
P	Matriz de transição de estados do modelo de Gilbert-Elliot
π_B	Probabilidade estacionária do estado Bom
π_R	Probabilidade estacionária do estado Ruim
x	Número consecutivo de pacotes perdidos
H	Tempo máximo entre as chegadas de pacotes
D	Atraso máximo definido
C	Maior número de estados de erro
$\pi_{D,1}$	Distribuição de atraso do pacote no único estado de perda
a	Probabilidade de chegada de um pacote
r_t	Probabilidade de transição para o estado de perda a partir do estado 0 ($c_{0,1}$)
s	Probabilidade de transição para o estado de não perda a partir do estado 1 ($c_{1,0}$)
m	Número máximo de estados do modelo de Gilbert estendido
S_m	Último estado do modelo de Gilbert estendido
$p_{L,cum}(x)$	Taxa de perda acumulada
$p_{L,x}$	Frequência relativa de ocorrência de perdas

L_K	Estado máximo de perda do processo LRL do modelo de Gilbert estendido
π_j	Probabilidade estacionário do modelo de Gilbert estendido
π_S	Probabilidade estacionário dos estados de sucesso
π_L	Probabilidade estacionário dos estados de perda
π_{L_0}	Probabilidade estacionária do estado inicial no modelo de Gilbert estendido
π_{L_i}	Probabilidade estacionária dos estados intermediários no modelo de Gilbert estendido
π_{L_k}	Probabilidade estacionária do estado final no modelo de Gilbert estendido
E_1	Estado de perda de pacotes do modelo hMM
C_i	Estados de recepção de pacotes do modelo hMM
$P_{i,j}$	Probabilidade de transição do estado i para j
$t_{i,j}$	Probabilidade de transição do estado i para j do modelo FSMC com SNR
$f_g(k)$	Distribuição de ocorrência de estados Bons
$f_b(k)$	Distribuição de ocorrência de estados Ruins
π_i	Probabilidade de transição entre os estados do modelo 4SMM com $i = \{1, \dots, 8\}$
α_g	Captura inclinação inicial curta da distribuição de probabilidade do estado Bom
α_b	Captura inclinação inicial curta da distribuição de probabilidade do estado Ruim
β_g	Captura inclinação da cauda longa da distribuição de probabilidade do estado Bom
β_b	Captura inclinação da cauda longa da distribuição de probabilidade do estado Ruim

p_g	Frequência relativa das ocorrências de estados Bons
p_b	Frequência relativa das ocorrências de estados Ruins
X_ℓ	Variável aleatória de série logarítmica
θ	Parâmetro da distribuição logarítmica
ρ	Coeficiente de correlação do método de Pearson
τ	Coeficiente de correlação do método de Kendal
φ	Coeficiente de correlação do método de Spearman
Y_n	Variável estocástica
S	Espaço de estados do modelo baseado em HMM
Π	Vetor de estado inicial do HMM
N	Número de estados do HMM
X_n	Variável estocástica do processo de observações do HMM
A	Matriz de transição entre os N estados do HMM
B	Matriz de probabilidade de observações do HMM
M	Número de símbolos na série de observações usadas no treinamento do HMM
Δ	Tripla do conjunto de parâmetros do HMM
Θ	Conjunto de dados utilizado no treinamento do HMM
$F(x)$	Função de distribuição acumulada
$\bar{F}(x)$	Distribuição complementar acumulativa
$f(x)$	Função de densidade probabilidade
λ	Parâmetro de escala da distribuição de Pareto Tipo II
α	Parâmetro de forma da distribuição de Pareto Tipo II
μ	Parâmetro de média da distribuição exponencial

SUMÁRIO

RESUMO	6
ABSTRACT	7
LISTA DE ILUSTRAÇÕES	9
LISTA DE TABELAS	10
1 INTRODUÇÃO	20
1.1 Objetivos	23
1.1.1 Objetivo geral	23
1.1.2 Objetivos específicos	23
1.2 Contribuições	24
1.3 Estrutura da tese	24
2 REDES WIFI	26
2.1 Camada Física	28
2.2 Camada de Enlace	31
2.3 Causas de Perdas de Pacotes	32
2.3.1 Perdas no meio físico	33
2.3.1.1 Interferência	33
2.3.1.2 Desvanecimento do canal	37

2.3.2	Perdas no Acesso ao Meio	40
2.3.3	Perdas por Falta ou Excesso de Memória	41
3	MODELOS DE PERDAS DE PACOTES	44
3.1	Bernoulli	44
3.2	Gilbert-Elliot (GE)	45
3.2.1	Gilbert-Elliot com parâmetro de atraso	48
3.3	Gilbert Estendido	49
3.3.1	LRL com número ilimitado de estados	50
3.3.2	LRL com limitado número de estados	51
3.4	Adaptação do modelo de Gilbert Estendido	52
3.5	Gilbert-Elliot com Sub-estados	55
3.6	Markov Hierárquico (hMM)	56
3.7	Canal de Markov com Estados Finitos (FSMC)	57
3.8	Markov Escondido (HMM)	58
3.9	Markov de Quatro Estados (4SMM)	61
3.10	Distribuição de Séries Logarítmicas	62
3.11	Comparação dos modelos de perdas em redes sem fio	63
4	MODELO PROPOSTO	65
4.1	Problema e lacunas em aberto	65
4.2	Metodologia	66
4.3	Cenário de coleta de dados	68
4.4	Escolha das variáveis	75

4.4.1	Variáveis de camada física	78
4.4.2	Atraso e Jitter	79
4.4.3	Variáveis de camada MAC	83
4.5	Modelo de Perdas	85
4.5.1	Hidden Markov Model (HMM)	85
4.5.2	Definição do número de estados	90
4.5.3	Modelagem das perdas por estado	93
4.6	Validação do Modelo	97
5	ANÁLISE DOS RESULTADOS E DISCUSSÃO	98
5.1	Treinamento do HMM e Parametrização do Modelo	98
5.2	Avaliação do Modelo Proposto	104
5.3	Comparação com Trabalhos Relacionados	108
6	CONCLUSÕES E TRABALHOS FUTUROS	116
A	PARAMETRIZAÇÃO MODELO FSMC	118
A.1	FSMC de 4 estados	118
A.2	FSMC de 5 estados	119
A.3	FSMC de 7 estados	119
A.4	FSMC de 10 estados	119
A.5	FSMC de 20 estados	120

CAPÍTULO 1

INTRODUÇÃO

Os avanços tecnológicos em redes sem fio permitiram um crescimento no desenvolvimento de aplicações de software e dispositivos de hardware que conectem os usuários à diferentes arquiteturas de redes e serviços. As redes sem fio tem sido amplamente utilizadas em ambientes residenciais, públicos, corporativos, industriais e instituições, promovendo interligação entre os dispositivos com maior flexibilidade e mobilidade (BAGHAEI; HUNT, 2004). Uma rede local, também conhecida como rede LAN (*Local Area Network*), é um conjunto de dispositivos conectados em um local físico, como um prédio, escritório ou casa. Uma LAN pode ser pequena ou grande, variando de uma rede doméstica com um usuário à uma rede corporativa com milhares de usuários e dispositivos em um escritório ou escola.

Dados publicados em 2020 no relatório anual da Cisco (CISCO, 2020) referente a evolução das redes de acesso desde 2018 indicam que o número total de usuários móveis globais aumentará de 5,1 bilhões em 2018 para 5,7 bilhões em 2023. Ainda no cenário global, as velocidades de conexão WiFi (*wireless fidelity*) originadas de dispositivos móveis devem triplicar até 2023, e a velocidade média de conexão da rede WiFi (30,3 Mbps em 2018) excederá 91,6 Mbps em 2023. A quantidade de pontos de acesso WiFi crescerá até quatro vezes a quantidade de 2018, onde haverá quase 628 milhões de pontos de acesso WiFi públicos até 2023, contra 169 milhões de 2018. Com essa expansão, também decorrente da demanda por aplicações de vídeo, resultou a necessidade por redes de comunicação que ofereçam maior largura de banda, baixo atraso (WU et al., 2016), melhores taxas de transmissão e maior confiabilidade. Em 2020 este cenário tende a ser diferente devido ao comportamento dos usuários durante o período de pandemia do COVID-19. Em todo o mundo, as

peças passaram a realizar atividades e/ou estudos de forma remota em suas residências (BLOOM; DAVIS; ZHESTKOVA, 2021; GONZALEZ et al., 2020), o que fez com que tecnologias como serviços de vídeo conferência, acessos remotos e também *streaming* de vídeo para entretenimento exigissem ainda mais das redes WiFi devido ao aumento da utilização das redes pelos usuários (CANDELA; LUCONI; VECCHIO, 2020).

Este cenário exige que as redes sem fio evoluam rapidamente permitindo um conjunto de benefícios, como ganhos de eficiência, mobilidade e menores custos de negócios. O principal padrão disponível hoje para redes WLAN (*Wireless LAN*) é o IEEE 802.11, também chamado de WiFi. Iniciado em 1990, o IEEE 802.11 especifica o gerenciamento de controle de acesso ao meio (MAC, *Media Access Control*) e da camada física (PHY, *Physical Layer*) para conectividade sem fio em estações fixas, portáteis e móveis dentro de uma área de cobertura local (PRASAD et al., 2001). O principal cenário de aplicações de redes WiFi são edifícios comerciais e ambientes residenciais, onde espera-se que o número de dispositivos aumente ainda mais rapidamente com a implementação de novas tecnologias como a Internet das Coisas (IoT, *Internet of Things*) (AL-FUQAHA et al., 2015).

A qualidade em redes WiFi é uma característica crucial para manter os serviços em perfeito funcionamento, garantindo assim níveis de qualidade de serviço (QoS, *quality of service*) aceitáveis. O QoS pode ser definido como a totalidade das características de um serviço de telecomunicações que atendam a capacidade de satisfazer as necessidades do usuário de um serviço (ITU-T, 2008). Os principais parâmetros relacionados à QoS em redes de comunicação são a probabilidade de perda de pacotes, o *throughput*, o atraso fim-a-fim e o *jitter* (FOSTER et al., 2004).

Um modelo é uma representação simplificada de um sistema complexo real e é projetado para exibir características significativas do sistema que se deseja estudar, prever, modificar ou controlar. Assim, um modelo normalmente inclui alguns aspectos, mas não todos, do sistema que está sendo modelado. Um modelo é valioso na

medida em que fornece intuições úteis, previsões e respostas para as perguntas que ele usa para abordar (KELLNER; MADACHY; RAFFO, 1999). Em redes WiFi, um modelo permite entender o comportamento da rede e auxilia a representar algumas características como as perdas de pacote, atraso ou *jitter*. Nestas redes, as perdas de pacotes ocasionadas durante a transmissão ocorrem por diferentes motivos, como colisões oriundas do compartilhamento do canal, a existência de terminais ocultos, interferências no sinal de transmissão, baixa intensidade do sinal, ruídos, efeito multicaminho (*multipath*), e estouro ou tamanho excessivo de *buffer* de memória.

Os modelos de perdas de pacotes permitem que Engenheiros realizem o planejamento de sistemas de comunicação para atender os requisitos de qualidade das aplicações. Desde os anos 60, os modelos de perda de pacotes são usados para simular esse comportamento e permitem a análise de desempenho de diferentes serviços e aplicações. O principal modelo em uso para modelagem da perda de pacotes é o modelo de Gilbert-Elliot (GE) (GILBERT, 1960) (ELLIOTT, 1963). O modelo de GE é representado por uma cadeia de Markov de dois estados em que cada estado é denominado Bom (B) ou Ruim (R) (CHEN; HUNG, 2011). Um estado é considerado B quando um pacote é transmitido com sucesso e R quando ocorre uma perda. Uma sequência de perda de pacotes é conhecida como perdas em rajadas (WANG; XU, 2007) e dentro de um modelo pode ser considerado como um estado ruim, de erro ou falha. As rajadas podem ser mensuradas pelo tempo de duração do estado ruim ou também pela quantidade de perdas ocorridas dentro de uma rajada (comprimento da rajada). O comprimento das rajadas, seja de falha ou de chegada, apresenta dependência temporal entre o comprimento da rajada atual e os anteriores o que prejudica o uso do modelo de GE sem memória (ROHLING et al., 2016), o que implica na necessidade de desenvolvimento de novas abordagens para modelagem de perdas em redes sem fio.

Modelos ocultos de Markov (HMM, *Hidden Markov Models*) têm sido usados como alternativa para modelar e analisar diferentes comportamentos em redes sem fio (HA-

SAN et al., 2010) (ALLAHDADI; MORLA, 2019) (KODA et al., 2017), incluindo técnicas de seleção de canais de comunicação (SENTHILKUMAR et al., 2018). Um HMM é um modelo estocástico formado por duas estruturas, em que a primeira é um processo estocástico não observável diretamente, mas só pode ser inferido por um segundo processo observável (RABINER; JUANG, 1986). Normalmente, os modelos baseados em HMM são usados para modelar sequências de dados obtidos de fenômenos observados em processos reais (ADAMS; BELING; COGILL, 2016).

1.1 Objetivos

1.1.1 Objetivo geral

Desenvolver um novo modelo de perdas de pacotes para redes IEEE 802.11b/g/n considerando simultaneamente diferentes variáveis de camada física e de enlace para aprimorar a acurácia na modelagem do comprimento das rajadas de perdas.

1.1.2 Objetivos específicos

Os objetivos específicos desta tese são:

- Realizar o levantamento das principais causas de perdas de pacotes em redes WiFi.
- Apresentar os principais modelos para caracterização de perda de pacote em rede WiFi disponíveis na literatura, bem como a eficiência desses modelos.
- Coletar amostras de perda de pacotes em redes típicas e em cenários controlados para parametrização do modelo proposto.
- Avaliar a correlação entre as perdas de pacotes e os diversos parâmetros de desempenho coletados, como: atraso, *jitter*, relação sinal ruído (SNR), potência

de sinal, qualidade do sinal, taxa de transmissão, percentual de ocupação e número de usuários no canal.

- Desenvolver um novo modelo de perda de pacotes para redes IEEE 802.11b/g/n baseado em HMM considerando simultaneamente características de camada física e de enlace.
- Comparar o desempenho do modelo proposto com os principais modelos disponíveis na literatura.

1.2 Contribuições

As principais contribuições deste trabalho podem ser resumidas da seguinte forma: (I) São apresentados vários testes com diferentes variáveis existentes em redes WiFi que são classificados como irrelevantes ou redundantes para modelar o comportamento das perdas de pacotes, permitindo assim a redução no número de variáveis simplificando a parametrização e uso do modelo. (II) Um novo modelo de perda de pacotes para redes IEEE 802.11b/g/n baseado em HMM é proposto. (III) O modelo proposto considera duas variáveis, a SNR e a ocupação do canal. Os modelos atuais não consideram essas duas variáveis simultaneamente para modelar as perdas. (IV) Ainda, é mostrado que para redes WiFi sob condições ideais de transmissão na camada física, em que a SNR é alto, as perdas têm grande influência da ocupação do canal. Neste caso, considerar apenas a SNR para modelos de perdas é falho porque não são considerados outros fatores importantes, ou seja, a ocupação do canal na camada de enlace.

1.3 Estrutura da tese

Além deste capítulo introdutório, esta tese está organizada da seguinte forma: No Capítulo 2 é apresentada uma visão geral sobre redes WiFi e as características

do protocolo IEEE 802.11 e suas versões, além das principais causas de perdas de pacotes nestas redes. No Capítulo 3 são descritos e comparados os principais modelos de perdas de pacotes disponíveis na literatura. O Capítulo 4 apresenta o modelo proposto, bem como o cenário de coleta de dados, o procedimento de definição das variáveis, a escolha do número de estados e a validação do modelo. Os resultados são apresentados e discutidos no Capítulo 5, no qual é avaliado o desempenho do modelo e também comparado com os demais trabalhos da literatura. Por fim as conclusões, trabalhos futuros e publicações realizadas são apresentadas no Capítulo 6.

CAPÍTULO 2

REDES WIFI

A conectividade sem fio através de redes WiFi para smartphones, tablets e computadores está atualmente bem estabelecida e as especificações do seu principal protocolo IEEE 802.11 se tornou o principal padrão para redes locais sem fio WLAN. Ao longo dos anos, várias versões do IEEE 802.11 foram lançadas e esses padrões foram projetados para serem compatíveis com versões anteriores em nível de camada física e de enlace. A compatibilidade com as versões anteriores é um requisito importante para garantir a interoperabilidade entre os dispositivos devido à grande diversidade de tipos de dispositivos e diferentes redes disponíveis. Além disso, a especificação do IEEE 802.11 fornece alguns requisitos opcionais que os fabricantes podem ou não implementar em seus produtos (SILVA; PEDROSO, 2019).

O IEEE 802.11 usa o controle de link lógico (LLC, *Logical Link Control*) 802.2 para fornecer uma camada física otimizada (PHY) e também especifica o controle de acesso ao meio (MAC) e subcamadas físicas para comunicação sem fio com diferentes esquemas de modulação de sinal (ZHU et al., 2004), conforme mostrado na Figura 2.1.

A arquitetura do IEEE 802.11 é construída por meio de um conjunto de serviços básicos (BSS, *basic service set*), conforme ilustrado pela Figura 2.2. Um BSS é definido como um grupo de estações sob controle direto de um único nó com função de coordenação, conforme será melhor definido posteriormente na subseção 2.2. A área de operação de um BSS é conhecida como área de serviço básico (BSA, *basic service area*), onde todas as estações podem se comunicar umas com as outras. O IEEE 802.11 define dois modos de operação: modo de infraestrutura onde as estações se comunicam por meio de um elemento central; e modo ad-hoc onde as estações se

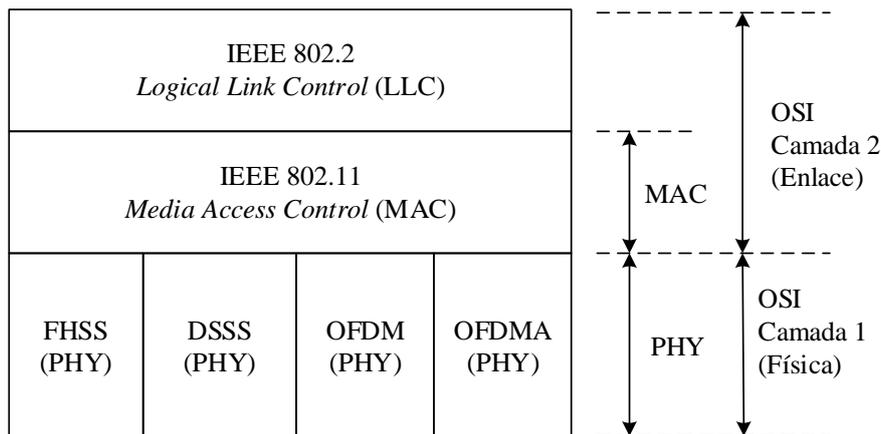


Figura 2.1: Padrão MAC e PHY para IEEE 802.11 (STALLINGS, 2004).

comunicam entre si diretamente. O modo ad-hoc também é conhecido como BSS independente (IBSS, *Independent BSS*) porque a comunicação entre as estações dentro do BSS é realizada sem a necessidade de gerenciamento por um nó central ou controlador.

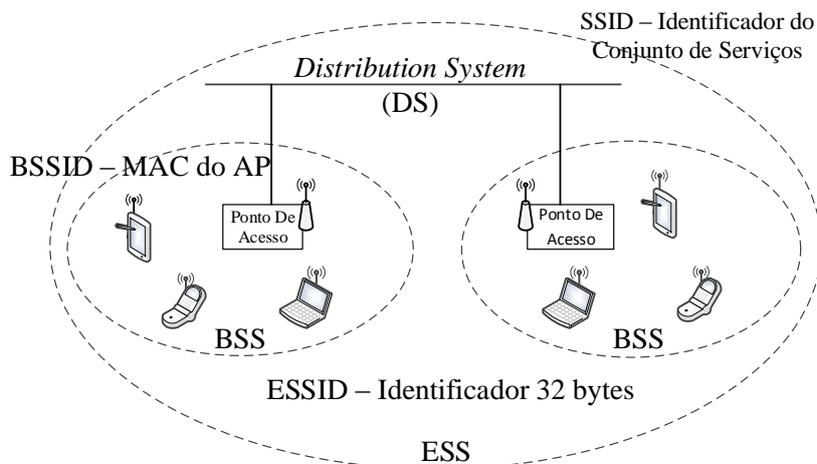


Figura 2.2: Arquitetura de conjunto de serviço estendido (SILVA; PEDROSO, 2019).

Por outro lado, o elemento central do BSS que na maioria das vezes é um ponto de acesso (AP, *access point*) no modo de infraestrutura geralmente permanece estacionário e concentra todo o tráfego entre as estações. O AP anuncia o BSS por meio de mensagens contendo seu endereço de camada de enlace (BSSID, *BSS identifier*) e o nome do identificador de rede (SSID, *service set identifier*). Dispositivos de estação associados ao AP deixam toda a coordenação da funcionalidade da rede para o AP (WIRTZ et al., 2014).

A utilização a partir de uma combinação dos modos infraestrutura e o modo ad-hoc é investigado como uma oportunidade para melhorar o desempenho das redes sem fio (CHEN; CHAN; LIEW, 2003). O AP pode fornecer a integração necessária para redes entre vários BSSs, resultando em um conjunto de serviço estendido (ESS, *extended service set*). O ESS pode ser descrito como um grande BSS para a subcamada LLC de cada estação e estende a área de cobertura a um único BSS. O ESS consiste em vários BSSs integrados em um sistema de distribuição comum (DS, *distribution system*). O DS pode ser entendido como um backbone responsável por transportar unidades de dados de serviço MAC e é especificado no IEEE 802.11 como independente de implementação. Um ESS pode fornecer um gateway para usuários sem fio à redes com fio, como a Internet. Isso é feito por meio de um dispositivo que especifica o ponto de integração no DS onde a rede WiFi se integra com uma rede não IEEE 802.11 (CROW et al., 1997).

2.1 Camada Física

A camada física em redes WiFi estabelece os requisitos para transportar os bits através de uma infraestrutura de equipamentos de rádio transmissão e mecanismos para modulação e equalização do sinal para transmissão de dados. O primeiro padrão IEEE 802.11 iniciado no anos 90, transmitia em uma frequência não licenciada de 2,4 GHz e especificava duas tecnologias de uso de frequência: espalhamento espectral em sequência direta (DSSS, *Direct Sequence Spread Spectrum*) e o espalhamento espectral por saltos de frequência (FHSS, *Frequency Hopping Spread Spectrum*) (VARSHNEY, 2003). No DSSS, cada bit no sinal original é representado por múltiplos bits no sinal transmitido, usando um código de espalhamento. O código de propagação espalha o sinal através de uma faixa de frequência mais ampla em proporção direta ao número de bits usados. No FHSS, o sinal é transmitido através de uma série pseudo-aleatória de radiofrequências. Este método consiste na mudança constante da portadora através de vários canais escolhidos a partir da sequência

pseudo-aleatória de frequências conhecidas por ambos os transmissores e receptores. Essa troca de portadora produz algumas vantagens significativas como a alta resistência a interferências, segurança de interceptação de dados e eficiência espectral (ALATABANI; ABDALLA, 2015).

Em versões mais atuais do IEEE 802.11 é utilizada a multiplexação por multiportadoras ortogonais (OFDM, *Orthogonal Frequency-Division Multiplexing*) (ZHU et al., 2004). No OFDM, a largura de banda acomoda em múltiplas portadoras ortogonais chamados sub-portadoras. As sub-portadoras ortogonais não permitem que as frequências se sobreponham, assim não interferindo umas nas outras. As principais vantagens do OFDM em relação às técnicas clássicas que utilizavam apenas uma portadora são: (i) possibilidade de obter as mesmas taxas de transmissão; (ii) diminuir os efeitos do desvanecimento do canal e oferecer facilidade de equalização; (iii) possibilidade de ajustar a técnica de modulação nas sub-portadoras, onde por exemplo, uma sub-portadora em condição espectral ruim utiliza o BPSK (*Binary Phase Shift Keying*) e para uma boa condição de frequência espectral utiliza o 64-QAM (64 bits *Quadrature Amplitude Modulation*) (HWANG et al., 2009). Informações adicionais para uma análise detalhada sobre técnicas de modulação e detalhes físicos de implementação podem ser consultadas em (HWANG et al., 2009) e (SGORA; VERGADOS; VERGADOS, 2015). A Tabela 2.1 apresenta as principais características de camada física e taxas de transmissão das principais versões do IEEE 802.11.

O IEEE 802.11a podia operar em 3,7 GHz ou 5 GHz usando FDMA (*Frequency Division Multiple Access*) e SISO (*Single Input, Single Output*). O modelo SISO permite apenas a comunicação entre um transmissor e um receptor. No entanto, hoje em dia é comum o uso de técnicas MIMO (*Multiple-Input Multiple-Output*), onde podem ser utilizados vários transmissores e receptores através da rede. Os padrões IEEE 802.11b/g/n utilizam frequência de comunicação de 2,4 GHz. O IEEE 802.11e permite transmissões em frequências de 5 GHz e introduz capacidades de QoS. O IEEE 802.11n permite taxas de comunicação de até 150 Mbps (STUDER; FATEH;

Tabela 2.1: Especificação histórica dos padrões IEEE 802.11 (ZHU et al., 2004)(ABDELRAHMAN; MUSTAFA; OSMAN, 2015)

Versão protocolo	Data de lançamento	Frequência	Largura de banda	Vazão	MIMO streams	Técnica de modulação
		GHz	MHz	Mim-Max Mbits/s		
802.11	Jun 1997	2,4	22	1-2	1	DSSS, FHSS
a	Set 1999	5,0	20	6-54	1	OFDM (SISO)
		3,7				
b	Set 1999	2,4	22	1-11	1	DSSS (SISO)
g	Jun 2003	2,4	20	6-54	1	OFDM, DSSS (SISO)
e	Set 2005	5	20	6-54		OFDM
n	Out 2009	2,4/5	20	7,2 - 72,2	4	OFDM (MIMO)
			40	15 - 150		
ad	Set 2010	60	2160	385-4620	>1	Single Carrier, OFDM
ac	Dez 2013	5	20	7,2 - 96,3	8	OFDM (MU-MIMO)
			40	15 - 200		
			80	32,5 - 433,3		
			160	65 - 866,7		
af	Fev 2014	0,47-0,71	6, 7 e 8	569	45	OFDM (MIMO)
ah	Mai 2017	0,9	1, 2, 4, 8 e 16	<347	4	OFDM (MIMO)
ax	2019	2,4/5	20, 40, 80 e 160	<9,6x10 ³		OFDMA

SEETHALER, 2010) e suporta o sistema MIMO capaz de aumentar a confiabilidade, as distâncias de comunicação e as taxas de transmissão (TRAMARIN et al., 2016). O IEEE 802.11ac, também conhecido de WiFi de quinta geração (WiFi5), pode usar as larguras de banda de 20, 40 ou 80 MHz (ONG et al., 2011), além disso, fornece suporte opcional para operação em canais de 160 MHz permitindo maiores taxas de transmissão. As larguras de banda de 80 e 160 MHz podem ser formadas por uma combinação de dois canais adjacentes não sobrepostos de 40 e 80 MHz, respectivamente (BEJARANO; KNIGHTLY; PARK, 2013). O IEEE 802.11ad, desenvolvido pela Samsung Electronics, opera em um frequência de 60 GHz com larguras de banda de pouco mais de 2 GHz, o que permite transmissões sem fio com links de vários Gbps (PERAHIA; GONG, 2011). Em 2019 foi lançada a versão 802.11ax, conhecida como

WiFi de sexta geração (WiFi6), que pode operar sobre as frequências de 2,4 e 5 GHz. Este novo padrão prevê maiores taxas de transmissão em relação as versões anteriores, onde a taxa de transmissão pode chegar até 9,6 Gbps operando sobre o OFDMA (*Orthogonal frequency-division multiple access*). No OFDMA as subportadoras podem ser usadas de forma independente pelos elementos de rede, diferente do OFDM onde um elemento utiliza todas as subportadoras enquanto transmite.

2.2 Camada de Enlace

A camada MAC é responsável por procedimentos de alocação de canal, endereçamento de unidade de dados de protocolo (PDU, *Protocol Data Unit*), formatação de quadro, verificação de erro, fragmentação e remontagem. O MAC IEEE 802.11 suporta dois protocolos básicos de acesso ao meio: a função de coordenação distribuída (DCF, *Distributed Coordination Function*) e a função de coordenação de ponto (PCF, *Point Coordination Function*). Quando o PCF está habilitado, o canal sem fio é dividido em vários superquadros. Cada superquadro consiste em um período sem contenção (CFP, *Contention-Free Period*) para um PCF e um período de contenção (CP, *Contention Period*) para um DCF. No início do CFP, o ponto de coordenação, geralmente representado por um AP gerencia a disputa pelo canal sem fio pelos dispositivos ou estações da rede. Uma vez adquirido o canal, realiza periodicamente a varredura de alta prioridade das estações para permitir privilégios de transmissão. O PCF é um serviço centralizado e opera apenas no modo de infraestrutura (ZHU et al., 2004). No modo de infraestrutura, a rede é controlada por um método centralizado que gerencia o acesso e a comunicação entre os dispositivos. Além do modo de infraestrutura, o WiFi pode operar no modo *ad-hoc*, onde a rede é composta somente por estações que se comunicam entre si sem a necessidade de um nó de controle central (RUBINSTEIN; REZENDE, 2002).

O DCF baseia-se no acesso múltiplo com sensoriamento de portadora com preven-

ção de colisão (CSMA/CA, *Carrier Sense Multiple Access with Collision Avoidance*). Diferente do CSMA com detecção de colisão (CSMA/CD, *CSMA with Collision Detection*), o CSMA/CA não permite que uma estação envie dados no canal durante uma transmissão já em execução, porém não é eficiente devido ao problemas de terminais ocultos (*hidden terminal*). No IEEE 802.11, o sensoriamento da portadora (CS) é realizado pelas camadas PHY e MAC. Inicialmente no CSMA/CA, antes de enviar um quadro no meio físico, o transmissor verifica se o meio está livre transmitindo um quadro com uma requisição de envio (RTS, *Request to Send*). Se o meio estiver ocupado, o transmissor aguarda um tempo aleatório chamado DIFS (*Distribution Interframe Space*). Quando o meio está livre, o receptor recebe o RTS e depois de esperar um tempo curto (SIFS, *Short Interframe Space*) responde com um quadro de sinalização de disponível para enviar (CTS, *Clear to Send*). As estações que recebem o CTS atualizam seu vetor de alocação de rede (NAV, *Network Allocation Vector*), que representa o tempo que a estação deve esperar antes de tentar usar o mesmo canal. Se o transmissor que inicialmente enviou o RTS recebe o CTS, é iniciado o envio da informação de dados para o receptor que retorna uma confirmação de mensagem positiva (ACK, *Acknowledgment*) após o recebimento dos dados (LAUFER; KLEINROCK, 2016). Este mecanismo auxilia na constante verificação do meio de transmissão identificando e evitando as colisões geradas pelo excesso de estações na rede.

2.3 Causas de Perdas de Pacotes

A perda de pacotes ocorre quando um ou mais pacotes não chegam ao seu destino. Entender e determinar as causas de uma perda de pacotes é crucial e permite identificar qual procedimento ou ação a ser tomada nas camadas física e enlace. Em alguns casos, colisões e problemas de potência de sinal levam a problemas no desempenho de redes sem fio e precisam ser entendidas separadamente (RAYANCHU et al., 2008).

As causas da perda de pacotes podem ser classificadas em três categorias diferentes (SILVA; PEDROSO, 2019). Primeiro, as características relacionadas à camada física como potência do sinal, ruídos e o fenômeno de multicaminhos. Segundo, a competição pelo acesso ao meio de transmissão, agravado pelo problema de terminais ocultos que podem resultar em colisões que levam à perda de pacotes. E terceiro, o congestionamento de rede sem fio gerado pela ocupação de áreas de *buffer*, decorrente do espaço de memória insuficiente nos dispositivos.

O IEEE 802.11 foi projetado para detectar colisões, enviando constantemente mensagens para verificar a condição do canal e estimar a probabilidade de colisões. No entanto, é difícil isolar as causas físicas ou da concorrência e controle do acesso ao meio, ou seja, diferenciar perdas decorrente da baixa potência no receptor em relação a terminais ocultos (RAYANCHU et al., 2008). Determinar essas causas, quando ocorrem e como são relacionadas é uma questão desafiadora e são a chave para melhorar o desempenho de redes WiFi (ABUSUBAIH, 2012) e serão tratados a seguir.

2.3.1 Perdas no meio físico

2.3.1.1 Interferência

É possível definir dois tipos de interferência: aquelas causadas por outras estações e as causadas por ruídos. A interferência de estações está relacionada ao processo de disputa e controle do meio de transmissão MAC que será tratada na seção 2.3.2, enquanto que a interferência de ruído é devida ao meio físico (TINNIRELLO; BIANCHI, 2010). Essa interferência no meio físico pode ser causada por outras estações que tentam acessar o meio ou outras tecnologias usando o mesmo canal.

Tecnologias como o Bluetooth (IEEE 802.15.1), o ZigBee (IEEE 802.15.4) e até mesmo alguns tipos de telefones sem fio são fontes de interferência porque utilizam as mesmas faixas de frequências não licenciadas. O uso da mesma frequência no IEEE 802.11 e no IEEE 802.15.1 gera uma interferência em ambas as tecnologias, e

nesses casos, pode causar baixa taxa de transferência quando os dispositivos estão próximos uns dos outros (MATHEW et al., 2010; GOLMIE; CHEVROLLIER; REBALA, 2003). O IEEE 802.15.1 e IEEE 802.15.4 são padrões de comunicação sem fio para redes pessoais sem fio (WPAN, *Wireless Personal Area Network*).

Estudos presentes em (YOON et al., 2007) e (SHIN; PARK; KWON, 2007b) mostram uma análise sobre a interferência de redes WPAN em redes WiFi, respectivamente usando os padrões IEEE 802.15.4 e IEEE 802.11b. A taxa de erro de pacote (PER, *Packet-error Rate*) do WiFi foi avaliada a partir da taxa de erro de bit (BER, *Bit Error Rate*) e do tempo de colisão. Para distâncias maiores que 4 metros, a interferência não prejudica significativamente o sinal WiFi e para distâncias menores, os autores sugerem mudar o tamanho da carga útil de dados para mitigar os efeitos da interferência da WPAN. O WPAN também pode prejudicar o desempenho do IEEE 802.11g/n como relatado em (PETROVA et al., 2007), sendo essa interferência causada por outras tecnologias que operam na mesma frequência chamada de interferência pulsada. A interferência do WPAN no WiFi também foi estudada em (ZARIKOFF; LEITH, 2013), que propõe uma técnica para detectar a presença de interferência pulsada e classificar a transmissão perdida em induzida por ruído, induzida por colisão ou perda de nó oculto. A coexistência entre o IEEE 802.16 (WiMAX, *Worldwide Interoperability for Microwave Access*) e o IEEE 802.11 também é um problema recorrente em redes sem fio, sendo que o acesso ao meio físico baseado em quadros do IEEE 802.16 requer proteção rigorosa contra interferência de redes locais sem fio para funcionar corretamente (BERLEMANN et al., 2006).

A interferência pode ser causada por ruído de outras fontes externas. Os fornos de micro-ondas, por exemplo, usam apenas uma pequena parte da largura de banda de 2,4 GHz e não estão ativos durante a maior parte do ciclo de energia de 20 milissegundos (KAMERMAN; ERKOCEVIC, 1997). Além de fornos de micro-ondas, outros dispositivos de rádio podem interferir no WiFi, como, por exemplo: o sistema de iluminação de plasma (PLS, *Plasma Lighting System*) que é instalado em hotéis e

outros lugares comuns; e os dispositivos de transmissão de vídeo sem fio usados na detecção de explosivos, que é operado em aeroportos (PARK et al., 2003; HUO et al., 2009, 2010).

A coexistência entre várias redes WiFi também é uma questão importante. Nos padrões IEEE 802.11 que utilizam 14 canais de 20 MHz cada, como, por exemplo no IEEE 802.11a/e/g/n que estão disponíveis na banda ISM (*Industrial Scientific and Medical*) de 2,4 GHz, podem ocorrer sobreposições entre canais de mesma faixa de transmissão. O IEEE 802.11 utiliza uma largura de banda de 20 MHz e como resultado, entre os 14 canais existem apenas três canais não sobrepostos. Dada a alta densidade da implantação de pontos de acesso WiFi é inevitável que uma rede compartilhe parte ou toda a sua banda de espectro operacional com redes WiFi vizinhas. O problema da coexistência WiFi torna-se ainda mais preocupante, já que o padrão IEEE 802.11ac prevê aumentar a largura de banda do canal para 80 e 160 MHz (ZHANG et al., 2017). A Figura 2.3 apresenta uma visão sobre algumas sobreposições de canais existentes entre diferentes padrões do WiFi e tecnologias que operam na mesma faixa de frequência. Observa-se que poucos são os canais que não se sobrepõem entre si quando comparado entre padrões iguais, 3 ou 4 canais, respectivamente para as versões b/g/n. No caso do padrão 802.11n onde a largura de banda dos canais pode ser de 40 MHz essa quantidade é reduzida a apenas 2 canais, respectivamente 3 e 11. Com o aumento da largura de banda, há um aumento na taxa de transmissão possível, contudo aumenta a possibilidade de interferências com outros dispositivos que compartilham da mesma faixa de frequência, ocasionando assim mais perdas. Já os padrões de baixo consumo de energia, como o Bluetooth e ZigBee por exemplo, mesmo possuindo canais de transmissão de 1 ou 2 MHz podem acarretar em perdas por interferência aos canais do WiFi como mostrado na figura.

O processo de escolha dos canais ideais ocorre durante o processo de configuração dos pontos de acesso de transmissão baseado no WiFi. Em muitos casos, os usuários configuram de forma que o próprio equipamento gerencie essa tarefa, onde o dispositi-

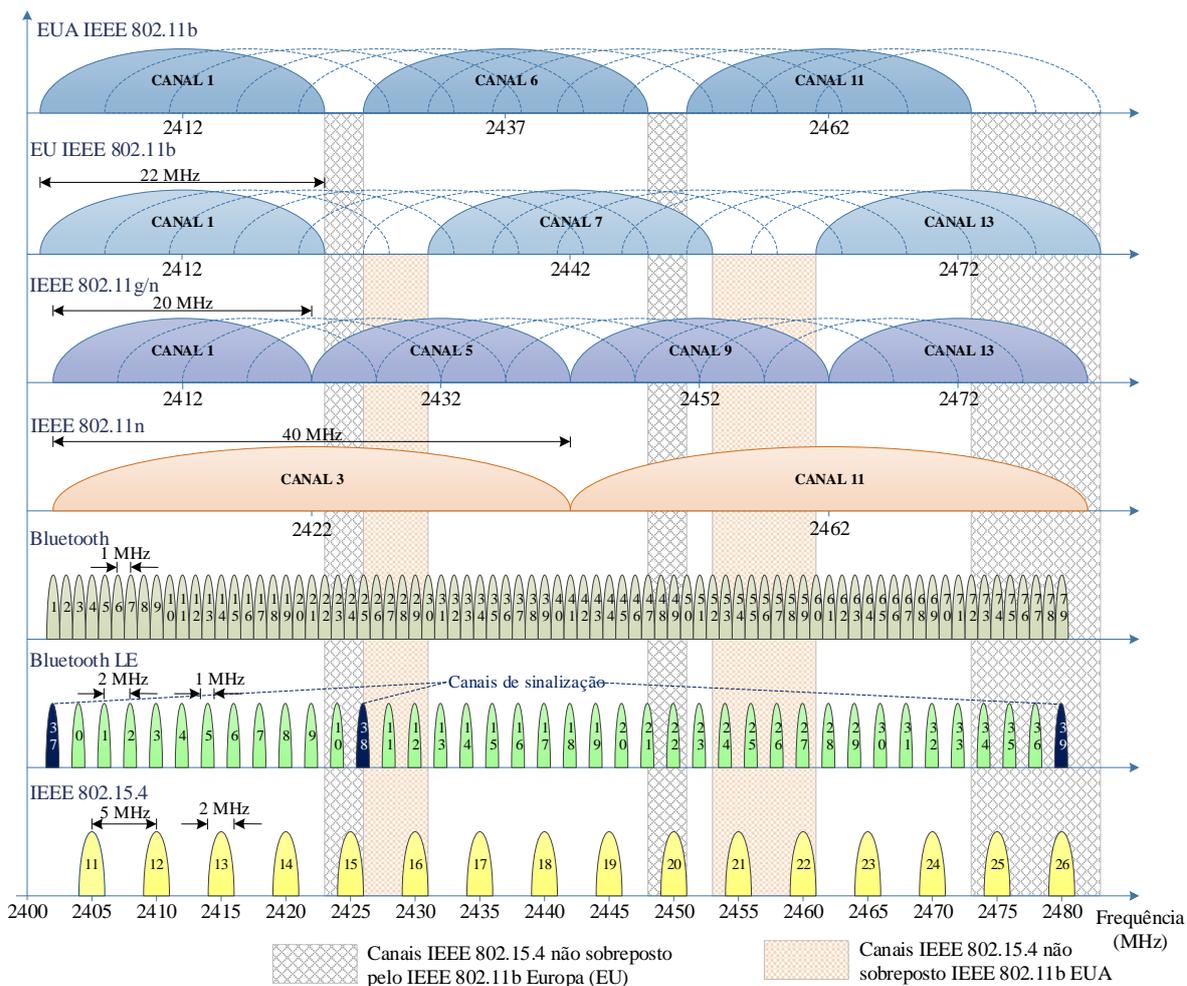


Figura 2.3: Mapa de sobreposição de canais em tecnologia sem fio WiFi (SILVA; PEDROSO, 2019).

tivo de forma automática escolhe o melhor padrão ideal para transmissão, assim como a largura de banda e canais a serem usados.

A Tabela 2.2 resume os trabalhos de pesquisa sobre interferência e coexistência entre várias tecnologias de rede sem fio com o IEEE 802.11. Observa-se que este assunto é abordado em diversos trabalhos e a rápida expansão de novas tecnologias traz novos desafios à desenvolvedores e projetistas desses equipamentos quando o assunto é sobreposição de canais em camada física.

Tabela 2.2: Trabalhos relacionados de interferência e coexistência entre WiFi e outras tecnologias

Rede sem fio	Interferência	Trabalhos relacionados
IEEE 802.11	IEEE 802.15.4	(CHIASSEIRINI; RAO, 2000), (LANSFORD; STEPHENS; NEVO, 2001), (CORDEIRO; AGRAWAL, 2002), (OPHIR; BITRAN; SHERMAN, 2004), (YOON et al., 2006),(SHUAIB et al., 2006), (YUAN; WANG; LIN-NARTZ, 2007),(ANGRISANI et al., 2007),(MYOUNG et al., 2007), (ANGRISANI et al., 2008),(POLLIN et al., 2008),(HOWITT; SHUKLA, 2008),(JUNG et al., 2008), (ROY; S., 2009), (LIANG et al., 2010),(HUANG et al., 2010), (YANG; XU; GIDLUND, 2011),(ZHANG; SHIN, 2011b),(ZHANG; SHIN, 2011a), (ZACHARIAS et al., 2012), (SINGH; SHARMA; TOMAR, 2013), (WINTER et al., 2014), (WINTER et al., 2015), (YANG et al., 2016), (LIU et al., 2017) e (SILVA; SANTOS; OSINSKI, 2018)
	IEEE 802.15.1	(HOWITT; MITTER; GUTIERREZ, 2001),(HOWITT, 2001), (CHIASSEIRINI; RAO, 2002), (CONTI et al., 2003) e (ZACHARIAS et al., 2012)
	IEEE 802.15.4a	(BELLORADO et al., 2003) e (HAMALAINEN; TESI; IINATTI, 2004)
	IEEE 802.15.6	(WANG; CAI, 2011) e (HAYAJNEH et al., 2014)
IEEE 802.15	IEEE 802.11	(SHIN et al., 2005), (BIANCHI; TINNIRELLO; SCALIA, 2005), (SIKORA; GROZA, 2005), (BERLEMANN et al., 2006), (SHIN; PARK; KWON, 2007b), (SHIN; PARK; KWON, 2007a), (SHIN et al., 2007), (HAN et al., 2007), (KANG et al., 2007), (YUN et al., 2008), (HUO et al., 2009) e (HAUER; HANDZISKI; WOLISZ, 2009)
LTE	IEEE 802.11	(MILOS et al., 2016)

2.3.1.2 Desvanecimento do canal

O desvanecimento do canal na rede sem fio, conhecido como *fading*, é definido como a atenuação da intensidade do sinal em função de variáveis como o tempo, a posição geográfica de dispositivos e a frequência de rádio (MIU et al., 2004). Conforme descrito na seção 2.3.1.1, o ruído e as interferências mitigam a qualidade do canal causando perdas na transmissão entre os dispositivos. Esses ruídos de frequência de rádio também afetam o sinal de origem, causando interferências no nó e fluxo de tráfego (QIU et al., 2006).

Em uma comunicação sem fio típica, a estação base é fixada em uma posição enquanto os dispositivos estão em movimento. Esta condição pode causar uma propaga-

ção entre a estação rádio base e os dispositivos em grande parte através da dispersão do sinal, seja por reflexão ou difração em edifícios, terrenos, ou objetos devido à obstrução do caminho da linha de visada (LOS, *Line-of-sight*). Este fenômeno de reflexão ou difração é conhecido como propagação multicaminho. O fenômeno de propagação multicaminho é causado por múltiplas cópias do sinal, de diferentes direções, com diferentes amplitudes, fases e atrasos chegando ao receptor devido a obstáculos e pode levar ao desvanecimento severo do sinal transmitido (SARKAR et al., 2003; DUEL-HALLEN, 2007).

O desvanecimento de sinal é um comprometimento de canal particularmente severo que pode ser mitigado através do uso de técnicas de mitigação, como por exemplo, a técnica de diversidade cooperativa (LANEMAN; TSE; WORNELL, 2004). Para uma comunicação sem fio, a estimativa de uma perda na trajetória (PL, *Path Loss*) é uma medida da atenuação média de radiofrequência (RF, *Radio Frequency*) sofrida por um sinal transmitido quando chega ao receptor, após percorrer um caminho de vários comprimentos de onda (SARKAR et al., 2003).

A mobilidade do usuário também pode ter um efeito negativo nas redes sem fio, aumentando ligeiramente a taxa de perda de pacotes (PRABHAKARAN; SANKAR, 2006). Neste caso um dispositivo pode trocar de estação rádio base enquanto sai de uma área de cobertura com menor qualidade de sinal para uma outra área melhor. O processo de transferência de uma chamada ou sessão de/para estações base é conhecido como *handoff*. O *handoff* pode ser classificado em dois tipos: *handoff* horizontal, o que significa que os dispositivos se movem dentro da mesma tecnologia de rede de acesso sem fio (como por exemplo de uma rede WiFi para outra), e *handoff* vertical, no qual os dispositivos se movem entre tecnologias de rede de acesso sem fio heterogêneas, por exemplo, do WiFi para WiMax, conforme ilustrado na Figura 2.4.

Para transferência vertical, a infraestrutura de *handoff* vertical deve fornecer uma sobrecarga mínima, capacidade de autenticação e baixo atraso para minimizar a perda de pacotes. Para evitar perdas de pacotes durante o *handoff* horizontal, vários métodos

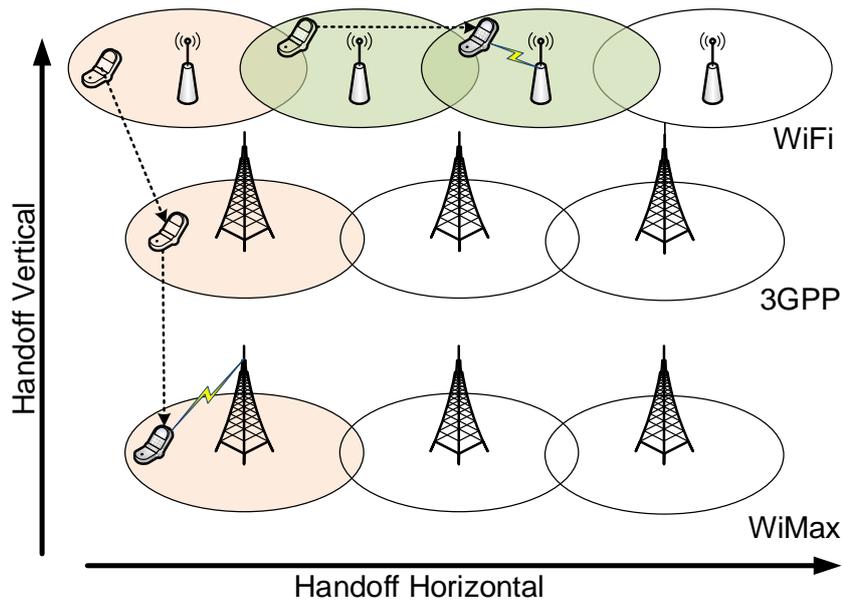


Figura 2.4: *Handoff* vertical e horizontal (SILVA; PEDROSO, 2019).

são propostos na literatura (LUGLIO et al., 2009).

Distância entre o AP e os dispositivos acarreta em diferentes níveis de intensidade de sinal, o que interfere na taxa de perdas de pacotes em redes sem fio. Em longas distâncias a atenuação do sinal de transmissão tende a ser maior, assim aumentando a probabilidade de perdas. Esta atenuação pode ser mitigada a partir de técnicas como o uso de múltiplos AP (CHANG; RASHIDZADEH; AHMADI, 2010). Em (XIAO et al., 2016) são apresentadas pesquisas e discussões para entender o impacto da posição do dispositivo na comunicação sem fio. Neste caso, a potência do sinal recebido (RSS, *Received Signal Strength*) é o recurso de sinal mais comumente usado para localização interna de dispositivos baseada em WiFi, onde o RSS diminui monotonicamente com a distância maior no espaço livre. Desta forma, a potência do sinal de transmissão é deteriorada gradativamente ao mesmo modo que os dispositivos emissor e receptor são afastados entre si, seja em espaço livre onde apenas a distância pode afetar essa potência ou em ambientes com presença de sinais multicaminhos que além da distância também são afetados pelos obstáculos físicos (NAKATANI et al., 2018).

2.3.2 Perdas no Acesso ao Meio

Existem três estados possíveis para um canal: ocupado devido à uma transmissão bem sucedida; ocupado devido à uma colisão; ou canal ocioso (GARETTO; CHIASSERINI, 2005). No primeiro estado, o canal fica ocupado enquanto uma transmissão é bem sucedida. No segundo estado, o canal fica indisponível devido a um evento de colisão e no último estado em que o canal permanece disponível, novas transmissões são permitidas. As colisões no acesso ao meio podem ser causadas por vários motivos, como o número de terminais que compartilham o canal ou terminais ocultos que competem pelo canal.

Em redes IEEE 802.11, o desempenho pode ser afetado devido ao número de estações concorrentes. A competição pelo canal ocorre quando vários terminais tentam simultaneamente enviar um pacote no meio compartilhado, causando uma colisão. Nas redes em modo de infraestrutura IEEE 802.11, a informação sobre os terminais é obtida apenas a partir do número de associações entre AP e dispositivos. Neste caso, é difícil estimar a interferência de dispositivos não associados à rede. Assim, o número de dispositivos concorrentes podem ser muito diferente das associações reais entre AP e dispositivos (BIANCHI; TINNIRELLO, 2003).

O fenômeno de terminal oculto ocorre quando um dispositivo transmissor não pode ouvir outros terminais que estão no alcance de um receptor. As distâncias entre os dispositivos podem reduzir a intensidade do sinal e permitir que os nós na mesma rede não ouçam esses dispositivos, como exemplificado na Figura 2.5. Os dispositivos A e C podem ouvir o dispositivo B e o dispositivo B pode ouvir ambos os dispositivos, porém A e C não podem se ouvir. Quando A transmite para B, C não consegue detectar esta transmissão. Desta forma, se C transmitir para B, ocorrerá uma colisão no dispositivo B. Assim, o aumento na probabilidade de ocorrência de colisões em uma rede WiFi pode ser correlacionada com o número de terminais competindo pelo canal.

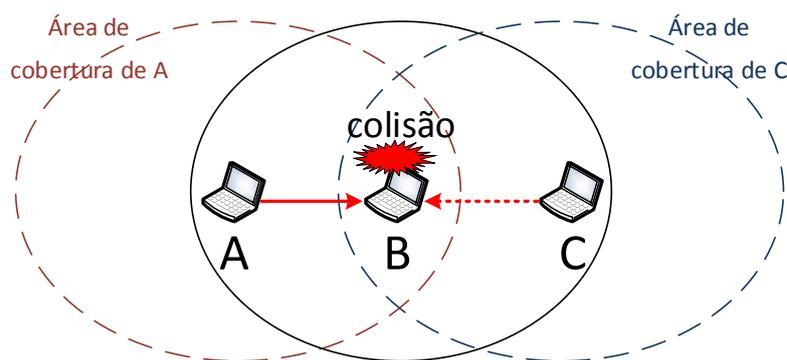


Figura 2.5: Cenário de colisão em redes WiFi.

A combinação de sinais ainda fracos de transmissores pode causar uma interferência suficiente para causar colisões (NASIPURI; ZHUANG; DAS, 1999). Devido à ausência de uma ligação direta entre os dois transmissores, que não ouvem um ao outro, só podem validar a transmissão indiretamente através dos pacotes de confirmação (ACK) enviados pelo receptor como uma confirmação de uma recepção correta de dados. Conseqüentemente, neste caso, o período crítico em que as colisões de pacotes podem ocorrer pode ser bastante longo e gerar perdas de pacotes e isso faz com que o desempenho das estações se degrade (TSERTOU; LAURENSEN, 2008). Em redes IEEE 802.11, o uso de CSMA/CA para evitar colisões em transmissões pode ser mais efetivo usando os mecanismos RTS/CTS contra terminais ocultos apenas para baixas taxas de transmissão e cenários de interferência severa (BORGIO et al., 2004).

2.3.3 Perdas por Falta ou Excesso de Memória

O congestionamento da rede é apontado como uma das razões para perdas de pacotes e a sua detecção pode ser difícil em redes sem fio (YADAV; SINGH, 2016). O congestionamento da rede ocorre quando uma rede (ou uma parte da rede) ou um nó da rede está sobrecarregado com dados. O congestionamento pode acontecer por vários motivos, mas, independentemente da causa, pode ser um grande problema para usuários de serviços que são facilmente afetados por este congestionamento. O problema pode ser resolvido usando dispositivos com processadores mais rápidos

ou também aumentando a quantidade de memória para armazenar todos os pacotes recebidos (POKHREL et al., 2016; JAMSHAD et al., 2014; ELAARAG, 2002). O tamanho do *buffer* de memória é inversamente relacionado à taxa de perda de pacotes e essa taxa é particularmente importante para o desempenho, por exemplo, de aplicações multimídia como o *streaming* de vídeo (SEQUEIRA et al., 2013).

Durante o tempo de duração de uma rajada de pacotes o dispositivo precisa de espaço suficiente no *buffer* para armazenar os pacotes e evitar descartes. Várias técnicas de controle de congestionamento foram desenvolvidas com o objetivo de melhorar a prestação de serviços e conteúdos (KAFI et al., 2014). Esse controle é feito através de protocolos como o TCP (*Transmission Control Protocol*) que fornece comunicação confiável orientada a conexão. Não é objetivo desta pesquisa descrever técnicas para evitar perdas decorrentes da baixa capacidade de armazenamento de *buffer*.

A disponibilidade de memória de baixo custo e a necessidade de evitar a ocorrência perdas de pacotes levaram ao desenvolvimento de *buffers* com maiores capacidade de armazenamento. Desta forma, espera-se que o aumento na capacidade de armazenamento melhore o desempenho da rede, diminuindo a taxa de perda de pacotes. No entanto, isso pode levar a um aumento na latência, o que pode prejudicar a qualidade de certos serviços, por exemplo em *streaming* de vídeo em tempo real. Além disso, *buffers* maiores podem causar um problema chamado *bufferbloat* (*buffers* persistentemente cheios) (GETTYS; NICHOLS, 2011). Nakayama e Sezaki (2018) afirmam que o *bufferbloat* refere-se ao fenômeno do excesso de *buffering* de quadros causando alta latência e baixo rendimento (NAKAYAMA; SEZAKI, 2018). Mesmo quando estudados principalmente no contexto de redes sem fio, *buffer* persistentemente cheios podem deteriorar a justiça da alocação de taxa e aumentar o RTT em redes sem fio (SHOWAIL; JAMSHAD; SHIHADA, 2014).

Ao longo dos anos, vários trabalhos envolvendo modelagem de perda de pacotes em redes sem fio permitem uma nova compreensão do comportamento da perda de

pacotes em redes. A próxima seção descreve os principais modelos de perda de pacotes disponíveis na literatura.

CAPÍTULO 3

MODELOS DE PERDAS DE PACOTES

Uma perda de pacote pode ser causada por uma série de fatores como apresentado no capítulo anterior do ponto de vista da camada física e de enlace e o motivo de perda nem sempre pode ser facilmente identificado. Um bom modelo de perdas do ponto de vista da camada de rede deve considerar todos os efeitos adicionados das camadas inferiores e são importantes para a simulação ou para o desenvolvimento de métodos como controle de admissão, qualidade de serviço, entre outros. Neste capítulo, os principais modelos existentes serão apresentados.

3.1 Bernoulli

No modelo de Bernoulli, o tempo de duração de chegada consecutiva de pacotes (*Good Run-Length*) e o tempo de duração de perdas consecutivas (*Loss Run-Length*) são representados por uma variável aleatória (VA) independente e identicamente distribuída (IID, *Independent and Identically Distributed*). No tempo de duração de chegada, todos os pacotes são recebidos com sucesso e no tempo de duração de perda todos os pacotes são perdidos. Para um dado pacote i a VA X_i pode ser 0 ou 1, com $X_i = 1$ indicando um pacote perdido. A distribuição de tempo de duração das chegadas consecutivas de pacotes para este modelo é indicada na Eq. 3.1 e a distribuição de tempo de duração das perdas é mostrada na Eq. 3.2 (YAJNIK et al., 1999)

$$f_c(j) = \hat{r}(1 - \hat{r})^{j-1} , j = 1, 2, \dots, \infty \quad (3.1)$$

$$f_p(j) = (1 - \hat{r})\hat{r}^{j-1}, \quad j = 1, 2, \dots, \infty \quad (3.2)$$

no qual, o modelo usa um único parâmetro, sendo r a probabilidade de X_i ser 1, estimado a partir de $\hat{r} = n_1/n$, onde n_1 é o número de vezes que o valor 1 ocorreu em uma série temporal observada $\{x_i\}_{i=1}^n$, n é o número de ocorrências na série temporal e \hat{r} é a taxa de perda média.

O modelo de Bernoulli apresenta desvantagens no processo de modelagem de perdas em redes sem fio. Isto porque o processo de Bernoulli não consegue apresentar a dependência temporal do estado futuro em relação ao estado do pacote atual (CHUNG et al., 2015). Para Nguyen et al. (2009) este modelo é claramente insuficiente para descrever muitos cenários do mundo real, mas destina-se apenas a captar a essência da transmissão sem fio usando uma baixa complexidade computacional (NGUYEN et al., 2009). O modelo de Bernoulli é comumente usado para comparar com outros modelos, como mostrado em (CEN; COSMAN; VOELKER, 2003; TANG; MCKINLEY, 2003; ROZNER et al., 2007).

3.2 Gilbert-Elliot (GE)

Em 1960, o matemático americano e teórico da codificação, Edgar Nelson Gilbert propôs o uso de uma cadeia de Markov de dois estados para modelar perdas consecutivas em canais de comunicação com ruídos em rajadas (GILBERT, 1960). A cadeia de Markov possui dois estados identificados como “Bom” (B) e “Ruim” (R), como mostrado na Figura 3.1. No estado R, a probabilidade de descarte é de $1 - h$ e no estado B a transmissão é livre de erros. Ao contrário do modelo de Bernoulli, este modelo de dois estados é capaz de capturar a dependência entre perdas consecutivas (YAJNIK et al., 1999). O fenômeno de perdas em rajadas é uma característica importante principalmente para comunicação sem fio.

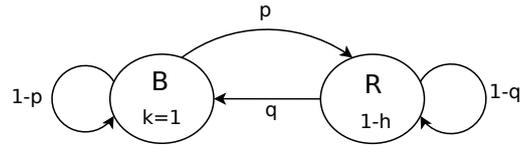


Figura 3.1: Modelo de Gilbert.

O modelo de Gilbert foi estendido por Elliot em 1963 (ELLIOTT, 1963), incluindo a possibilidade de perda em ambos os estados do modelo. O modelo de Gilbert-Elliot (GE) é ilustrado na Figura 3.2. O parâmetro p é a probabilidade de transição do estado Bom para Ruim e q é a probabilidade de transição oposta. Uma perda pode acontecer em cada estado como eventos independentes com probabilidade de $1 - k$ e $1 - h$, respectivamente para os estados Bom e Ruim (HASSLINGER; HOHLFELD, 2008). Geralmente $p + q < 1$ e se $p + q = 1$ o modelo é reduzido ao processo de Bernoulli.

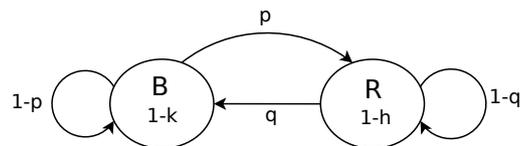


Figura 3.2: Modelo de Gilbert-Elliot.

A matriz de transição de probabilidades do modelo de Gilbert-Elliot \mathbf{P} é dada por

$$\mathbf{P} = \begin{matrix} & \begin{matrix} B & R \end{matrix} \\ \begin{matrix} B \\ R \end{matrix} & \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \end{matrix} . \quad (3.3)$$

A probabilidade estacionária para os estados Bom e Ruim é dada respectivamente por $\pi_B = q/(p + q)$ e $\pi_R = p/(p + q)$. A taxa de perda de pacotes (PLR, *Packet Loss Ratio*) é obtida usando as probabilidades de estado estacionário (LI et al., 2009b) (LI et al., 2009a) e é dada por

$$\text{PLR} = (1 - k)\pi_B + (1 - h)\pi_R . \quad (3.4)$$

O conjunto de parâmetros para os modelos de Markov com dois estados é apre-

sentado na Tabela 3.1. O modelo de Gilbert simples é sempre livre de erros no estado Bom com $k = 1$ e as perdas no estado Ruim podem acontecer com probabilidade $1 - h$. Ainda, como mostrado na tabela, os principais modelos baseados em cadeias de Markov de dois estados apresentam escalas de complexidade diferentes na estimação de seus parâmetros, no qual o modelo de GE possui complexidade alta quando comparado com os modelos mais simples de Gilbert.

A estimativa dos parâmetros do modelo de GE apresenta uma complexidade alta se comparada com os demais modelos também baseados em cadeias de Markov de dois estados.

Tabela 3.1: Comparação dos parâmetros dos modelos de Gilbert simples, Gilbert e Gilbert-Elliot (HASSLINGER; HOHLFELD, 2008).

Modelo	Parâmetro	Complexidade de estimação	Simplificação
Gilbert simples	p, q	baixa	$k=1, h \in \{0, \dots, 0,5\}$
Gilbert	p, q, h	média	$k=1$
Gilbert-Elliot	p, q, h, k	alta	$k, h \in \{0, \dots, 1\}$

Além de modelar a sequência de erros em rajadas no canal, o modelo de GE pode ser usado como um gerador de perda aleatória em redes de comunicação (MO; GARONE; SINOPOLI, 2013). O modelo GE é utilizado para modelar as perdas de pacotes em diversas aplicações, em especial as redes *wireless* (WILLIG et al., 2002; WILLIG, 2002; KONRAD et al., 2003; KOPKE; WILLIG; KARL, 2003; MCDUGALL; MILLER, 2003). Em (HASSLINGER; HOHLFELD, 2008) é avaliado o desempenho dos modelos de Gilbert e GE. Os autores esclarecem que o modelo de Gilbert não é apropriado para pequenas amostras e o modelo de GE apresenta melhor ajuste com processos de perda de pacotes considerando redes com fio.

Em Russ e Haghani (2009a) é apresentado uma distribuição do comprimento de rajadas de erros consecutivos a partir de uma combinação do modelo clássico de Gilbert-Elliot e um modelo de cauda pesada (RUSS; HAGHANI, 2009). Os autores sugerem que os comprimentos das rajadas podem ser expressos por esses dois diferentes modelos, dependendo do número consecutivo de n_P pacotes perdidos: onde para $n_P \leq 3$ é melhor utilizar o modelo de GE e para $n_P > 3$ por uma distribuição de

cauda pesada longa (*Long Heavy-tail*). Os resultados indicam que o padrão observado de rajadas de pacotes descartados se encaixa na combinação do modelo de GE para pequenos comprimentos de rajadas de erros e em uma distribuição de cauda pesada longa para tamanhos maiores de surtos de pacotes descartados em redes sem fio. Os autores sugerem ainda uma investigação mais detalhada para determinar um modelo adequado de cauda pesada longa.

3.2.1 Gilbert-Elliot com parâmetro de atraso

Normalmente, os modelos de perda de pacotes não usam parâmetros como o *jitter* ou o atraso para modelagem do comportamento de perdas. Porém, modelos que permitam utilizar e investigar esses parâmetros aumentam a confiabilidade de novos métodos de estimativa de perdas de pacotes. Lee e Chanson (2002) propõem o uso de um modelo baseado em cadeias de Markov de dois estados com a utilização do parâmetro de atraso.

Uma cadeia de Markov bidimensional com espaço de estados $\{(i, j) | -H + 1 \leq i \leq D, 0 \leq j \leq C\}$ é usada neste modelo, em que H é o tempo máximo de chegada entre os pacotes, D é o atraso máximo definido e C é o maior número de estados de erro.

Assim como para o modelo de GE onde os estados são descritos como B ou R, o modelo sugerido em (LEE; CHANSON, 2002) denota o estado de sucesso pelo estado 0 (um pacote pode ser transmitido com sucesso) e o estado de falha identificado como estado 1 (uma transmissão de pacote falha). A matriz bidimensional de transição c dos estados correlacionados é descrita por $c = \{c_{0,0}, c_{0,1}, c_{1,0}, c_{1,1}\}$, com $\sum_{l=0}^C c_{j,l} = 1 (0 \leq j \leq C)$ (LEE; CHANSON, 2002).

O parâmetro de desempenho mais importante neste modelo, isto é, o PLR é simplesmente representado pela expressão

$$\text{PLR} = \frac{\pi_{D,1}}{a} . \quad (3.5)$$

Note que, $D,1$ é o único estado em que um pacote pode ser (e deve ser) descartado devido ao atraso limite D e a é necessário para obter a fração de pacotes perdidos (LEE; CHANSON, 2002).

O PLR é expressado em função de parâmetros básicos (a, r_t, s, D, H) , onde a é a probabilidade de chegada de um pacote, r_t é a probabilidade de transitar para o estado com perda a partir de um estado 0 ($c_{0,1}$) e s é a probabilidade de estar no estado 1 e transitar para o estado 0 após receber um pacote com sucesso ($c_{1,0}$).

Neste modelo, o PLR é afetado pelo limite de atraso D imposto e é independente da taxa de chegada, exceto quando a taxa está próxima de 1. Assim como em outros modelos já descritos e que serão mencionados, esse também assume que a perda de pacotes ocorre em rajadas e durante um estado sem perdas todos os pacotes são recebidos, e o tamanho da pilha de chegada se reduz a zero. Resultados mostraram que aumentando o atraso limite diminui exponencialmente a probabilidade de perdas desde que o atraso limite utilizado não acarrete no estouro do buffer finito aplicado no modelo. Ao aumentar a taxa de chegada H de pacotes ocorre o aumento da probabilidade de perdas do modelo. Os autores afirmam que a estimação dos parâmetros possui uma maior complexidade computacional, sendo aí uma grande desvantagem do modelo. Contudo, não realizam testes ou comparações com outros modelos da literatura.

3.3 Gilbert Estendido

Um modelo estendido do modelo de Gilbert com múltiplos estados é apresentado na Figura 3.3. Nesta abordagem há duas categorias do modelo de Gilbert original, descritos como *Reception Run-Lengths* (RRL) e *Loss Run-Lengths* (LRL). Para um

RRL há $m + 1$ estados $\{S_0, S_1, \dots, S_m\}$, onde cada evento de recepção levará o sistema ao estado sucessivo até que RRL atinja m , exceto para o estado S_m em que o sistema permanece enquanto ocorrer o recebimento do pacote ou retorna para o estado S_0 depois da ocorrência de uma perda. Para cada perda ocorrida, o retorno para o estado inicial fará com que o processo seja livre de dependência do histórico passado de perdas (SANNECK; CARLE, 1999; LI et al., 2009a).

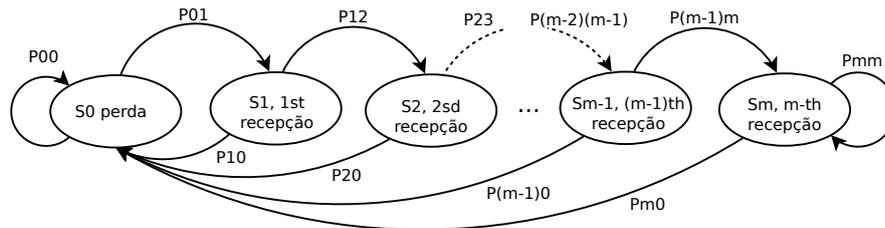


Figura 3.3: Modelo de Gilbert Estendido (YOUNESIAN et al., 2014).

O processo LRL é definido pela sequência de perdas de pacotes sucessivas. Esse modelo LRL pode ser descrito em dois tipos diferentes de abordagens: a primeira com um número limite de estados $m + 1$ e a segunda com um número infinito de estados ($m \rightarrow \infty$), sendo que ambas são descritas a seguir.

3.3.1 LRL com número ilimitado de estados

Este modelo é definido a partir de uma variável aleatória X , onde $X = 0$ representa o recebimento de um pacote e $X = x$ a perda consecutiva de x pacotes. Com esta condição é estabelecido o modelo LRL como representado na Figura 3.4 com infinitas possibilidades de número de $m + 1$ estados, no qual a probabilidade de perda é dependente do tamanho das rajadas que é aumentada toda vez que ocorre uma perda.

Se um pacote é recebido o sistema retorna para o estado $X = 0$, logo pode-se assumir que a probabilidade do sistema para $x > 0$ é $P(X \geq x)$ e pode ser aproximada pela taxa de perda acumulada

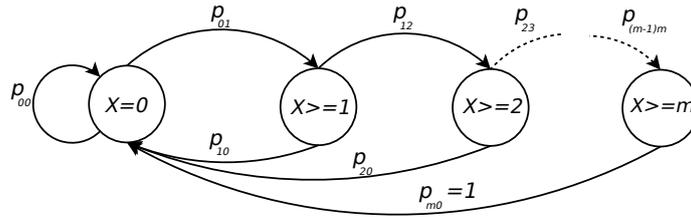


Figura 3.4: Modelo LRL com número ilimitado de estados ($m \rightarrow \infty$)(SANNECK; CARLE, 1999).

$$P_{L,cum}(x) = \sum_{n=x}^{\infty} p_{L,n} \quad (3.6)$$

A taxa média de perdas é expressa como

$$PLR = \sum_{x=1}^{\infty} x \cdot p_{L,x} \quad (3.7)$$

onde, $p_{L,x} = o_x/a$ é a frequência relativa para a ocorrência de falhas, a partir da LRL x dada por o_x o número de ocorrência de rajadas de perda e a o número determinado de pacotes de chegada.

3.3.2 LRL com limitado número de estados

Neste modelo, o número de estados limitados ($m + 1$) é derivado do modelo básico LRL e é apresentado na Figura 3.5. Os parâmetros de configuração desse modelo são semelhantes do LRL com estados infinitos, contudo a probabilidade para o último estado m é $P(X = m)$, e ainda é adicionado a transição do estado m para m . Para o último estado, $X = m$ é o número consecutivos de m pacotes perdidos e também o limite do sistema com memória.

O PLR neste modelo pode ser obtido a partir da equação:

$$PLR = \sum_{n_P=1}^{\infty} \frac{n_P \cdot o_{n_P}}{a} \quad (3.8)$$

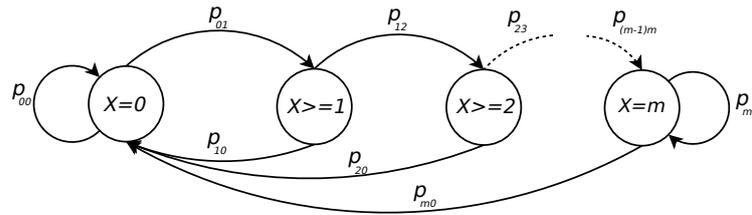


Figura 3.5: Modelo LRL com limitado número de estados $[(m + 1)$ estados](SANNECK; CARLE, 1999).

Para $0 < n_P < m$, $X = n_P$ apresenta o número consecutivo de perdas de pacotes. a é o número de chegadas de pacotes e o_{n_P} é o número de ocorrência de comprimento de perdas em rajadas.

Os autores em (SANNECK; CARLE, 1999) não destacam a aplicação da parametrização desse modelos em uma rede WiFi, sendo apenas indicado que as medições foram realizadas para uma transmissão de voz sobre IP (VoIP, *Voice over IP*). Ainda concluem que para o suporte ao QoS da Internet em aplicativos é necessário um modelo intermediário, que deve ser mais complexo que o modelo simples de Gilbert. A complexidade necessária de tal modelo intermediário é determinada pelos requisitos da aplicação.

3.4 Adaptação do modelo de Gilbert Estendido

Nesta adaptação da versão de Gilbert Estendido, o modelo é baseado em cadeias de Markov com um processo RRL e outro LRL, com mecanismos de ordem m e k respectivamente. A Figura 3.6 apresenta a transição de estados do mecanismo proposto (YOUNESIAN et al., 2014).

- Em RRL, todo pacote recebido levará o sistema para o próximo estado com probabilidade de transição de $P_{S_i, L_{i+1}}$ até atingir o estado S_m . No último estado de RRL, se o pacote é enviado com sucesso o sistema permanece no mesmo estado com probabilidade P_{S_m, S_m} . Caso contrário, se o pacote for perdido, o sistema retorna para o estado inicial L_0 e o mesmo ocorre caso esteja em qual-

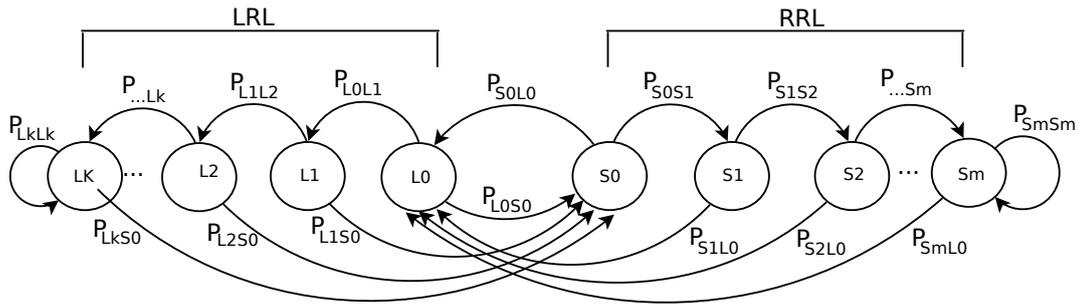


Figura 3.6: Adaptação do modelo de Gilbert Estendido (YOUNESIAN et al., 2014).

quer outro estado de RRL sendo que esta probabilidade de transição é definida por P_{S_i, L_0} . Essa transição faz com que o sistema fique livre do processo de dependência temporal de estados passados (YOUNESIAN et al., 2014);

- No LRL, cada pacote perdido levará o sistema para o próximo estado de perda com probabilidade de $P_{L_i, L_{i+1}}$ até atingir o último estado de perda L_k . No último estado, se o pacote é perdido, o sistema permanece no mesmo estado até que um pacote seja recebido e neste caso fará com que o sistema retorne para o estado S_0 com probabilidade de transição P_{L_i, S_0} . Para qualquer outro estado de LRL, caso um pacote seja recebido o sistema retorna para o estado S_0 com mesma probabilidade de transição (YOUNESIAN et al., 2014).

A probabilidade de estado estacionário (π_j) é definida da seguinte forma, no qual se $\forall j \in S$ onde S é o número de estados e $0 \leq \pi_j \leq 1$

$$\sum_{j \in S} \pi_j = 1. \quad (3.9)$$

$$\pi_j = \sum_{i \in S} \pi_i P_{ij} . \quad (3.10)$$

A probabilidade de estado estacionário do modelo é definida como a soma de todas as probabilidades de estado estacionário de RRL e LRL

$$\sum_m \pi_S + \sum_k \pi_L = 1. \quad (3.11)$$

onde π_S e π_L são as probabilidades de estado estacionário de estados de sucesso e perda para os m e k estados, respectivamente. PLR é dado por

$$\text{PLR} = \sum_k \pi_L. \quad (3.12)$$

As probabilidades de estado estacionário (π_L) para cada estado do LRL podem ser calculadas diferentemente para o estado inicial, os estados intermediários e o estado final.

Para o estado inicial

$$\pi_{L_0} = \pi_{S_0} \cdot p_{S_0, L_0} + \pi_{S_1} \cdot p_{S_1, L_0} + \dots + \pi_{S_m} \cdot p_{S_m, L_0}. \quad (3.13)$$

Para os estados intermediários

$$\pi_{L_i} = \pi_{L_{i-1}} \cdot p_{L_{i-1}, L_i}, i \neq 0, k. \quad (3.14)$$

E para o estado final

$$\pi_{L_k} = \pi_{L_k} \cdot p_{L_k, L_k} + \pi_{L_{k-1}} \cdot p_{L_{k-1}, L_k}. \quad (3.15)$$

Ainda, esse modelo foi avaliado usando três diferentes distribuições de probabilidade para transição entre os estados: uniforme, Gaussiana e exponencial. Os cenários considerados foram em redes 6LoWPAN e Wi-Fi. Em ambos os cenários, o PLR foi calculado variando as probabilidades de transição entre os estados com os seguintes resultados observados:

- Distribuição de probabilidade de transição constante: se a probabilidade de transição bem-sucedida for próxima de 100%, então a PLR é baixo. Caso contrário, uma probabilidade de transição bem-sucedida menor (1%) resulta em um PLR maior
- Distribuição de probabilidade de transição gaussiana: se a probabilidade de transição bem-sucedida for maior, então PLR é menor. Nesse caso, o PLR mínimo é inferior a 2%. Por outro lado, o maior valor de PLR ocorre quando a probabilidade de transição bem-sucedida tem o menor valor. Nessa situação, a perda de pacotes é de aproximadamente 45%.
- Distribuição de probabilidade de transição exponencial: um PLR máximo de 30% foi identificado para uma probabilidade de transição de perda próxima de 100% e a menor probabilidade de transição bem-sucedida. Por outro lado, uma probabilidade mínima de transição de perda de pacote e um valor máximo de probabilidade de transição bem-sucedida resultam em um PLR próximo a 0,4%.

3.5 Gilbert-Elliot com Sub-estados

Wolter et al., (2012) argumentam que o modelo clássico de Gilbert-Elliot assume uma interferência em rajada no qual não representa muito bem o tempo de duração de um estado. Desta forma, os autores propõem um modelo de perda baseado em uma distribuição em fase (*PH-distribution*) porém a um alto custo computacional e ainda há a introdução de mais estados (WOLTER et al., 2012). Para resolver o problema da complexidade e o acréscimo de estados adicionais, Feng, Liu e Ji (2014) sugerem a implantação de um modelo com sub-estados, como mostrado na Figura 3.7.

O modelo de múltiplos estados apresentado é baseado em um estado Bom formado por quatro sub-estados chamados de $\{G_1, G_2, G_3, G_4\}$. A probabilidade de perda de pacotes p_i neste modelo é aplicada para cada sub-estado do estado Bom, com transição para o estado Ruim. No estado Bom, todo pacote é recebido o que significa

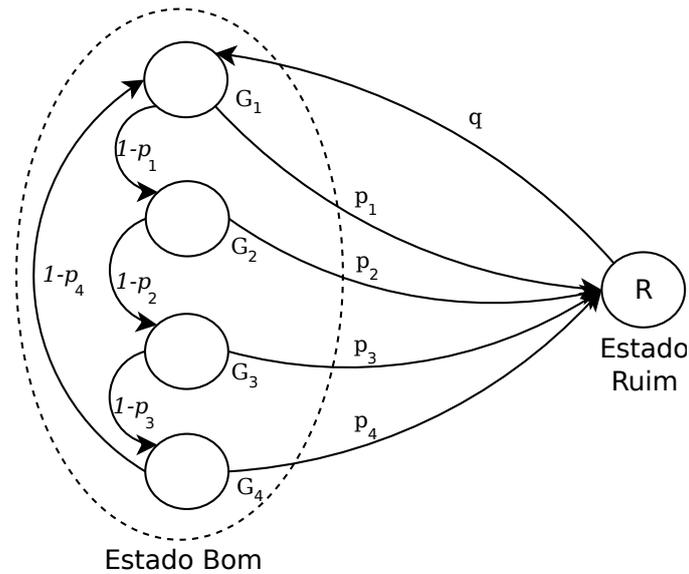


Figura 3.7: Modelo de Gilbert-Elliot com sub-estados (FENG; LIU; JI, 2014).

que a taxa de perda de pacotes no modelo de Gilbert-Elliot é igual a zero. No estado Ruim, apenas é possível transitar para o estado G_1 com probabilidade q . Esse modelo de sub-estados é comparado com o modelo clássico de Gilbert-Elliot e a distribuição PH. Resultados mostraram que o modelo proposto representa melhor comportamento em relação ao modelo clássico de Gilbert-Elliot e quando usada distribuição PH com 4 e 25 estados.

3.6 Markov Hierárquico (hMM)

Métodos de modelagem hierárquica são baseados na independência entre as rajadas e seções aleatórias. Sub-modelos de rajadas e sub-modelos aleatórios são construídos a partir de distribuições de perdas e são interconectados a partir de probabilidades de transição entre os estados (TAO; LU; CHUANG, 2001). O modelo de Markov hierárquico (hMM, *Hierarchical Markov Model*) apresenta duas camadas hierárquicas de Markov, onde a primeira é apresentada como sub-modelo de rajada e a segunda como um sub-modelo aleatório para canais tipicamente com desvanecimento Rayleigh para curtas distâncias. Esse modelo é apresentado na Figura 3.8

No sub-modelo em rajadas as perdas são representadas no estado E_1 e a correta

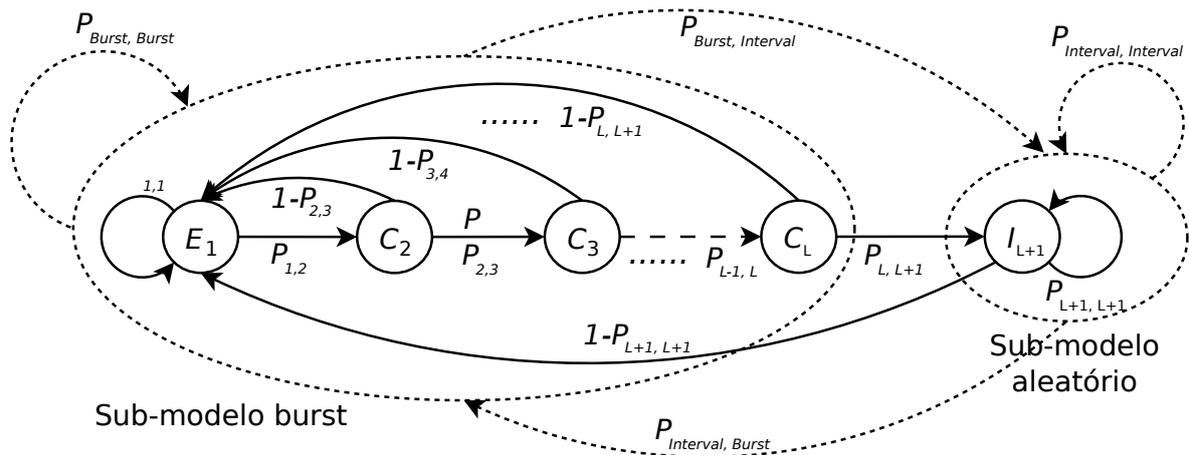


Figura 3.8: Modelo hMM de duas camadas (TAO; LU; CHUANG, 2001).

recepção dos pacotes apresentada pelos estados $C_i (i = 2, 3, \dots, L)$. O sub-modelo aleatório consiste em um único estado simétrico binário que gera erros aleatórios de baixa probabilidade. Resultados mostram a comparação entre dados simulados e o modelo hMM em termos de distribuição acumulada do tamanho de rajadas e seções aleatórias. Ao separar as rajadas de erro das seções aleatórias é possível verificar que ambas são independentes e que o modelo pode descrever estaticamente o canal de desvanecimento de Rayleigh em termos de distribuição acumulada de tamanho de rajadas e seções aleatórias. O hMM ainda apresentou melhorias na caracterização dos parâmetros das rajadas de erro quando comparadas a outros modelos como o modelo simplificado de Fritchman² (SFM, *Simplified Fritchman Model*) (TAO; LU; CHUANG, 2001).

3.7 Canal de Markov com Estados Finitos (FSMC)

O modelo de estados finitos chamado de FSMC (*Finite-State Markov Channel*) é proposto por Wang e Moayeri (1995) e é apresentado como um canal de desvanecimento de Rayleigh para redes sem fio. O PLR do canal de transmissão é calculado a

²O modelo de Fritchman, também conhecido como modelo N-estados particionados de Fritchman, é uma generalização do modelo de Gilbert para N estados (FRITCHMAN, 1967). Neste modelo os estados são divididos em k estados de ocorrência de erros e $N - k$ estados livres de erros, onde as transições e as probabilidades só podem ocorrer entre os estados (FERNANDES, 2002).

partir da SNR obtida do canal e as transições de estado são apresentadas na Figura 3.9.

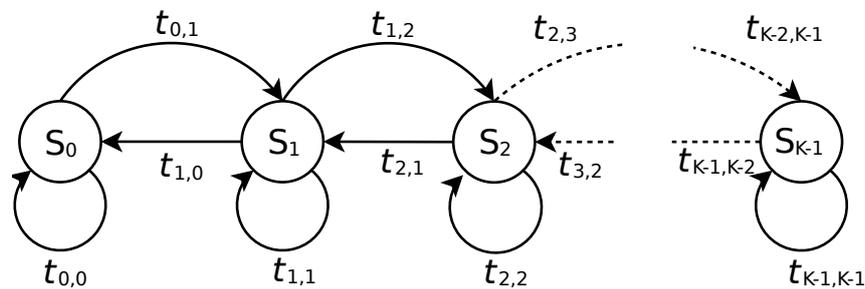


Figura 3.9: Modelo de cadeias de Markov com estados finitos para um canal de desvanecimento Rayleigh (WANG; MOAYERI, 1995).

Esse modelo permite analisar o efeito do desvanecimento para a eficiência da codificação e a capacidade do canal a ser investigado. E ainda, o efeito de particionamento da SNR tem impacto direto no modelo FSMC e na probabilidade de perdas. Uma solução para esse problema do particionamento da SNR é proposto por Hassan, Krunz e Matta (2004), onde para cada faixa de SNR é aplicado a um estado diferente onde é analisada a perda de pacotes e o desempenho do atraso em redes sem fio (HASSAN; KRUNZ; MATTA, 2004). Resultados mostraram que o número de estados do modelo Markoviano para um FSMC depende do esquema de modulação de cada canal (FSK- *frequency shift keying*, DPSK-*differential phase shift keying* ou BPSK-*binary phase shift keying*), da média da SNR e do intervalo (*thresholds*) do BER. Modelos de alta ordem como cadeias FSM (*Full-state Markov*) são aplicadas para modelar erros e perdas em redes de comunicação sem fio como proposto por Khayam e Radha (2006), onde apresentaram melhor aproximação de variação do tempo comparada a amostras reais para altas ordens de 4, 16 e 4096 estados (KHAYAM; RADHA, 2006; KHAYAM et al., 2007).

3.8 Markov Escondido (HMM)

Um modelo de Markov escondido (HMM, *Hidden Markov Model*) é um processo duplamente estocástico, constituído por um processo também estocástico subjacente

que não é observável diretamente (conceito escondido - *hidden*), mas que só pode ser observado através de outro conjunto de processos estocásticos que produzem a sequência de ocorrências observadas (RABINER; JUANG, 1986). Desta forma, o primeiro processo é uma cadeia de Markov e o segundo é um processo de observação a partir de uma distribuição qualquer e é totalmente determinada pelo estado atual da cadeia de Markov (SILVEIRA; SILVA, 2012).

Em (CARDOSO; REZENDE, 2009) é proposto um modelo generalizado do modelo de Gilbert-Elliot com três estados com duas estruturas diferentes: uma geral com transição entre todos os pares de estados (HMM3g), e uma *birth-death* com transições 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 do processo de perdas se aproximou mais da ACF dos dados amostrados e coletados de um ambiente real. Esse comportamento é melhor explicado devido a natureza da perda de pacotes que mudam rapidamente, mas não de forma abrupta (CARDOSO; REZENDE, 2009). Comparado ao modelo de Gilbert-Elliot, o HMM3bd também apresentou melhores resultados em relação a ACF e a função de distribuição acumulada complementar do tamanho das rajadas de perdas. Contudo, o HMM3bd ainda não é suficiente para descrever adequadamente o processo de perda, sendo necessário aumentar o número de estados de maneira a aumentar a precisão do modelo. Esse aumento acarreta na oscilação do processo de convergência a um resultado ideal, necessitando de uma investigação de número de estados adequado.

Uma adaptação baseada no acréscimo do modelo HMM em processos incorporados duplos é chamado de DEPHMM (*double embedded process HMM*) (SALIH et al., 2009). O DEPHMM é apresentado na Figura 3.10, onde para obter o comportamento das rajadas de perdas desse modelo basta isolar o estado (0) que indica o único estado livre de erros de fácil criação. Assim, o número de pacotes perdidos é variado de acordo com o tamanho das rajadas de perda que conseqüentemente torna

o DEPHMM mais complexo para ser modelado.

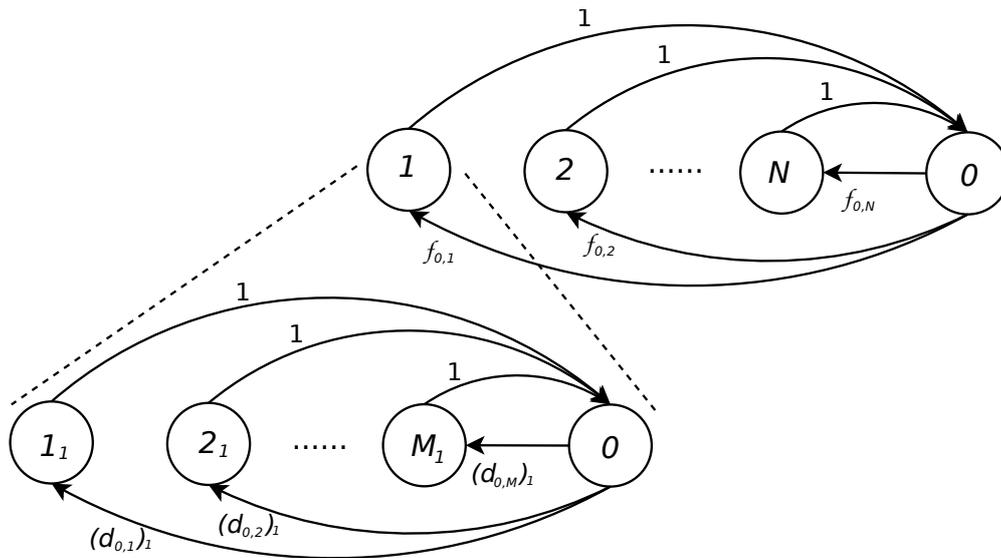


Figura 3.10: Modelo DEPHMM.

O DEPHMM é comparado ao modelo DPBGM (*Deterministic Process Based Generative Model*) e ao modelo FSMC. O DPBGM proposto em (WANG; XU, 2005) e melhorado também em (WANG; XU, 2007) descrevem uma nova classe de modelos generativos usando um processo determinístico adequadamente parametrizado e amostrado para caracterização de rajadas de erro. Comparado ao DEPHMM, o modelo FSMC errou na maioria das estimativas das falhas em rajadas e o DPBGM aproximou consideravelmente as estimativas de falhas porém não superior.

Uma simples HSMM (*Hidden Semi-Markov Model*) derivada do HMM considera o processo de perda de pacotes como um processo *On-Off*, que permite a extração dos comprimentos de estados com sucessivas chegadas de pacotes a partir de distribuições não exponenciais (diferente do modelo HMM em que são obtidos de distribuições exponenciais) (NGUYEN; ROUGHAN, 2013).

Em redes sem fio, como no padrão IEEE 802.11b, o HMM é usado na caracterização das perdas de quadros termos de estatística (GÓMEZ et al., 2012). Para redes *indoor* que utilizam o padrão 802.11, Hartwell e Fapojuwo (2004) sugerem a utilização do modelo FSMM de cinco estados para a criação de um modelo de erros de quadros, onde resultados mostram que modelos de alta ordem treinados com o algoritmo

de Baum-Welch superam o desempenho do modelo de Gilbert (HARTWELL; FAPOJUWO, 2004). Resultados também 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.

3.9 Markov de Quatro Estados (4SMM)

Para Yu e Miller (2007) uma 4SMM (*Four-state Markov Model*) apresenta melhor desempenho em relação a modelos 2SMM, onde resultados de simulações comparados a medições reais coletas em uma rede 802.11 são mais próximas para os modelos com quatro estados. O modelo 4SMM é apresentado na Figura 3.11.

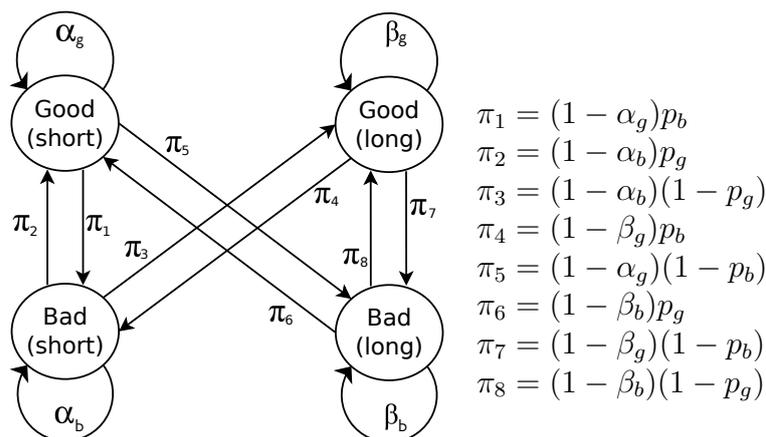


Figura 3.11: Modelo de quatro estados 4SMM (YU; MILLER, 2007).

Nesta expressão, α e β são usadas para capturar a inclinação inicial (cauda) da distribuição e descreve o comportamento curto ou longo de ocorrências, enquanto $p(0 \leq p \leq 1)$ é o parâmetro que controla a frequência relativa dessas ocorrências. A distribuição de ocorrência de estados Bons, $f_g(k)$, é governada por um conjunto de três parâmetros, $\{\alpha_g, \beta_g, p_g\}$ e, da mesma forma, a distribuição de ocorrência de erros, $f_b(k)$, é governada por um conjunto similar de parâmetros para o estado ruim, $\{\alpha_b, \beta_b, p_b\}$.

O 4SMM utiliza distribuições para caracterizar a duração de estados Bons e Ruins

relativos a intervalos de valores da SNR decorrente do desvanecimento do canal de transmissão sem fio para uma rede IEEE 802.11a. Ao comparar o 4SMM com o 2SMM e o FSMC, o mesmo resulta na redução da complexidade do modelo e obtém resultados razoavelmente melhores. O 2SMM apresenta baixa complexidade porém não consegue retratar efetivamente a distribuição da duração dos estados Bons e Ruins, diferente do FSMC que reproduz corretamente as distribuições porém a uma complexidade maior que o 4SMM.

3.10 Distribuição de Séries Logarítmicas

Em sistemas em tempo real, Carvalho, Angeja e Navarro (2005) propõe modelar a duração das rajadas de chegadas e perdas de pacotes transmitidos sobre a pilha de protocolos UDP/IP/802.11g. Para isso é realizado um teste de ajuste no modelo de Gilbert-Elliot e proposto uma distribuição de séries logarítmicas para um conjunto de dados experimentais.

O modelo considera X_ℓ como sendo o comprimento da rajada de n pacotes perdidos e supõe que X_ℓ é uma variável aleatória de série logarítmica. A função de massa de probabilidade (PMF, *Probability Mass Function*) de X_ℓ é dada por $P[X_\ell = n] = -\frac{\theta^n}{n \ln(1-\theta)}$, onde θ é o parâmetro da distribuição que deve ser estimado a partir de uma amostra de dados. O PLR é dado por

$$\text{PLR} = \frac{1}{\frac{P[X_\ell=0]}{E[X_\ell](1-P[X_\ell=0])} + 1} . \quad (15)$$

Resultados de testes de verificações e comparações com o modelo de primeira ordem de Gilbert-Elliot mostraram que o modelo proposto apresenta melhor acurácia na representação de dados coletados em uma rede real.

3.11 Comparação dos modelos de perdas em redes sem fio

Nesta seção são apresentadas as principais diferenças entre os modelos de perdas conhecidos. Na Tabela 3.2 esses modelos são apresentados quanto a sua classificação em função de diversos aspectos. Na coluna *métricas* são indicados os principais parâmetros e características observadas em cada modelo analisado, onde se observou a grande incidência da análise das perdas em rajadas. Alguns parâmetros não foram inseridos devido a sua pouca ou única aplicação no modelo analisado, como por exemplo, a largura de banda do canal, utilização do canal, interferência de rádio e latência. A coluna *validação* apresenta as condições ao qual cada modelo foi validado, sendo a grande maioria realizada através de simulações computacionais, aplicação de técnicas analíticas ou testes com ambientes e condições reais de transmissão de dados em redes sem fio.

Conforme apresentado neste capítulo, o uso de cadeias de Markov é bastante comum na modelagem de perdas. Assim, nesta coluna é descrito o tipo de cadeia de Markov utilizado, classificado em dois estados, com um número finito mas variável de estados e os modelos com número fixo de estados porém maior que dois. Nos casos onde esta informação não é apresentada o modelo não utilizou cadeias de Markov para a modelagem da perda de pacotes. A coluna de *redes de comunicação* aborda o tipo de cenário que foi utilizado para o desenvolvimento do modelo. Neste caso, alguns trabalhos não foram claros quanto à aplicação específica do modelo para redes WiFi IEEE 802.11 e sim apenas para comunicações *wireless* o que abre discussão quanto a efetiva acurácia do modelo para qualquer rede sem fio. A última coluna, indica a utilização de amostras reais em redes sem fio para realização e teste dos modelos. Esse item é importante para o desenvolvimento de um modelo de perda, pois a partir dele é possível averiguar o real comportamento destas redes.

Tabela 3.2: Comparação modelos de perdas de pacotes em redes sem fio (SILVA; PEDROSO, 2019)

Referência	Métricas						Validação			Cadeias de Markov			Rede de comunicação		Amostras	
	Atraso/Jitter	SNR	Throughput	Distância entre dispositivos	Tempo entre chegadas pacotes	Tamanho Payload de dados	Perdas em rajadas	Testes	Simulação	Análítico	2-estados	Estados variados e finitos	Estados fixos e não 2-estados	Padrão IEEE 802.11	Wireless em geral	Uso de dados reais
(WANG; MOAYERI, 1995)		✓					✓	✓	✓	✓	✓				✓	
(KRUNZ; KIM, 2001)	✓						✓	✓	✓	✓	✓				✓	
(LEE; CHANSON, 2002)	✓				✓			✓		✓					✓	
(TANG; MCKINLEY, 2003)			✓	✓			✓	✓	✓	✓	✓			✓		✓
(KARANDE et al., 2003)			✓				✓	✓	✓	✓				✓		✓
(ARAUZ; KRISHNAMURTHY, 2003)		✓	✓				✓	✓	✓	✓	✓			✓		
(LEE; CHANSON, 2004)	✓				✓		✓	✓	✓	✓				✓		
(HARTWELL; FAPOJUWO, 2004)			✓			✓	✓	✓	✓	✓		✓			✓	
(CARVALHO; ANGEJA; NAVARRO, 2005)			✓		✓		✓	✓	✓	✓				✓		✓
(YU; MILLER, 2007)		✓	✓			✓	✓	✓	✓	✓		✓		✓		✓
(HAN; LEE, 2007)						✓	✓	✓		✓				✓		✓
(WANG; XU, 2007)							✓	✓	✓						✓	
(CARDOSO; REZENDE, 2009)							✓	✓	✓			✓		✓		
(XIE et al., 2009)	✓			✓				✓	✓					✓		
(YOUNESIAN et al., 2014)							✓	✓	✓		✓			✓		✓
(FENG; LIU; JI, 2014)							✓	✓	✓	✓					✓	

CAPÍTULO 4

MODELO PROPOSTO

4.1 Problema e lacunas em aberto

Um dos desafios na modelagem das perdas de pacotes em redes WiFi é a representação correta dessas perdas sob condições semelhantes à aquelas onde essas redes geralmente operam. Desta forma, é primordial não apenas observar o resultado das perdas como elemento principal no processo de modelagem desse comportamento, mas também observar e considerar o cenário crítico ou não, controlável ou não no qual essas redes estão condicionadas a operação. Assim, considerar o número ideal de variáveis para representar o cenário semelhante ao real onde ocorrem as perdas é ponto chave para o desenvolvimento de um modelo de apresentação desse comportamento. Em redes WiFi, diversas variáveis podem influenciar nas perdas, sendo essas variáveis classificadas como de origem física ou da rede de enlace. Ao não considerar simultaneamente essas variáveis, os modelos tornam-se menos precisos por ignorar aspectos relevantes que influenciam a perda de pacotes. Partindo desse pressuposto, questiona-se sobre como desenvolver um modelo de perdas de pacotes para redes WiFi que considere simultaneamente fatores de camada física e de enlace para representar as perdas.

O problema é que os modelos atuais não consideram simultaneamente variáveis físicas e de enlace na modelagem das perdas de pacotes. Os modelos, na sua grande maioria, consideram apenas variáveis físicas como a SNR (ARAUZ; KRISHNAMURTHY, 2003; YU; MILLER, 2007; WANG; MOAYERI, 1995) ou a distância dos dispositivos, ou variáveis de enlace como o número de usuários que competem ao meio de acesso. Nesta primeira abordagem, os modelos baseado apenas em variáveis de

camada física deixam de capturar as condições de disputa no meio de acesso, sendo que o mesmo ocorre nos modelos que deixam de considerar as variáveis de camada física.

Ainda, há o conjunto de modelos que utilizam séries de dados ou sequências de perdas de pacotes para a sua parametrização, como por exemplo os modelos baseados apenas em cadeias de Markov como o modelo de GE e (FENG; LIU; JI, 2014; SANNECK; CARLE, 1999; TAO; LU; CHUANG, 2001; YOUNESIAN et al., 2014). Nos modelos baseados em cadeias de Markov, o comprimento das rajadas de perdas não apresentam autocorrelação temporal, o que já é observado em redes WiFi (RUSS; HAGHANI, 2009). Contudo, modelos que sugerem o uso de distribuições de cauda pesada para capturar a autocorrelação temporal das perdas não consideram o uso simultâneo de variáveis de camada física e de enlace na modelagem. Nos modelos que utilizam o HMM na modelagem de perdas em redes WiFi como em (CARDOSO; REZENDE, 2009; SALIH et al., 2009; NGUYEN; ROUGHAN, 2013; GÓMEZ et al., 2012; HARTWELL; FAPOJUWO, 2004) observa-se a ausência do uso das características físicas e de acesso do meio. Nestes modelos, são considerados unicamente as séries de perdas de pacotes observadas.

4.2 Metodologia

Neste capítulo é apresentado um novo modelo de perda de pacotes, bem como a metodologia utilizada. A sequência de atividades realizadas durante o processo de desenvolvimento e validação do modelo são apresentadas na Figura 4.1. Cada etapa é apresentada individualmente em seção específica dentro deste capítulo.

Inicialmente são apresentadas as características do cenário utilizado para coleta das amostras de dados utilizadas. As amostras de tráfego real foram usadas para identificar as variáveis a serem usadas para caracterizar o comportamento de perdas de forma mais precisa (KUHN; JOHNSON, 2013).

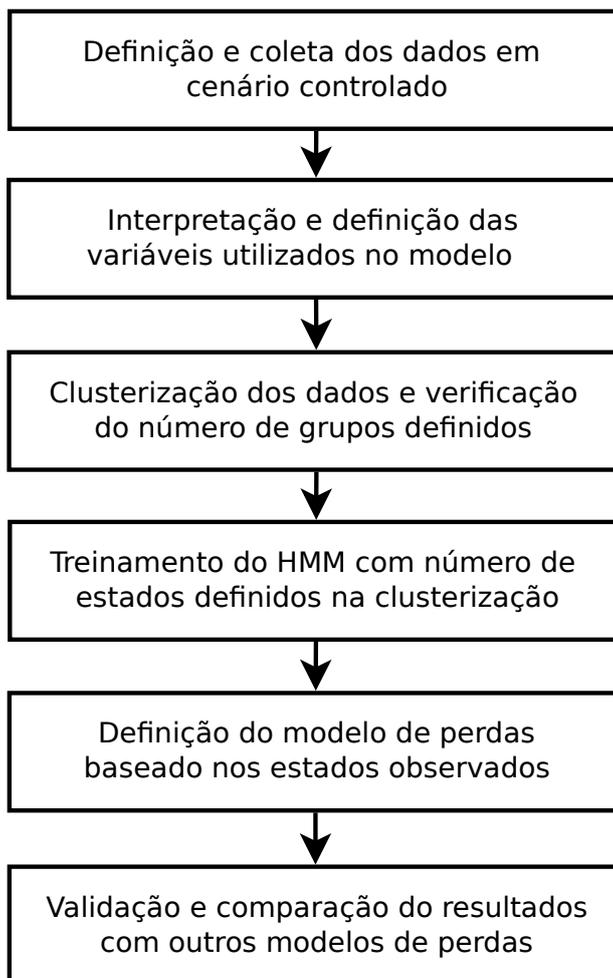


Figura 4.1: Sequência de atividades para realização do trabalho.

Em seguida, as variáveis coletadas no cenário são interpretadas e a partir de análises específicas são definidas quais das variáveis candidatas são utilizadas na definição do modelo de perdas. Um dos grandes desafios foi selecionar as variáveis e correspondentes parâmetros relevantes que permitisse confiar que o modelo construído pudesse descrever o comportamento de perdas com um grau de precisão semelhante na amostra de validação, coletada de forma independente da amostra usada na criação do modelo. Remover variáveis irrelevantes ou redundantes é importante para tornar o modelo mais simples e conciso, reduzindo a complexidade computacional da parametrização e também no uso do modelo para simulações computacionais. A identificação das variáveis relevantes, irrelevantes e redundantes foi realizada através da utilização da correlação computacional entre as variáveis.

Após definidas as variáveis a serem usadas no modelo de perdas baseado no HMM, estas foram agrupadas a partir de técnicas de clusterização para identificação do número de grupos de similaridade a partir de dendrogramas. O dendrograma é um diagrama de árvore que exhibe os grupos formados por agrupamento de observações em cada passo e em seus níveis de similaridade. A quantidade de grupos prováveis foi utilizado como o número de estados para treinamento do HMM. O treinamento não supervisionado da HMM detecta similaridade e anomalias, agrupando os dados a partir de características observáveis. Uma análise posterior permite que os estados identificados sejam interpretados e as rajadas de perda modeladas.

A partir dos estados identificados após o treinamento do HMM, foi identificado e interpretado o comportamento das perdas dentro de cada estado. Em seguida foram definidas as distribuições de probabilidade que melhor representassem esse comportamento. A validação do modelo foi avaliada em amostras que não foram usadas para construir ou ajustar o modelo, de modo que forneçam um senso imparcial de eficácia do modelo (KUHN; JOHNSON, 2013). Os resultados serão comparados com os modelos mais atuais de perdas disponíveis na literatura.

4.3 Cenário de coleta de dados

A análise do comportamento de perda de pacotes em redes sem fio IEEE 802.11 foi realizada a partir da coleta de amostras reais em um ambiente *indoor*. Na maioria dos casos é difícil determinar um ambiente físico que represente o comportamento real de redes locais sem fio, isto porque em redes internas deve-se considerar fatores como a existência de paredes, portas, pessoas, dispositivos e outros obstáculos que podem afetar as propriedades físicas do canal de transmissão (CARDOSO; REZENDE, 2009). Baseando-se nestas condições, foi utilizado um cenário de transmissão em rede IEEE 802.11 em modo infraestrutura como mostrado na Figura 4.2.

Neste cenário proposto para coleta dos dados, o meio de transmissão utilizado

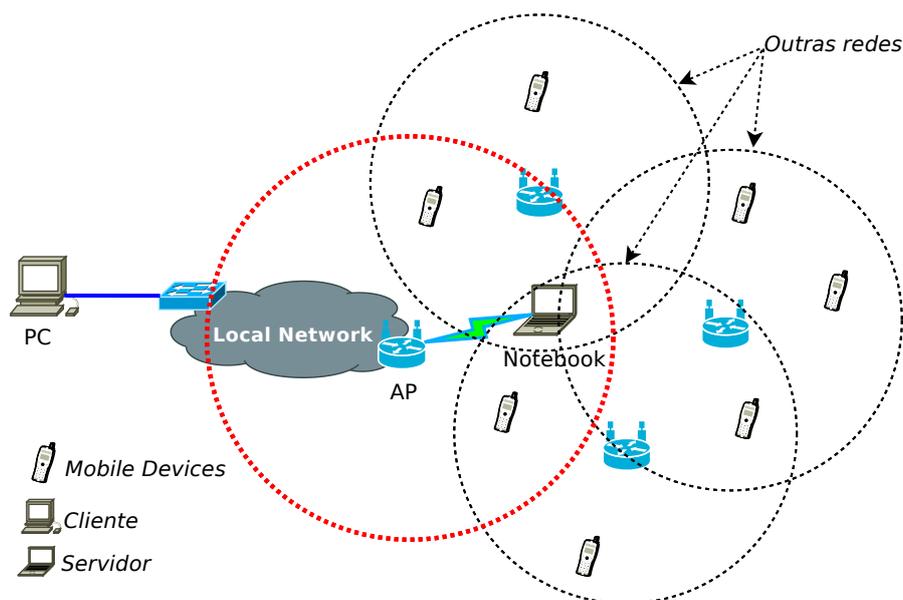


Figura 4.2: Cenário utilizado para coleta de dados em rede sem fio IEEE 802.11.

é compartilhado com diversos dispositivos sem fio (*smartphones*, notebooks e outros *mobile devices*) conectados no mesmo AP principal ou por outros dispositivos conectados em outros APs que compartilham do mesmo canal para transmissão de dados. Nestes casos, pode ocorrer a sobreposição de canais de mesma frequência ocasionando interferências na camada física ou colisões de camada MAC devido a existência de terminais ocultos, como mostrado na seção 2.3.1. As perdas de pacotes ocorrem devido a uma combinação de problemas de camada física e MAC. Assim, essas perdas não podem ser estimadas puramente com base em variáveis de camada física, como por exemplo a SNR (GIUSTINIANO et al., 2010). Neste caso é necessário a inserção de características inerentes à camada MAC. No cenário proposto, não é levado em consideração o valor numérico da distância e das interferências de paredes ou outras características que podem ocasionar o desvanecimento do canal entre o AP e o dispositivo teste, pois a SNR é uma variável afetada por ambos os fatores.

As coletas foram realizadas em dias e horários variados de maneira a capturar diferentes situações de comportamento da rede em estudo. Ainda, para capturar as variações nos parâmetros de camada física do canal de transmissão, o dispositivo de teste (*notebook*) com interface Wi-Fi conectado à rede foi movido em diferentes

pontos de distância em relação ao AP durante a coleta das amostras. Os enlaces onde foram realizadas as coletas se encontravam em constante utilização, apresentando sempre uma quantidade variada de usuários conectados no AP ou em outros APs de mesmo canal de transmissão. Em modo infraestrutura, o AP gerenciou os pacotes transmitidos entre o PC conectado à rede Ethernet por um *switch* e o dispositivo de teste conectado à rede sem fio. A configuração do AP é apresentada na Tabela 4.1.

Tabela 4.1: Características de configuração do AP.

Parâmetro do AP	Descrição
Fabricante/Modelo	TP-Link TD-W8961N
Características	300Mbps Wireless N ADSL2+ Modem Router
Modo de operação	Bridge
Versão IP	IPv4
Serviço DHCP	Desabilitado
Canal	Fixo para cada coleta (1, 6 ou 11)
Intervalo de Beacon	100 milisegundos
RTS/CTS Threshold	2347 bytes
Threshold de fragmentação	2346 bytes
Protocolo Wireless	IEEE 802.11b+g+n
Largura de banda canal	11b/g - 20 MHz 11n - 20 ou 40 MHz

O fluxo de dados transmitido é enviado do PC configurado como cliente a partir de um aplicação cliente-servidor UDP (*User Datagram Protocol*). O dispositivo de teste configurado como servidor recebeu o fluxo de dados *unicast* de forma que não ocasionasse congestionamento na rede. Os pacotes foram gerados em uma taxa de mil pacotes por segundo com tamanho fixo de 162 bytes (128 bytes do cabeçalho UDP, protocolo IP e *payload*, e por fim 34 bytes do protocolo IEEE 802.11), o que não afetou significativamente o desempenho da rede sob o protocolo IEEE 802.11b/g/n utilizado nas coletas de dados. Retransmissões decorrentes de perdas de pacotes não foram consideradas neste processo de coleta para análise.

Foram coletados aproximadamente 7 horas de transmissões. Os parâmetros foram coletados a partir da observação dos canais de transmissão utilizando o analisador de redes sem fio AirCheck™ Wi-Fi Tester desenvolvido pela Fluke Networks. As características do sinal de transmissão foram obtidas por um *bash script* no *notebook*

(Sistemas Operacional Linux Ubuntu 18.4). O esquemático que representa o cenário de coleta dos dados é apresentado na Figura 4.3.

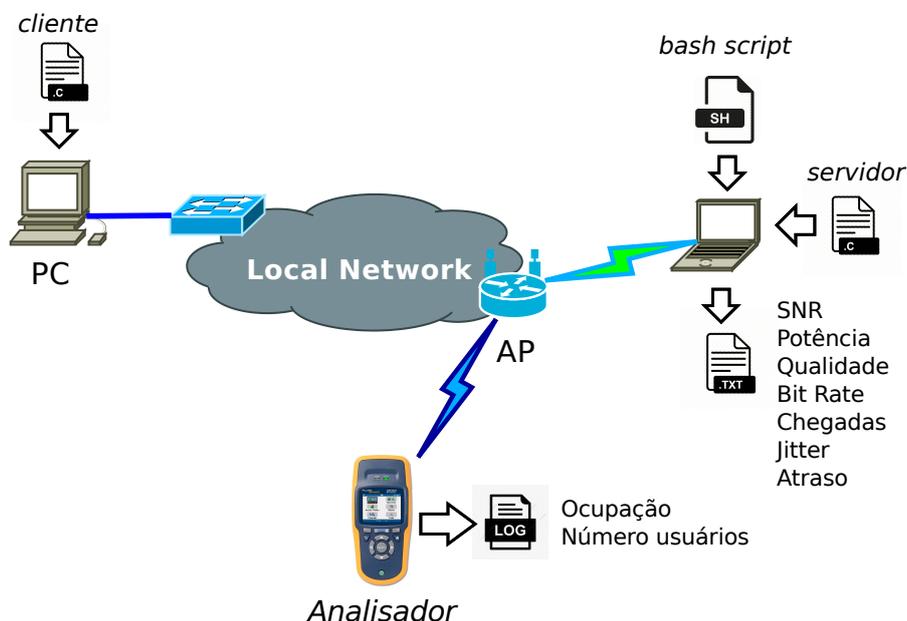


Figura 4.3: Cenário de coleta das amostras e arquivos de saída.

Foram coletadas informações da SNR (dB), da potência do sinal (dBm), da taxa de transmissão (Mbps) e da qualidade do canal (0 até 1). A SNR é relação entre a potência do sinal e a potência do ruído observada no receptor. A qualidade do canal é uma razão entre a maior SNR possível e a SNR instantânea. Na prática essa razão tem influência direta da SNR que é usada na sua estimativa.

O analisador de rede permitiu a coleta dos parâmetros de ocupação (%) do canal de transmissão e do número de usuários conectados no AP e nos demais APs que utilizam o mesmo canal de transmissão. Em uma rede WiFi com padrão 802.11b/g/n, quando um elemento decide usar o meio ele faz a transmissão usando todas as subportadoras do canal. A ocupação do canal é definida pela fração de tempo que o canal está ocupado (FUKUHARA et al., 2009). A ocupação do canal é resultado direto do algoritmo de controle de acesso ao meio. Essa ocupação é impactada pelos dispositivos conectados no canal observado, assim como também o comportamento desses dispositivos em relação a quantidade de dados que está sendo transmitida por eles.

Quando é realizado o armazenamento dos parâmetros da rede em um *log* de registro, o equipamento apenas registra uma média desde o momento que o analisador é ligado.

A estimação da ocupação foi realizada com o equipamento Fluke Air Check em períodos de 10 segundos de coleta. As demais variáveis também foram estimadas durante o período de 10 segundos para permitir a análise conjunta.

A Figura 4.4 apresenta uma amostra coletada em 10 minutos de captura de SNR, qualidade do sinal, potência do sinal e taxa de transmissão da interface de rede sem fio. Esses valores são obtidos de arquivos específicos do sistema operacional do dispositivo teste que calcula uma estimativa desses valores. Nesta amostra, inicialmente o receptor foi posicionado ao lado do AP e afastado até um limite máximo que permitisse que o receptor continuasse conectado ao AP. Esse distanciamento se reflete na qualidade do sinal ilustrado na Figura 4.4-B, que mostra uma qualidade superior à 90% no início da amostra e qualidades inferiores à 70% no decorrer da coleta. Em algumas amostras, ao final da coleta, o receptor foi movido para próximo do AP, no qual apresentou uma melhora na qualidade, potência do sinal e na SNR.

A máxima taxa de transmissão na interface de rede também variou devido a estimativa da taxa baseado na variação da SNR, como indicado na Figura 4.4-D. A taxa de transmissão está relacionada ao tipo de modulação do sinal utilizado para transmitir os bits no canal de transmissão (por exemplo, BPSK, QPSK, 16-QAM, 64-QAM) e a taxa de codificação FEC (*forward error correction*) (por exemplo, 1/2, 2/3, 3/4, 5/6). A utilização do tipo de modulação é baseado na estimação da razão sinal-ruído do canal de transmissão, no qual para um canal de transmissão de SNR alto permite o envio de mais bits por símbolo, aumentando assim a taxa de transmissão. Por exemplo, para um link 802.11n de baixa SNR utiliza modulação BPSK a uma taxa de FEC de 1/2, no qual transmite respectivamente para 20 e 40 MHz de largura de banda, 7,2 e 15 Mbps. Em outro caso, para um SNR alto que utiliza modulação 64QAM a uma taxa de FEC de 5/6 permite transmissão de 72,2 e 150 Mbps, respectivamente para 20 e 40 MHz

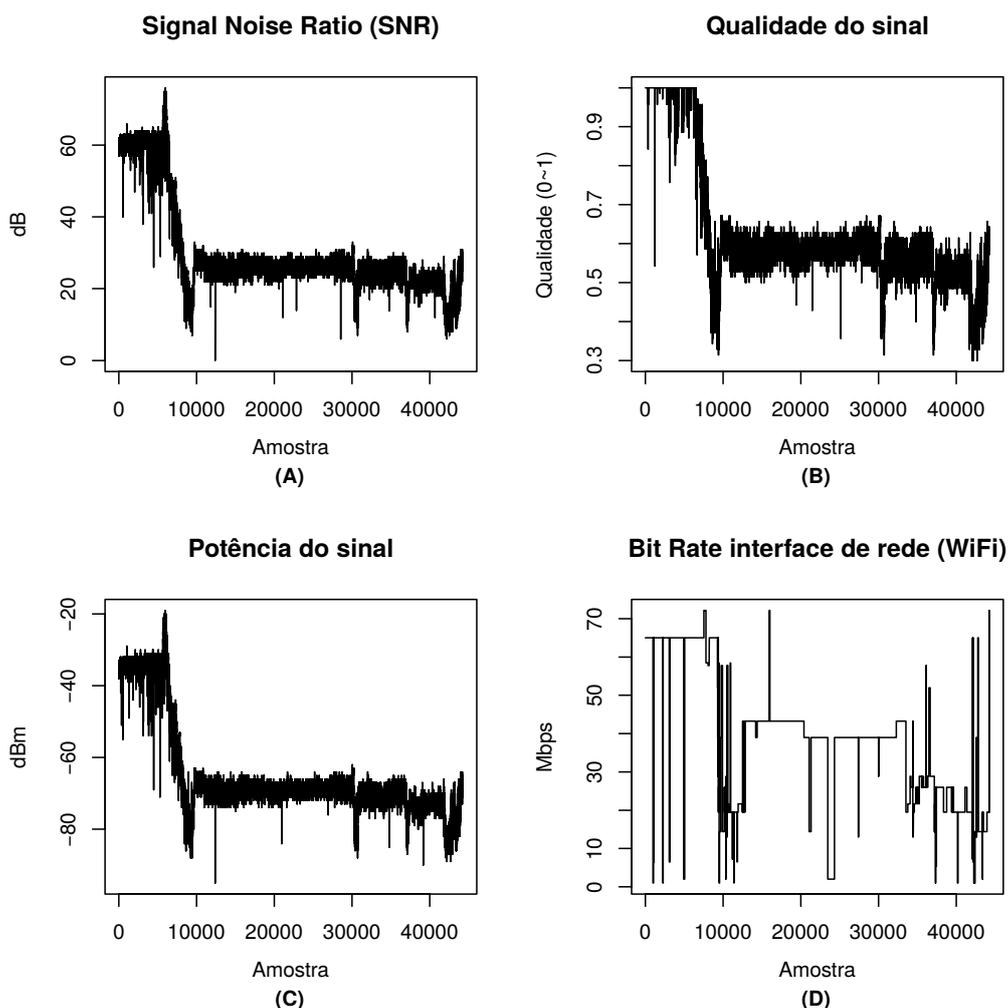


Figura 4.4: Dados coletados na interface WiFi do Notebook: (A) SNR, (B) qualidade do sinal, (C) potência do sinal e (D) taxa de transmissão.

de largura de banda (KARMAKAR; CHATTOPADHYAY; CHAKRABORTY, 2017).

Ainda para a mesma amostra coletada da Figura 4.4, é apresentado na Figura 4.5 a ocupação e o número de usuários do canal utilizado para realizar a transmissão dos dados da aplicação cliente-servidor. Como já indicado anteriormente, devido a limitações do analisador de redes sem fio foi necessário realizar uma amostragem a cada 10 segundos. Desta forma, os valores apresentados correspondem a média calculada no analisador em cada janela de amostragem.

A identificação das perdas de pacotes das amostras coletadas foi realizada a partir do registro no receptor dos pacotes recebidos. Cada pacote enviado da origem possui um identificador numérico único e sequencial, no qual quando este pacote é ausente

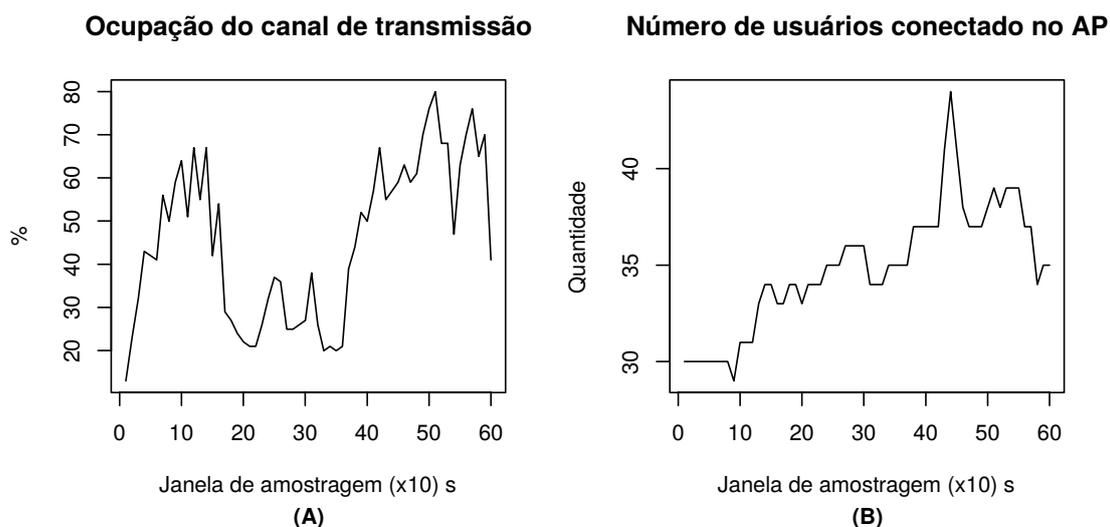


Figura 4.5: Dados coletados no analisador de redes: (A) ocupação do canal e (B) número de usuários conectados no mesmo canal de transmissão.

ou não observado no receptor ele é considerado como perdido. Os pacotes recebidos são identificados por 0 e os perdidos com 1. A Figura 4.6 apresenta a sequência de chegadas e perdas para a mesma amostra realizada na coleta dos dados das Figuras 4.4 e 4.5. Como os valores apresentados na Figura 4.6-A são mostrados individualmente por pacotes, foi calculado a taxa de perdas também em janelas de 10 segundos (Figura 4.6-B), assim como realizado para os valores de ocupação e número de usuários conectados no canal. Essa amostra possui dois momentos de alta concentração de perdas de pacotes, no qual pode-se visualizar observando a Figura 4.6-B com as taxa de perdas superiores a 60%.

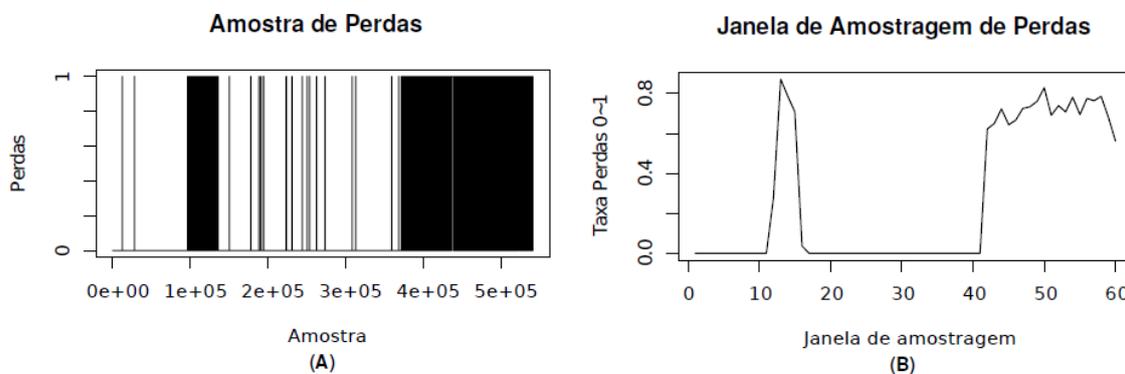


Figura 4.6: Dados observados: (A) sequência das chegadas e perdas de pacotes e (B) taxa de perdas por janela de amostragem.

Antes de realizar o treinamento do HMM a partir dos dados observados que foram coletados durante a aquisição das amostras, foi necessário realizar um estudo sobre quais destas variáveis poderiam ser utilizadas. Na próxima seção são apresentados os testes de aderência e as justificativas para escolha das variáveis relevantes a serem empregadas no modelo proposto.

4.4 Escolha das variáveis

Como já indicado na seção anterior, durante as coletas de dados foram obtidos diversos parâmetros do dispositivo móvel e da rede que serão analisados para utilização no treinamento do HMM. Um dos grandes desafios nesta etapa é definir quais as variáveis são ideais para serem utilizados no modelo de perdas. Algumas abordagens são aplicadas neste processo, no qual as variáveis (em alguns casos, também chamado de características) que foram analisados como candidatos para não compor o modelo são classificados em irrelevantes e redundantes (BERMINGHAM et al., 2015). Os conjuntos de variáveis irrelevantes e redundantes resultam em baixa precisão dos modelos baseados e devem ser removidos dos conjuntos de dados de treinamento para obter um bom desempenho dos modelos. Eliminar variáveis irrelevantes melhora a precisão e reduz o tempo computacional para execução de um modelo proposto, enquanto a remoção de recursos redundantes também reduz o tempo computacional e a precisão do conjunto de dados do modelo (DUANGSOITHONG; WINDEATT, 2009).

Redundante e irrelevante são duas noções bem diferentes, pois uma variável relevante pode ser redundante na presença de outra variável relevante com a qual está fortemente correlacionado. Desta forma, foi definido que em um conjunto de variáveis que apresentam alta correlação entre si podem ser reduzidos a escolha de somente uma variável relevante, isso porque deseja-se observar a relação dessa variável nas perdas que podem ocorrer na transmissão. De outra forma, uma variável pode ser considerada irrelevante quando este dado não possui correlação com o fenômeno ob-

servado que se deseja modelar.

O teste de correlação das variáveis foi realizado a partir do uso de funções disponíveis no Software R (R CORE TEAM, 2019). Essas funções possuem como entrada dois vetores de tamanhos iguais, no qual será testada a correlação entre esses vetores. É possível utilizar diferentes métodos para o teste de correlação, como os coeficientes de Pearson, Kendall ou Spearman.

O método de Pearson utiliza um coeficiente de correlação (ρ) que pode assumir valores no intervalo de -1 à +1, no qual $\rho = +1$ apresenta uma correlação positiva perfeita entre as duas variáveis analisadas, para $\rho = -1$ apresenta uma correlação negativa perfeita e $\rho = 0$ indica que as duas variáveis não dependem linearmente uma da outra. Esse método é comumente aplicado à conjuntos de dados que possuem relação linear direta. O coeficiente ρ de Pearson é obtido com a Equação 4.1

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (4.1)$$

onde, x_1, x_2, \dots, x_n e y_1, y_2, \dots, y_n são os valores das duas variáveis que estão sendo testadas. \bar{x} e \bar{y} são as médias aritméticas de ambas as variáveis, no qual $\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$ e $\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^n y_i$.

O coeficiente de correlação Kendal (τ) avalia o grau de similaridade entre duas variáveis. Para comparar dois conjuntos de dados ordenados, a abordagem do Kendall é contar o número de pares diferentes entre esses dois conjuntos ordenados. Esse número fornece uma distância entre os conjuntos, chamada de distância da diferença simétrica (*symmetric difference distance*). Uma vantagem de τ sobre o coeficiente de Spearman, que será apresentado a seguir, é que τ pode ser generalizado para um coeficiente de correlação parcial. O coeficiente τ de Kendal é expressado pela Equação 4.2 (CHOK, 2010)

$$\tau = \frac{\sum_{i=1}^n \sum_{j=1}^n \text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j)}{n(n-1)} \quad (4.2)$$

onde,

$$\operatorname{sgn}(x_i - x_j) = \begin{cases} 1 & \text{se } (x_i - x_j) > 0 \\ 0 & \text{se } (x_i - x_j) = 0 \\ -1 & \text{se } (x_i - x_j) < 0 \end{cases} \quad (4.3)$$

$$\operatorname{sgn}(y_i - y_j) = \begin{cases} 1 & \text{se } (y_i - y_j) > 0 \\ 0 & \text{se } (y_i - y_j) = 0 \\ -1 & \text{se } (y_i - y_j) < 0 \end{cases} \quad (4.4)$$

Semelhante ao coeficiente de correlação anterior, o τ de Kendall varia de -1 à +1, com o valor absoluto de τ indicando a força de correlação entre as duas variáveis.

O coeficiente de correlação de ordem de classificação de Spearman (φ) é uma versão baseada em classificação do coeficiente de correlação de Pearson. Seu coeficiente de correlação da amostra φ pode ser escrito como mostrado na Equação 4.5:

$$\varphi = \frac{\sum_{i=1}^n ((\operatorname{rank}(x_i) - \overline{\operatorname{rank}(x)}) (\operatorname{rank}(y_i) - \overline{\operatorname{rank}(y)}))}{\sqrt{\sum_{i=1}^n ((\operatorname{rank}(x_i) - \overline{\operatorname{rank}(x)})^2 \sum_{i=1}^n (\operatorname{rank}(y_i) - \overline{\operatorname{rank}(y)})^2)}} \quad (4.5)$$

onde, $\operatorname{rank}(x_i)$ e $\operatorname{rank}(y_i)$ são as *ranks* da observação na amostra. As *ranks* são pesos atribuídos em ordem crescente para os valores dos dados analisados na série, inicialmente também ordenados de forma crescente. Por exemplo, para o menor valor da série x é atribuído o peso 1 e para o maior valor da mesma série é atribuído o peso igual ao número máximo de dados da série x .

Assim como nos dois primeiros modelos de verificação de correlação o coeficiente de correlação de Spearman também varia de -1 à +1 e o valor absoluto de φ descreve a força de correlação entre os dois conjuntos de dados analisados. Quanto mais próximo o valor de 0 for, mais fraca é a correlação entre as duas variáveis. Caso contrário, as duas variáveis são correlacionadas quanto mais próximas à -1 ou à +1

elas forem.

4.4.1 Variáveis de camada física

Dentre as variáveis de camada física, foram analisados a SNR, a qualidade do sinal e a potência do sinal. Foram calculados os valores dos coeficientes de correlação ρ , τ e φ de similaridade dessas variáveis. Para isso foram escolhidas quatro amostras de dez minutos para analisar esses valores. Contudo, foram também testadas outras amostras que também apresentaram resultados semelhantes, no qual cada amostra representa o tempo de dez minutos de coleta. Os resultados dessas quatro amostras são apresentados na Tabela 4.2 e mostram alto nível de similaridade relacionada entre os dados.

Tabela 4.2: Teste de correlação utilizando os métodos Pearson, Kendall e Spearman.

Amostra	Parâmetros comparados	Pearson (ρ)	Kendall (τ)	Spearman (φ)
1	SNR+Qualidade	0,989	0,823	0,902
	SNR+Potência	0,996	0,875	0,931
	Qualidade+Potência	0,989	0,840	0,913
2	SNR+Qualidade	0,990	0,808	0,894
	SNR+Potência	0,989	0,781	0,877
	Qualidade+Potência	0,991	0,816	0,900
3	SNR+Qualidade	0,967	0,879	0,960
	SNR+Potência	0,961	0,863	0,953
	Qualidade+Potência	0,971	0,889	0,964
4	SNR+Qualidade	0,987	0,904	0,963
	SNR+Potência	0,997	0,935	0,976
	Qualidade+Potência	0,986	0,894	0,958

Como observado, os coeficientes de correlação são próximos à +1, principalmente para o método Pearson, o que demonstra que essas três variáveis são fortemente correlacionadas. Ao testar a correlação entre o percentual de perda de pacotes com a SNR foi observado um coeficiente de correlação $\rho < -0,7230$. Esses testes foram realizados em diferentes trechos da amostra coletada, apresentando em alguns deles coeficientes $\rho < -0,9$. Isso demonstra que o valor da SNR possui um impacto inverso nas perdas, no qual quanto maior a SNR menor é o percentual de perdas devido as

melhores condições do canal físico para transmissão dos dados. Desta forma, foi considerado utilizar somente a SNR para treinamento do HMM, dentre estas variáveis que foram testadas inicialmente. Embora essas variáveis estejam correlacionados como apresentado por exemplo entre a potência do sinal e SNR em (ZHU et al., 2004), esses testes de correlação reforçam a escolha das variáveis do modelo a ser apresentado. Esta escolha corrobora não somente pela análise de correlação das variáveis observadas, mas também por outros autores que descrevem a SNR como o atributo principal de um sistema para modelar de perdas de pacotes que ocorrem em rajadas (GUHA; SARKAR, 2008).

Para os valores da taxa de transmissão na interface de rede, os mesmos são estimados em função da largura de banda utilizada e da SNR. Assim, devido ao uso da SNR ter sido definida como variável de entrada do treinamento do HMM a taxa de transmissão não será usada.

4.4.2 Atraso e Jitter

A verificação das variáveis de atraso e jitter foi realizada a partir da coleta de um fluxo de transmissão de pacotes ICMP *echo request (ping)* de 64 bytes a cada 200 milissegundos da origem para o dispositivo destino conectado à rede Wi-Fi. O atraso foi estimado a partir do tempo de duração do envio da requisição ICMP e o recebimento da resposta dessa requisição (*echo reply*) no dispositivo de origem.

O tempo de envio e recebimento da resposta é chamado de RTT (*Round Trip Time*), no qual segundo Mirkovic et al. (2018), pode ser influenciado por três fatores: primeiro, pelo tempo necessário para que o sinal se propague por um meio físico específico de transmissão; segundo, pelo atraso de serialização no qual consiste no tempo necessário para transmitir um pacote de dados para o link; e por último, pelo atraso na fila, definido pelo tempo que um pacote espera na fila nos dispositivos enquanto aguarda que os pacotes à frente sejam desenfileirados e transmitidos (MIR-

KOVIC; ARMITAGE; BRANCH, 2018). O jitter foi obtido a partir da variação do atraso dos pacotes recebidos.

Além dos valores de atraso e jitter obtidos durante a transmissão do ICMP, também foram coletados os pacotes recebidos da transmissão da aplicação cliente-servidor, a SNR no dispositivo destino e a ocupação do canal. O resultado dessa transmissão é apresentado na Figura 4.7. Como mostrado, há três momentos de alta taxa de perdas de pacotes. Essas perdas ocorrem em momentos de baixos níveis de SNR como mostrado na Figura 4.7(B), no qual apresenta mudanças abruptas devido a movimentação dos dispositivos teste durante a coleta de dados. Ao observar o atraso e o jitter, é possível identificar um aumento dessas duas variáveis também para baixos níveis de SNR.

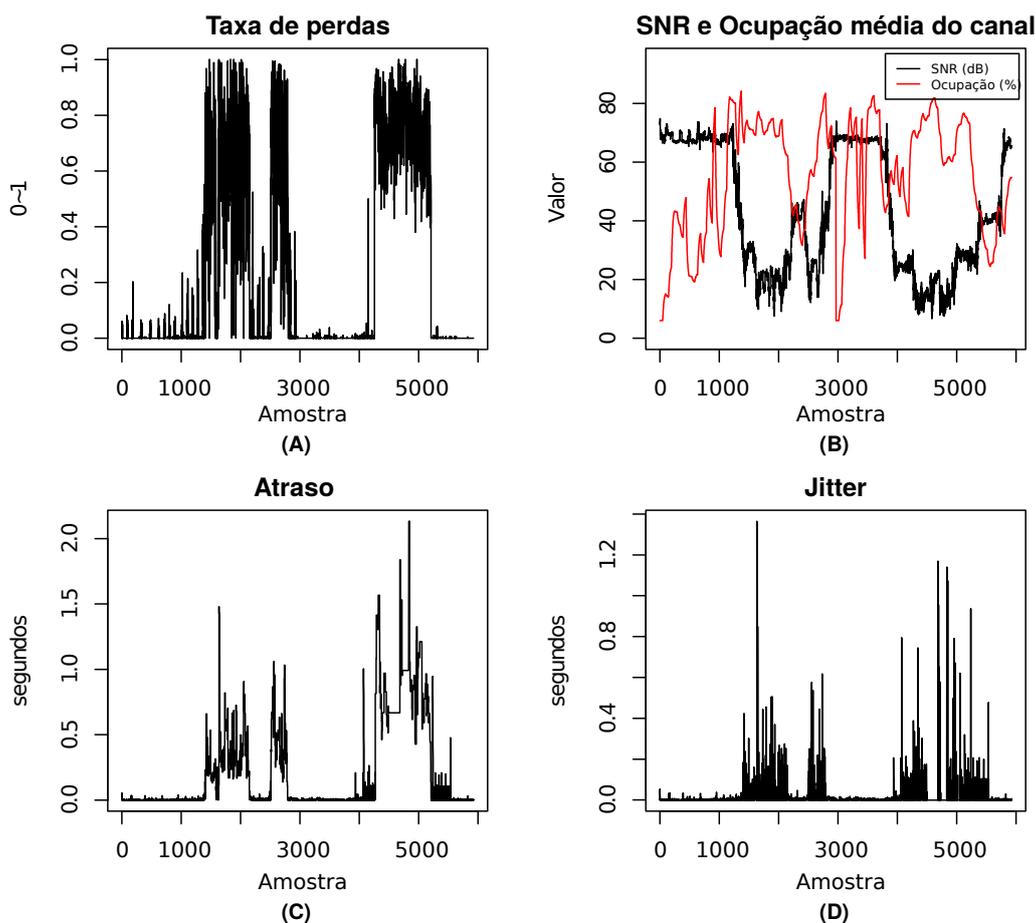


Figura 4.7: Amostra para teste do atraso e jitter: (A) taxa de perdas, (B) SNR e ocupação do canal, (C) atraso fim-a-fim, e (D) jitter.

No IEEE 802.11, a taxa de transmissão depende de muitos fatores, como a lar-

gura de banda do canal, número de fluxos espaciais (*spatial streams*), o intervalo de guarda, a taxa de codificação e o tipo de modulação de rádio frequência. A combinação de todos esses recursos determina a taxa de transmissão na camada física que varia de 6,5 Mbps a um máximo de 600 Mbps, como por exemplo no caso do IEEE 802.11n (KARMAKAR; CHATTOPADHYAY; CHAKRABORTY, 2017).

A Tabela 4.3 apresenta as taxas de transmissão para diferentes valores limite de SNR. Essa tabela demonstra que quanto mais baixa o valor da SNR mais baixa é a taxa de transmissão devido a utilização de esquemas de modulação menos eficientes como o BPSK, bem como o uso maior de informação redundante. A taxa de transmissão é afetada até 25 dB de SNR, no qual a partir dessa medida a taxa se mantém constante dependendo apenas da largura de banda e do intervalo de guarda configurados no AP.

Tabela 4.3: Taxas de transmissão IEEE 802.11n.

SNR mínimo	Modulação	Taxa de codificação	Taxa de transmissão (Mbps)			
			20 MHz		40 MHz	
			800 ns	400 ns	800 ns	400 ns
2	BPSK	1/2	6,5	7,2	13,5	15,0
5	QPSK	1/2	13,0	14,4	27,0	30,0
9	QPSK	3/4	19,5	21,7	40,5	45,0
11	16-QAM	1/2	26,0	28,9	54,0	60,0
15	16-QAM	3/4	39,0	43,3	81,0	90,0
18	32-QAM	2/3	52,0	57,8	108,0	120,0
20	32-QAM	3/4	58,5	65,0	121,5	135,0
25	32-QAM	5/6	65,0	72,2	135,0	150,0

Ao correlacionar os dados obtidos e apresentados na Figura 4.7, foi observado que as maiores taxas de perdas, superiores a 20% dos pacotes, ocorreram entre os valores de SNR inferiores à 25 dB. Neste intervalo o atraso e o jitter se mostraram fortemente correlacionado com a SNR, apresentando um fator de correlação $\rho = -0,9442$ e $\rho = -0,90255$ respectivamente. Ainda para o mesmo intervalo, a SNR indicou um $\rho = -0,64364$ em relação as perdas.

Para valores de SNR superiores à 25dB, as perdas de pacotes apresentaram respectivamente correlação de $\rho = 0,82219$ e $\rho = 0,79525$ em relação ao atraso e o jitter

observados. Essa correlação é esperada pois a taxa de transmissão neste intervalo de SNR foi máxima e constante, independente de outros fatores, o que resultou em taxas baixas e constantes de perdas. Neste intervalo as perdas se mostraram inferiores a 10%.

Desta forma, observa-se uma forte influência da SNR nos tempos de atraso e também no jitter. Diferentes trabalhos publicados mostram que ambas as variáveis de atraso e jitter estão relacionados com fatores físicos do canal de transmissão, neste caso a SNR. Essas justificativas são apresentadas a seguir.

Um estudo realizado em uma rede IEEE 802.11g utilizando um cenário similar ao da Figura 4.2, porém para uma aplicação cliente-servidor VoIP sobre TCP, mostrou que a SNR afeta significativamente o atraso e o jitter (RIVERA-LARA et al., 2008). Resultados mostraram que tanto o jitter quanto a quantidade de pacotes perdidos aumenta significativamente quando a SNR começa a diminuir a partir de 20dB até 11dB (limite mínimo de SNR testado). Para valores de SNR maiores que 20dB essas variáveis apresentaram valores aproximadamente lineares ao longo das coletas de amostras. Em contrapartida, o atraso diferente do jitter e das perdas, aumenta gradativamente à medida que a SNR diminui a partir de 90dB até 11dB (RIVERA-LARA et al., 2008).

Para baixos valores de SNR, Liu *et al.* (2011) apresentam resultados no qual o atraso aumenta à medida que a SNR diminui a partir de de 15dB para 9dB. Essa diminuição da SNR também aumentou a taxa de perdas no link 802.11b testado. Para Karmakar, Chattopadhyay e Chakraborty (2016) a taxa de perdas e o atraso dos pacotes em uma rede 802.11ac passam a aumentar quando a SNR está abaixo de 30dB (KARMAKAR; CHATTOPADHYAY; CHAKRABORTY, 2016).

O jitter de pacotes UDP é de grande importância quando usado, por exemplo, para *streaming* de vídeo em tempo real, pois os dados geralmente são armazenados em *buffers* de memória pelo aplicativo de vídeo. Mukherjee *et al.* (2009) apresentaram em sua pesquisa um aumento do jitter para baixos valores de SNR a partir de 20dB, em contrapartida para valores superiores desse SNR o jitter apresentou menores valores

para uma rede 802.11b/g (MUKHERJEE; PENG; GAO, 2009). Embora as perdas de pacotes sejam mínimas para valores altos do jitter observados neste estudo, elas podem ter um efeito considerável em um fluxo de vídeo, causando artefatos de vídeo, como o congelamento de quadros.

Como indicado nos diferentes trabalhos anteriores e nos testes de correlação, a SNR afeta diretamente os principais parâmetros de QoS como o atraso, jitter e a taxa de perdas. Baseado nestas condições, decidiu-se pela utilização apenas da SNR como o fator dominante na modelagem de perdas na camada física.

4.4.3 Variáveis de camada MAC

Nesta seção são testadas as variáveis de ocupação e do número de usuários compartilhando o mesmo canal. Inicialmente foram calculados os coeficientes de correlação da ocupação do canal em relação ao número de usuários compartilhando o canal. Para quatro diferentes trechos de amostras foi observado os valores dos coeficientes ρ , τ e φ como mostrados na Tabela 4.4. Como observado, a taxa de ocupação do canal não necessariamente está relacionada ao número de usuários conectados. Isso porque os usuários compartilhando um canal podem ter um perfil de transmissão diverso, com maior ou menor uso da rede. Da mesma forma, um canal com alto nível de ocupação pode estar sendo afetado fortemente apenas por um usuário, que pode estar transmitindo ou recebendo *streaming* de vídeo em alta resolução, por exemplo.

Tabela 4.4: Teste de correlação da ocupação do canal com o número de usuários.

Trecho de Amostra	Pearson (ρ)	Kendall (τ)	Spearman (φ)
1	0.348	0.271	0.401
2	-0.067	-0.001	0.008
3	0.809	0.359	0.475
4	0.648	0.287	0.410

Ao analisar a correlação da ocupação do canal e o número de usuários em relação as taxas de perdas de pacotes nas mesmas amostras usadas nos resultados da Tabela 4.4 é possível verificar que as perdas são visualmente predominantes em al-

tas ocupações, como mostrado na Figura 4.8. A Figura 4.8(A) mostra as séries de ocupação e número de usuários, enquanto que a Figura 4.8(B) mostra a série do percentual de perdas. As perdas são mais evidentes em percentuais maiores de ocupação do canal, enquanto que o número de usuários não afeta esse percentual de perdas.

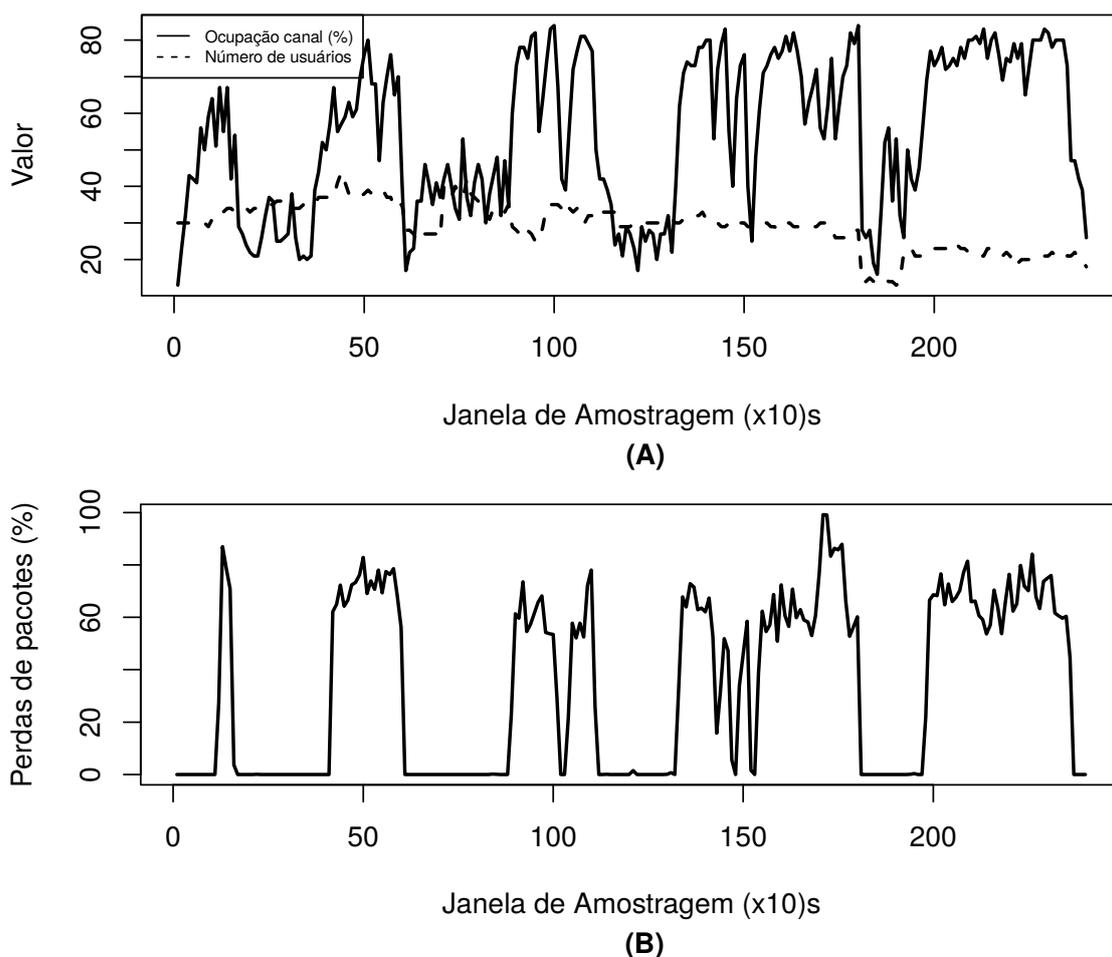


Figura 4.8: Dados observados: (A) ocupação em relação ao número de usuários conectados no canal e (B) taxa de perdas por janela de amostragem.

A Figura 4.8(A) apresenta a sequência de dados da ocupação e número de usuários das quatro amostras utilizadas no cálculo dos coeficientes de correlação da Tabela 4.4. A Figura 4.8(B) mostra a taxa de perdas de pacotes no mesmo intervalo observado da Figura 4.8(A). Ao verificar a correlação entre esses dados foi constatado que: o número de usuários em relação a taxa de perdas apresenta $\rho = -0.0778914$, $\tau = -0.07251272$ e $\varphi = -0.07891022$, logo esses dados não estão relacionados; a

ocupação do canal em relação a taxa de perdas apresentou $\rho = 0.8162346$, $\tau = 0.4975907$ e $\varphi = 0.7232222$, desta forma neste último caso, essas duas variáveis apresentaram considerável nível de correlação entre si.

A ocupação do canal é utilizado como principal variável de entrada em diferentes métodos aplicados em redes sem fio (GANDARILLAS et al., 2014). Desta forma, a utilização do comportamento de ocupação do espectro de WiFi é vital para o planejamento do uso do espectro e o desenvolvimento de dispositivos sem fio na faixa ISM de 2,4 GHz, incluindo modelos de representação dessas redes (RAJAB; BALID; REFAI, 2015).

Baseado na baixa correlação das perdas com o número de usuários conectados no canal de transmissão, essa variável não foi considerada como uma entrada do modelo para treinamento no HMM, diferente da ocupação do canal que será considerado devido ao alto nível de correlação observado.

4.5 Modelo de Perdas

4.5.1 Hidden Markov Model (HMM)

Um HMM é composto de dois processos estocásticos acoplados. O primeiro é uma cadeia de Markov e o outro é um processo de observação cuja distribuição a qualquer momento é totalmente determinada pelo estado atual da cadeia de Markov do primeiro processo (SILVEIRA; SILVA, 2012). Em outras palavras, o HMM é usado na modelagem de processos Markovianos que geram comportamentos observáveis de forma indireta, em função das transições entre os estados da cadeia de Markov que governa o processo inicial, mas que não pode ser diretamente observado sem a observação do segundo processo. Desta forma, o comportamento do processo da cadeia de Markov está escondida do observador.

Para ilustrar a estrutura do HMM, a Figura 4.9 mostra o diagrama de transição

de estado de um HMM de dois estados. Cada um dos dois estados ocultos (y_i) é mapeado para um dos três resultados observáveis (x_k) com alguma probabilidade (b_{ik}). As probabilidades de transição de estado (a_{ji}) são as probabilidades de passar de um estado oculto para outro.

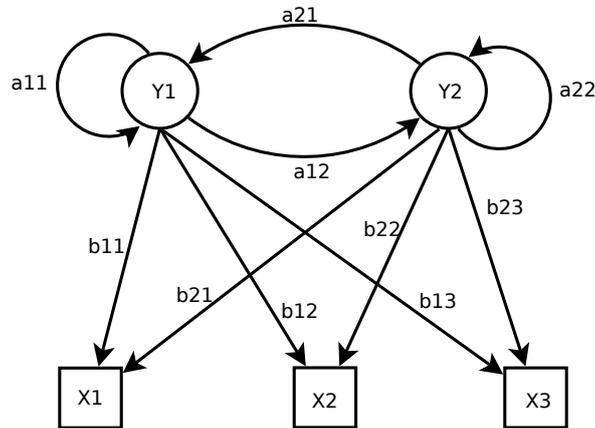


Figura 4.9: Diagrama de transição de estado de um HMM (TOLEDO; KATZ, 2009).

A variável estocástica Y_n , denota a cadeia de Markov de N estados subjacentes que pode assumir valores do espaço de estados Y em diferentes instantes de tempo t . O espaço de estado é indicado por $Y = \{y_1, y_2, y_3, \dots, y_{N-1}, y_N\}$ (SILVEIRA; SILVA, 2012). O número de estados do modelo de perdas de pacotes para rede WiFi será definido na subseção 4.5.2 a seguir. Cada estado representa o estado atual do sistema em observação, onde cada estado representa o comportamento das perdas de pacotes do modelo no instante t . Os estados são importantes para indicar, em condições de percentual de perdas, qual o melhor e pior estado, onde ainda é possível obter estados intermediários de perdas quando $N > 2$.

Uma vez que uma cadeia de Markov não mantém nenhum histórico, nem o HMM, então a única coisa que um modelo oculto de Markov pode lembrar é em qual estado ele está no momento t que é observado e não é possível modelar dependências entre eventos distantes (SCHUSTER-BÖCKLER; BATEMAN, 2007). Esta é a limitação mais importante do HMM que resulta da própria propriedade de cadeias de Markov. Os estados de um HMM são ocultos e nunca são observados diretamente (SILVEIRA; SILVA, 2012).

A probabilidade de transição entre os N estados do HMM é descrita por uma matriz quadrada $N \times N$ igual a $\mathbf{A} = \{a_{ij}\}$, onde

$$a_{ij} = P(Y_n = y_j | Y_{n-1} = y_i) \quad (4.6)$$

Dentro do modelo de perdas, o sistema transita entre os N estados do HMM, indicando o comportamento no qual o sistema pode sofrer com as diversas condições de qualidade de uma rede WiFi. Desta forma, de um bom estado com baixos níveis de perdas de pacotes o sistema pode migrar para um estado ruim ou intermediário, onde as perdas ocorrem com mais ou menos intensidade. Ainda, o sistema pode permanecer no mesmo estado de perdas quando as condições de qualidade da rede se mantêm estáveis.

Sabendo que é impossível definir o estado atual do sistema no instante t de forma direta, faz-se necessário o uso de um conjunto de observações que permita dizer qual o estado atual do sistema. O processo de observação é denotado pela variável estocástica $\{X_n\}$ que possui M símbolos. Como $\{X_n\}$ é o processo de observação, seus símbolos são geralmente determinados pelo que está sendo modelado. O treinamento estabelece as probabilidades b_{ij} , o que permite fazer seleção posterior do estado mais provável y_n a partir dos valores observados de x_k . Essa variável pode assumir valores do conjunto de observáveis também em diferentes instantes de tempo t (SILVEIRA; SILVA, 2012). A probabilidade de observação é representada por uma matriz $N \times M$ igual a $\mathbf{B} = \{b_{ij}\}$, no qual

$$b_{ij} = P(X_n = x_j | Y_n = y_i) \quad (4.7)$$

Após cada transição ser feita entre os N estados, um símbolo M de saída de observação é produzido de acordo com a distribuição de probabilidade que depende do estado atual definida na matriz \mathbf{B} . Essa distribuição de probabilidade é mantida fixa para o estado, independentemente de quando e como o estado é inserido.

No modelo de perdas, o conjunto de observações x_n é definido pelos símbolos definidos pelos valores inteiros da SNR e da ocupação do canal da rede WiFi. Isso quer dizer que para cada combinação das variáveis de SNR e ocupação do canal no instante t o HMM poderá definir qual o estado mais provável Y_n que representa o sistema naquele momento t . Observe que neste momento o HMM apenas define qual o estado mais provável atual do sistema sem considerar o real comportamento das perdas de pacotes do estado. Para o HMM esta análise de cada estado é irrelevante, o que faz com que o modelador do sistema necessite analisar e caracterizar o comportamento das perdas dentro de cada estado. O estado mais provável é obtido utilizando-se o algoritmo de Viterbi (VITERBI, 1967).

A distribuição que descreve o estado inicial do HMM no instante $t = 1$ é indicado pelo vetor $\Pi = \{\pi_i\}$ no qual possui N dimensões, com (SILVEIRA; SILVA, 2012)

$$\pi_i = P(Y_1 = s_i) \quad (4.8)$$

A Tabela 4.5 apresenta um exemplo de uma típica sequência de observação do modelo proposto. Neste exemplo, T representa o tamanho da sequência de observação baseado nas observações das variáveis de SNR e ocupação. No início da sequência de observação, com instante $t = 1$, no qual a SNR é igual a 22dB e a ocupação do canal é igual a 10% o HMM indica que o estado atual do sistema é igual a s_1 (estado 1) baseado no vetor de inicialização Π . A cada tempo, t , um novo estado é inserido com base na probabilidade de transição da matriz \mathbf{A} que depende do estado anterior (a propriedade Markoviana). Observe que a transição pode ser tal que o processo permanece no estado anterior, como visto no instante $t = 3$ no qual o sistema permaneceu no estado 2 em relação ao estado anterior.

No HMM o conjunto desses parâmetros do modelo é indicado como uma tupla dada por $\Delta = (\Pi, A, B)$, e dado os seus significados probabilísticos, onde $\sum_{i=1}^N \pi_i = 1$, $\sum_{j=1}^N a_{ij} = 1$ e $\sum_{j=1}^M b_{ij} = 1$ são restrições que devem ser respeitadas.

Tabela 4.5: Exemplo típico da sequência de observação baseado no modelo proposto.

Instante de tempo (t)	1	2	3	4	5	6	7	8	...	T
Estado escondido	y_1	y_2	y_2	y_1	y_3	y_3	y_1	y_2	...	y_{N-2}
SNR (dB) (x_1)	22	63	62	44	41	23	46	56	...	77
Ocupação do canal (%) (x_2)	10	24	14	16	41	56	85	45	...	23

Uma importante tarefa dentro da abordagem de modelagem do HMM é o problema de estimativa dos parâmetros, ou seja, como é realizada a inferência dos valores de Δ dado um conjunto de dados amostrais da sequência de treinamento. Para solucionar este problema, são utilizadas técnicas de estimativa iterativa dos parâmetros de máxima verossimilhança do HMM. Normalmente, neste caso é aplicado o algoritmo Baum-Welch (ADAMS; BELING; COGILL, 2016) e durante o processo de treinamento do modelo proposto foi usado este algoritmo. Esse algoritmo é um caso especial de uma classe de algoritmos iterativos para estimativa de máxima verossimilhança conhecido coletivamente como o método *Expectation-Maximization* (EM) (SILVEIRA; SILVA, 2012). Qualquer algoritmo baseado em EM começa com uma estimativa inicial de parâmetros e produz estimativas sucessivamente melhores, garantindo convergência para um máximo local na função de verossimilhança sempre que houver. Neste caso, o critério adotado para definição do melhor treinamento do HMM obtido neste trabalho é o maior valor de verossimilhança. Para o treinamento foram usadas bibliotecas disponíveis no software R (R CORE TEAM, 2019).

O HMM utiliza treinamento baseado em aprendizado não supervisionado, no qual os dados utilizados não possuem nenhum tipo de identificação quanto a sua característica e neste aprendizado o HMM tenta identificar similaridades entre os objetos durante as iterações em que este processo ocorre. Nesta tese, o treinamento da rede é utilizado como um método para parametrização do modelo proposto. Outros modelos de perdas tem dificuldades de parametrização devido a complexidade computacional dos algoritmos e técnicas utilizadas neste processo. A redução em esforço computacional é uma das vantagens do modelo proposto baseado em HMM. Desta forma, é possível disponibilizar parametrizações pré definidas para diversas situações

da rede (SNR e ocupação do canal) para o caso de simulações.

O outro problema é encontrar o estado provável da cadeia de Markov a partir de uma sequência de observações. Para encontrar o estado mais provável a partir de um conjunto de observações, Rabiner (1989) sugere a utilização do Algoritmo de Viterbi (VITERBI, 1967) e desde então usado em uma grande gama de aplicações, como por exemplo, em tarefas de reconhecimento de fala (LOU, 1995). Esta abordagem é uma solução ótima recursiva para o problema de estimar a sequência de estados de um processo Markoviano de estado finito e tempo discreto. Ao final, o Algoritmo de Viterbi encontra a sequência de transição de estados mais provável em uma sequência de símbolos, tendo como entrada as variáveis da HMM treinada.

4.5.2 Definição do número de estados

A definição do número de estados ideal para treinamento do HMM foi realizada utilizando técnicas de clusterização de dados. A clusterização é uma técnica de aprendizagem não supervisionada utilizada para o agrupamento de dados. O objetivo da clusterização é agrupar um conjunto de objetos de dados em vários grupos, também chamados de agrupamentos, para que os objetos dentro do mesmo agrupamento sejam semelhantes entre si e que tenham grande dissimilaridade com os objetos de outros clusters (LI et al., 2019). São conhecidas várias técnicas para o agrupamento, como por exemplo a distância Euclidiana, a distância Mahalanobis e a distância Manhattan (BIANCHESI et al., 2019). Neste trabalho foi aplicado o método da distância Euclidiana.

A apresentação desses agrupamentos é realizada através de gráficos do tipo dendrogramas. O dendrograma é uma estrutura de representação em árvore, como mostrado no exemplo da Figura 4.10, na qual os pontos de conjuntos de dados com similaridade estão próximos uns dos outros formando o agrupamento (URMELA; NANDHINI, 2017).

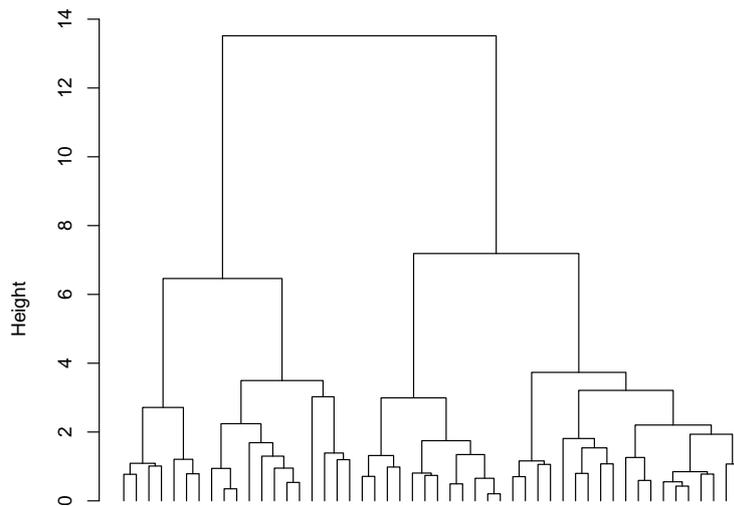


Figura 4.10: Exemplo de dendrograma com clusterização usando método distância Euclidiana.

Para Han, Pei e Kamber (2011) um dendrograma é uma estrutura em árvore comumente usada para representar de forma gráfica o processo de clusterização hierárquica de um conjunto de objetos e mostra quais são agrupados a partir de um método aglomerativo ou particionado em um método de divisão. Como apresentado na Figura 4.10, o dendrograma permite visualizar como os clusters são formados em cada passo de distância (*height*) e para avaliar os níveis de similaridade dos agrupamentos que são formados (HAN; PEI; KAMBER, 2011). A forma de como a similaridade ou os valores de distância mudam de um passo para outro auxilia no processo de escolha do agrupamento mais adequado para os dados analisados. A definição desse melhor agrupamento é dada a partir de um corte do dendrograma, no qual é inserido uma ou mais linhas horizontais por todo o dendrograma. No exemplo da Figura 4.10 é possível aferir dois agrupamentos principais, contudo ainda é possível identificar a existência de quatro grupos, como retificado na Figura 4.11.

Baseado nas séries de SNR e ocupação do canal de transmissão obtidos no processo de coleta de amostras, como mostrado na Figura 4.12, será realizado o processo de clusterização nessas séries. Essas duas séries são as sequências de observáveis do modelo HMM que serão utilizadas no treinamento de forma a obter os parâmetros de otimização da tupla $\Delta = (\Pi, A, B)$. A partir do dendrograma resul-

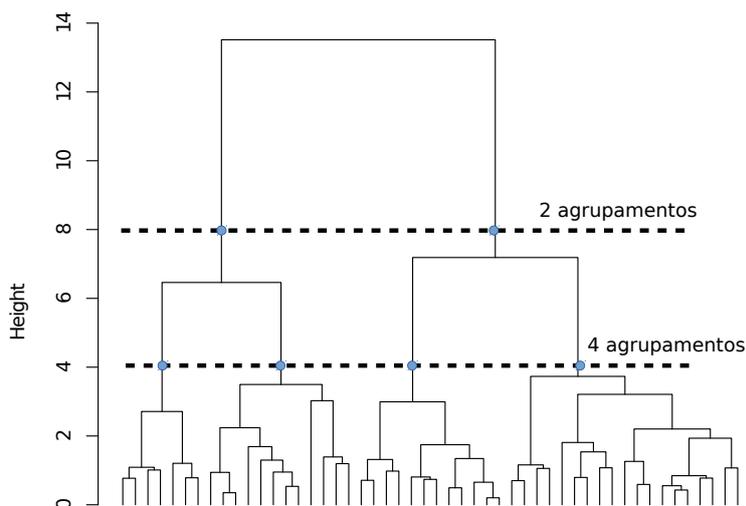


Figura 4.11: Representação linhas de corte em dendrograma gerado a partir da distância Euclidiana.

tante da clusterização dos dados será verificado o número mais provável de grupos, que consequentemente será utilizado como estimativa principal do número de estados do processo de treinamento do HMM. A Figura 4.12 apresenta a série de dados da ocupação do canal e também a da SNR utilizada na clusterização.

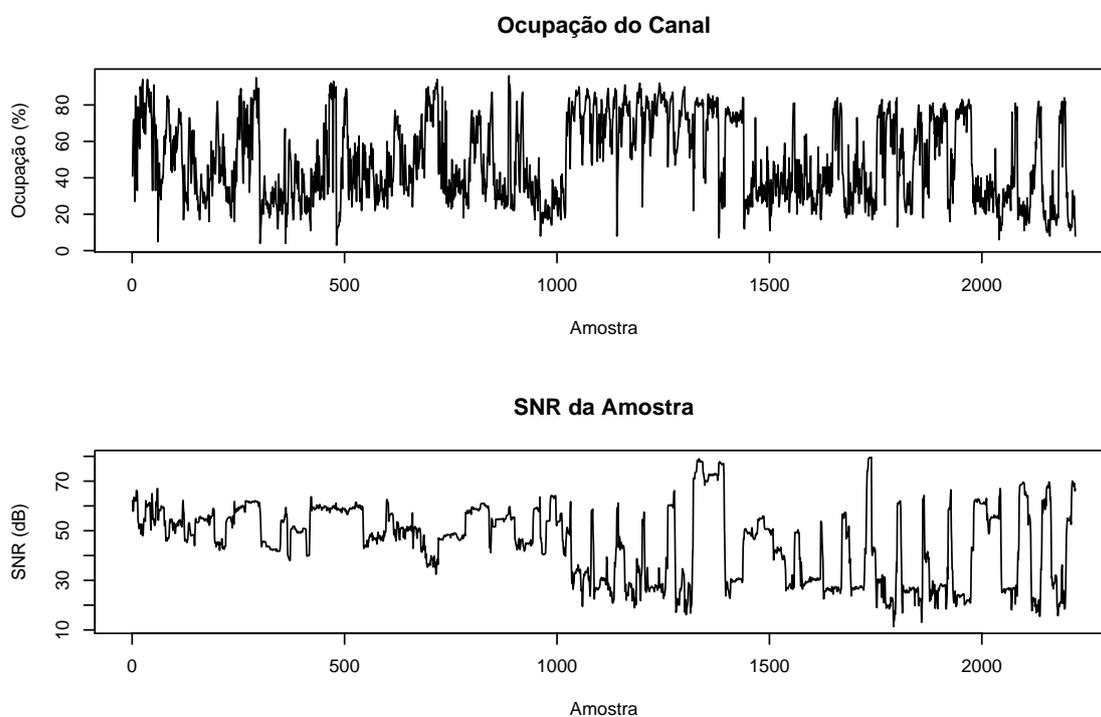


Figura 4.12: Séries de dados principais de SNR e ocupação do canal.

4.5.3 Modelagem das perdas por estado

Após o treinamento do HMM considerando como critério de parada o maior valor de verossimilhança (*likelihood*) obtido do uso das bibliotecas “depmixS4” (VISSER; SPEEKENBRINK et al., 2010), “quantmod” e “qcc” disponíveis no R (R CORE TEAM, 2019), foi analisada a sequência de estados prováveis que resultou deste processo. Durante a coleta de amostras, foram identificados um total de 2400 pontos de SNR e mais 2400 pontos de ocupação do canal, sendo que cada ponto representa o intervalo de 10 segundos de coleta de chegadas ou perdas de pacotes (aproximadamente 10 mil pacotes). A sequência ideal de estados do HMM treinado é representada pelo conjunto $\Theta = \{s_1, s_2, s_3, s_4, \dots, s_{2399}, s_{2400}\}$, no qual s pode assumir qualquer estado provável, considerando o número de estados selecionados no processo de treinamento como estado 1, estado 2, estado 3 até o valor máximo de N .

Para cada conjunto de estados iguais as sequências de chegadas e perdas (0's e 1's) de pacotes são concatenadas de forma a permitir a extração do comprimento das rajadas de perdas. A sequência de perdas de pacotes comumente ocorre em rajadas (WANG; XU, 2007) e dentro de um modelo de perdas pode ser considerado como um estado ruim, de erro ou de falha. Neste trabalho, o comprimento das rajadas foi calculado pela quantidade de perdas ocorridas de forma consecutiva até que o próximo pacote tenha sido recebido. Foram computados os tamanhos de rajadas observados em cada um dos estados.

A abordagem adotada neste trabalho é baseada na modelagem estatística do comprimento das rajadas observadas. Observou-se que o tamanho das rajadas de perda apresentou comportamento de cauda pesada, assim como já apresentado em trabalhos relacionados a esta pesquisa (ROHLING et al., 2016) e também por outros autores (RUSS; HAGHANI, 2009) e (YU; MILLER, 2007). As distribuições que foram usadas neste trabalho têm a propriedade de serem de cauda pesada. Uma distribuição tem cauda pesada se (CROVELLA; BESTAVROS, 1997)

$$P[X > x] \sim x^{-\alpha}, \quad \text{com } x \longrightarrow \infty, \quad 0 < \alpha < 2 . \quad (4.9)$$

Ou seja, independentemente do comportamento da distribuição para pequenos valores da variável aleatória, se a forma assintótica da distribuição for hiperbólica, ela é de cauda pesada.

A distribuição de cauda pesada mais simples é a distribuição de Pareto. A distribuição de Pareto é hiperbólica em toda a sua extensão e sua função de massa de probabilidade é (CROVELLA; BESTAVROS, 1997)

$$p(x) = \alpha k^\alpha x^{-\alpha-1}, \quad \alpha, k > 0, \quad x \geq k . \quad (4.10)$$

A função de distribuição cumulativa é dada por (CROVELLA; BESTAVROS, 1997)

$$F(x) = P[X \leq x] = 1 - (k/x)^\alpha . \quad (4.11)$$

O parâmetro k representa o menor valor possível da variável aleatória (CROVELLA; BESTAVROS, 1997).

Essa distribuição de cauda pesada tem um número diferente de propriedades comparada a distribuições como a exponencial, Gaussiana ou Poisson. Se $\alpha \leq 2$ então a distribuição possui variância infinita e se $\alpha \leq 1$, então a distribuição possui média infinita. Assim, à medida que α diminui, uma porção arbitrariamente grande da massa de probabilidade pode estar presente na cauda da distribuição. Em termos práticos, uma variável aleatória que segue uma distribuição de cauda pesada pode dar origem a valores extremamente grandes com probabilidade não desprezível (CROVELLA; BESTAVROS, 1997).

Para analisar esse comportamento nas amostras coletadas, foi empregado o uso de gráficos de distribuição complementar log-log (LLCD, *Log-log Complementary Dis-*

tribution). Esses gráficos são obtidos através de uma distribuição complementar acumulativa $\bar{F}(x) = 1 - F(x) = P[X > x]$ no eixo log-log. Traçado desta forma, as distribuições de cauda pesada têm a propriedade de

$$\frac{d \log \bar{F}(x)}{d \log x} = -\alpha, \quad x > \theta. \quad (4.12)$$

para qualquer valor de θ . Logo, para verificar a presença de caudas pesadas na prática, é formado o gráfico LLCD e procurado um comportamento aproximadamente linear ao longo de uma faixa significativa com três décadas ou mais na cauda (CROVELLA; BESTAVROS, 1997).

Foram testadas diferentes distribuições de probabilidade que são amplamente utilizadas na modelagem ou representação de características de causa pesada em séries de dados. Foi verificado a aderência de funções como Weibull, Pareto Tipo II, Log-Normal, Cauchy and Log-Cauchy. A parametrização de cada distribuição de probabilidade foi realizada através de estimadores disponíveis em diversas bibliotecas também no software R (R CORE TEAM, 2019). Dentre os estimadores de parâmetros utilizados neste processo pode-se citar o *Maximum likelihood* (MLE), *Moment Matching* (MLE), *Quantile Matching* (QME) e *Maximizing Goodness-of-fit Estimation* (MGE).

Ao verificar a função de distribuição acumulada (CDF, *cummulative distribution function*) do comprimento das rajadas de perdas a partir das distribuições utilizadas no processo de verificação da aderência observou que a distribuição de Pareto Tipo II apresentou melhor aderência em relação as demais. O uso da distribuição de Pareto Tipo II já é sugerida em outros trabalhos na modelagem das rajadas de perdas de pacotes, com em (CROVELLA; BESTAVROS, 1997; BORELLA, 2000; TANG et al., 2006; MORSY; SOWAILEM; SHALABY, 2010; ROHLING et al., 2016).

A distribuição Pareto Tipo II também é conhecida como Lomax. Uma variável aleatória Lomax X com parâmetro de escala (λ , *scale*) e parâmetro de forma (α , *shape*) tem função de densidade probabilidade:

$$f(x) = \frac{\lambda\alpha}{(1 - \lambda x)^{\alpha+1}} \quad x > 0, \quad (4.13)$$

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

A CDF é dada por:

$$F(x) = P(X \leq x) = 1 - (1 + \lambda x)^{-\alpha} \quad x > 0 . \quad (4.14)$$

No entanto, a distribuição de Pareto Tipo II foi incapaz de representar totalmente a cauda pesada presente na CDF do comprimento das rajadas em alguns estados. Neste caso, a série foi separada em dois conjuntos utilizando um limite e modelados por duas distribuições de probabilidade, a distribuição exponencial e a distribuição de Pareto Tipo II com sucesso.

A quantidade de dados modelada pela distribuição exponencial corresponde a menos de 1% do total de dados da amostra analisada. Uma variável aleatória exponencial possui uma função densidade probabilidade dada por:

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \quad x > 0 , \quad (4.15)$$

onde μ é a média do comprimento das rajadas.

A CDF da função exponencial é dada por:

$$F(x) = P(X \leq x) = 1 - e^{-x/\mu} . \quad (4.16)$$

A partir das distribuições indicadas anteriormente, são realizadas simulações no qual são comparadas com as amostras reais de cada estado a partir de gráficos do tipo Quantile-Quantile Plot (QQ-Plot). O QQ-Plot é um método gráfico simples usado para comparar coleções de dados ou distribuições teóricas e ajudar na identificação

da função de distribuição. Em estatística, um gráfico QQ-Plot é uma representação visual de probabilidades, que é um método gráfico para comparar duas distribuições de probabilidade, traçando seus quantis uns contra os outros.

4.6 Validação do Modelo

O modelo proposto foi validado por meio de seis amostras de dados coletadas que representam aproximadamente 1 hora cada. Essas séries são chamadas de amostras de validação. Cada amostra possui os níveis de SNR, percentual de ocupação do canal, incluindo também a sequência de chegadas e perdas do pacote relacionada a cada SNR e ocupação. Com o HMM treinado, cada amostra é submetida a verificação da sequência de estados ideal utilizando o algoritmo de Viterbi, também disponível na biblioteca *depmixS4* do software R (R CORE TEAM, 2019). Os valores de média e os desvios-padrão do comprimento das rajadas em cada estado das amostras treinada no HMM são comparados com o comprimento das rajadas das amostras de validação e também de uma série simulada do modelo.

Os resultados apresentados neste modelo também são comparados com os resultados apresentados por trabalhos relacionados da literatura. É verificado o erro médio quadrático (MSE, *mean square error*) do comprimento das rajadas estimadas a partir da simulação do modelo proposto com os demais trabalhos relacionados.

CAPÍTULO 5

ANÁLISE DOS RESULTADOS E DISCUSSÃO

Inicialmente são apresentados os resultados do treinamento do HMM em relação as amostras coletadas e da estimação dos parâmetros do modelo proposto. Em seguida, a avaliação do modelo proposto em relação a aderência dos resultados simulados e os dados reais é realizada e discutida. Por último, o modelo proposto é comparado aos modelos disponíveis na literatura.

5.1 Treinamento do HMM e Parametrização do Modelo

A Figura 5.1 apresenta o dendrograma para as duas séries resultantes da clusterização das séries de níveis de SNR e percentuais de ocupação do canal. Estas séries são as mesmas apresentadas na Figura 4.12. Visualmente é possível verificar a existência de 4 grupos utilizando SNR e ocupação do canal. Desta forma, neste trabalho assume-se que o número ideal de estados é igual a $N = 4$. Assim, o modelo baseado no HMM para quatro estados passa a ser chamado de HMM4, no qual também utilizará a sequência de treinamento formada pelas variáveis de SNR e ocupação do canal. Ainda, também foram realizados testes para 3 e 5 estados, que apresentaram resultado inferior em relação ao HMM4.

Durante o treinamento, o HMM4 ajustou os parâmetros do modelo $\Delta = (\Pi, \mathbf{A}, \mathbf{B})$ baseado na sequência de treinamento utilizando o método iterativo de Baum-Welch. A dispersão dos níveis de SNR e ocupação do canal em relação aos estados definidos pelo HMM4 são apresentadas na Figura 5.2. Uma alternativa seria dividir esses grupos em quatro quadrantes de forma direta sem a necessidade de treinamento do HMM, contudo ao fazer isso não seria possível obter o processo de transição entre

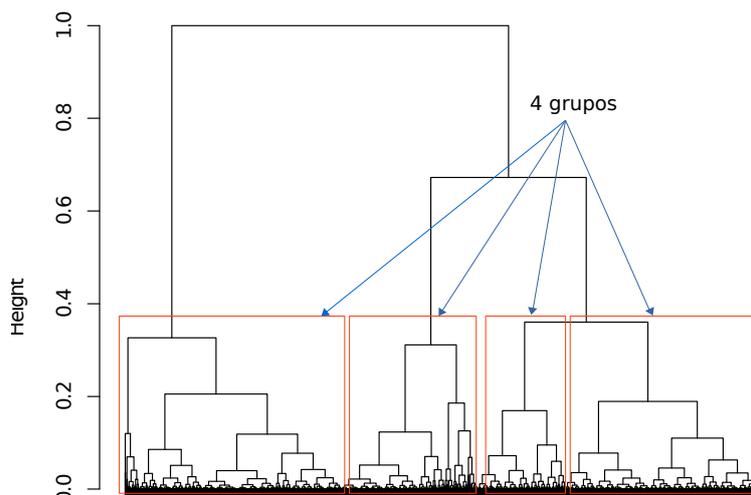


Figura 5.1: Dendrograma da clusterização com distância Euclidiana.

os estados do modelo. Desta forma, aferir os estados apenas separando os níveis de ocupação e SNR torna o modelo inadequado para modelagem de perdas de pacote.

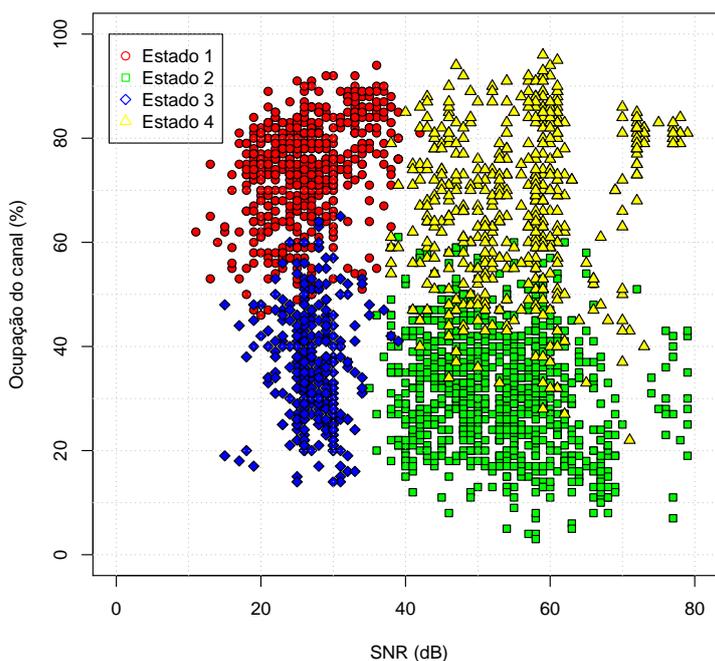


Figura 5.2: Estados treinados a partir do HMM4.

Observa-se que o treinamento não supervisionado do HMM4 foi capaz de identificar os quatro estados a partir das variáveis observáveis (sequência de treinamento). O estado 1 é representado por altos percentuais de ocupação do canal e também por baixos níveis de SNR. Considerando os valores dessas duas variáveis é esperado que este seja o pior estado possível em relação as perdas de pacotes. O estado 2

é representado pelos melhores níveis de SNR e com baixa ocupação do canal, no qual é esperado que este estado seja o melhor estado possível também em relação as perdas de pacotes. Os estados 3 e 4 são considerado estados intermediários, no qual representam respectivamente, baixos níveis de SNR e ocupação do canal, e altos níveis de SNR e de ocupação. As médias de SNR e ocupação do canal para cada estado são mostradas na Tabela 5.1.

Tabela 5.1: Informação da média de SNR e ocupação do canal por estado.

Estado	Média de SNR	Média de ocupação
1	26,5 dB	73,7%
2	52,4 dB	31,6%
3	27,0 dB	35,6%
4	55,1 dB	67,3%

Lembrando que até este momento as características e o comportamento das séries de perdas de pacotes não foram utilizadas no processo de treinamento. A Tabela 5.2 apresenta as características de perdas de pacotes dos estados analisados definidos pelo HMM4. Nesta etapa para cada estado é extraída a sequência de perdas respectivas de forma que seja possível analisar o comportamento dessas perdas, como a taxa de perda de pacotes (PLR) respectiva de cada estado. 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 novamente a Figura 5.2 e Tabela 5.1 no qual os níveis referentes ao estado 1 estão concentrados nos menores níveis de SNR e maiores taxas de ocupação do canal. 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 com 0,55% de média. Os demais estados 3 e 4 são considerados intermediários. Observa-se que o estado 4 apresenta considerável taxa média de perdas de 12,78% mesmo com níveis de SNR considerados ideais para uma rede Wi-Fi, porém essa taxa é influenciada pela ocupação do canal. Isso mostra que em modelos baseados somente em características físicas, como a SNR, não é considerada a influência dos

aspectos da camada de enlace. Desta forma, esses modelos falham em reproduzir as perdas de pacotes existentes quando considerado que as condições de camada física como SNR e potência do sinal são melhores ou ideais. O estado 3 apresenta média de perdas de pacotes de 2,02% com baixos níveis de SNR, porém com níveis baixos e considerado bons de ocupação do canal.

Tabela 5.2: Informação da média e desvio padrão do percentual de perdas por estado.

Estado	Média de perdas	Desvio padrão das perdas	Definição
1	60,97%	24,93%	Ruim
2	0,55%	4,09%	Bom
3	2,02%	8,36%	Intermediário 1
4	12,78%	21,76%	Intermediário 2

O processo de treinamento HMM4 resultou na matriz de transição **A** entre os quatro estados da sequência de treinamento, conforme mostrado a seguir

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

A matriz mostra que a probabilidade de permanecer no mesmo estado é maior que 95% para os estados 1 e 2, e ainda nesses mesmos estados existe a possibilidade de transição para todos os outros. Ao verificar os estados intermediários 3 e 4, a probabilidade de permanecer no mesmo estado é maior que 92%, porém, não há possibilidade de transição entre esses dois estados, apenas para o pior e melhor estado, respectivamente 1 e 2. O vetor do estado inicial é igual a $\mathbf{II} = [0 \ 0 \ 0 \ 1]$, em que para a sequência de treinamento utilizada o estado inicial proposto pelo HMM4 é o estado 4.

Com a definição do estados prováveis, foi extraído para cada estado as séries que correspondem ao tamanho das rajadas de perdas de pacotes. Essas séries são

apresentadas na Figura 5.3, no qual é possível constatar comprimentos de rajadas de perdas maiores no estado 1 e menores no estado 2. Ainda, é possível identificar a presença dos "spikes", mais evidentes nos estados 1 e 2.

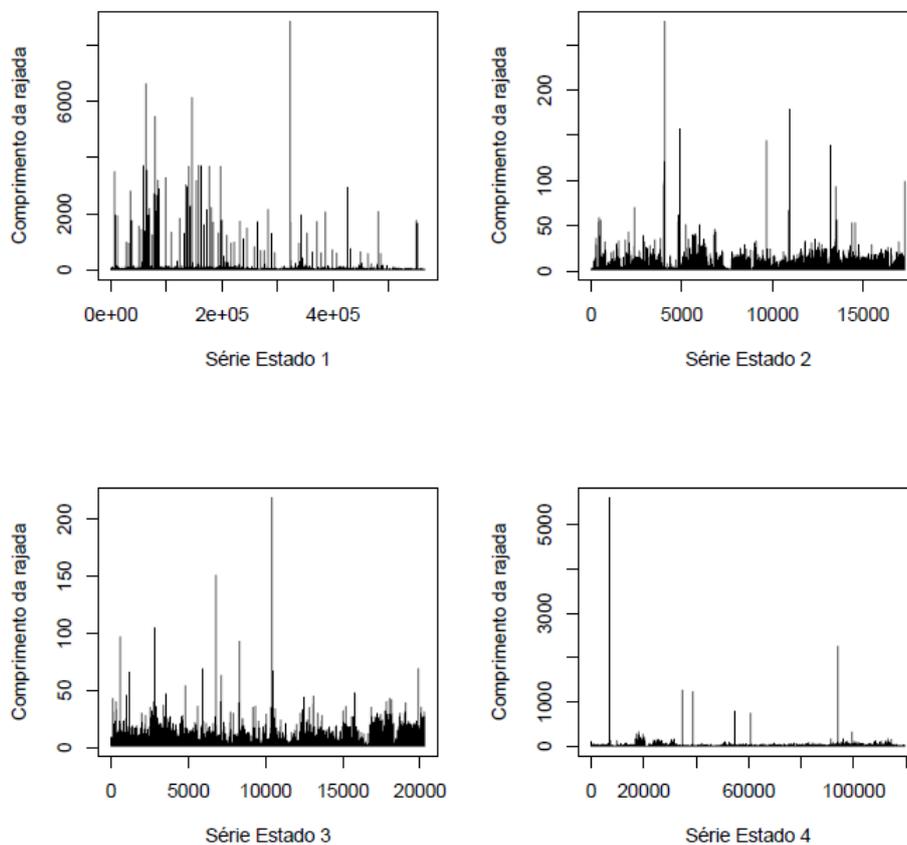


Figura 5.3: Série de comprimento das rajadas de perdas por estado.

A representação do comprimento das rajadas através de distribuições de cauda pesada foi realizada através de estimadores dos parâmetros de escala (λ) e forma (α) da distribuição de Pareto Tipo II. A Figura 5.4 apresenta os gráficos de QQ-Plot por estado, no qual mostra os quantis gerados pela distribuição de Pareto Tipo II em comparação aos dados da série do comprimento das rajadas de perdas. Caso os pontos estejam colocados exatamente sobre a linha na diagonal com inclinação de 45° as variáveis aleatórias dos eixos horizontal e vertical são idênticas. Esses gráficos mostram que a distribuição utilizada no modelo conseguiu representar adequadamente, dentro de um intervalo de confiança de 95%, os dados das séries para os estados 3 e 4. Embora esse tipo de gráfico seja muito sensível nas variações de valores, os quantis para os estados 1 e 2 não conseguiram ser representados totalmente pela distribuição

de Pareto Tipo II. Neste caso, essa distribuição conseguiu representar corretamente até aproximadamente o comprimento da rajada de 400 e 40, respectivamente para os estados 1 e 2.

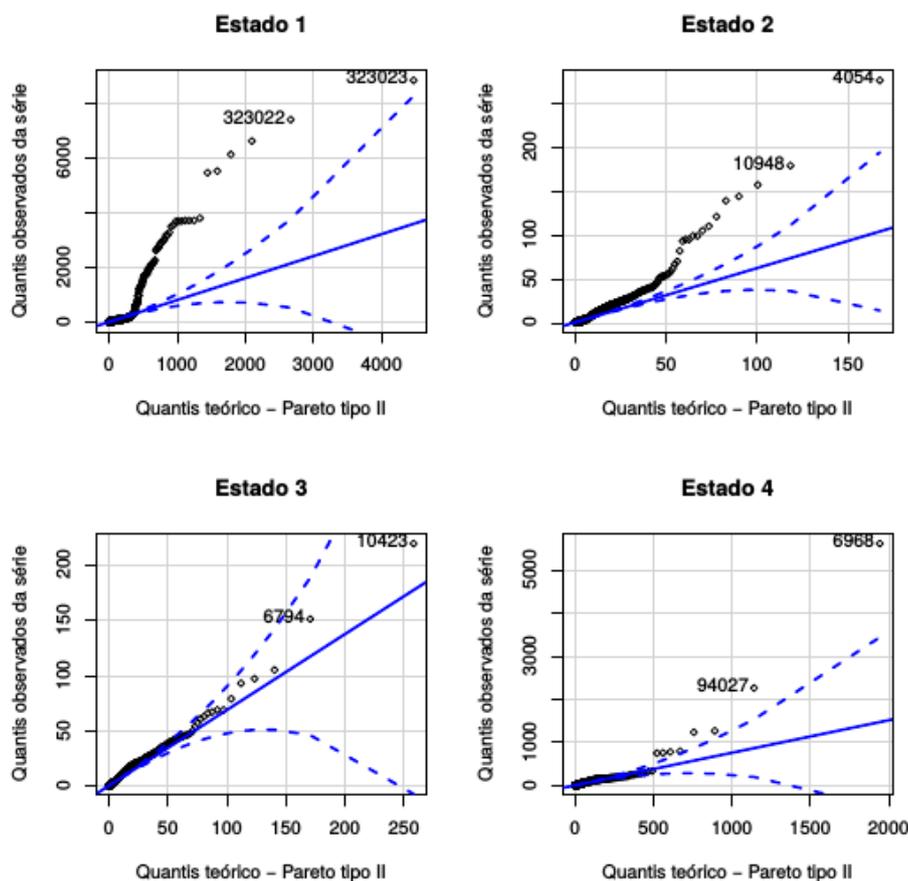


Figura 5.4: Gráficos de QQ-Plot para a distribuição de Pareto Tipo II por estado.

O ajuste apresentado na subseção 4.5.3, no qual os spikes são separados e modelados por uma distribuição exponencial, resolve o problema do ajuste do modelo identificado no QQ-Plot da figura anterior. Esse procedimento foi aplicado nos estados 1 e 2, no qual os gráficos de QQ-Plot para cada estado é mostrado na Figura 5.5. O comprimento das rajadas superiores ao limites (400 e 40) definidos para esses dois estados foram separados e os valores médios extraídos para compor o parâmetro μ da distribuição exponencial. Desta forma, a cauda da CDF é representada pela distribuição de Pareto Tipo II e os spikes por uma distribuição do tipo Exponencial.

Os valores dos parâmetros de escala, forma e média das distribuições são apresentadas na Tabela 5.3.

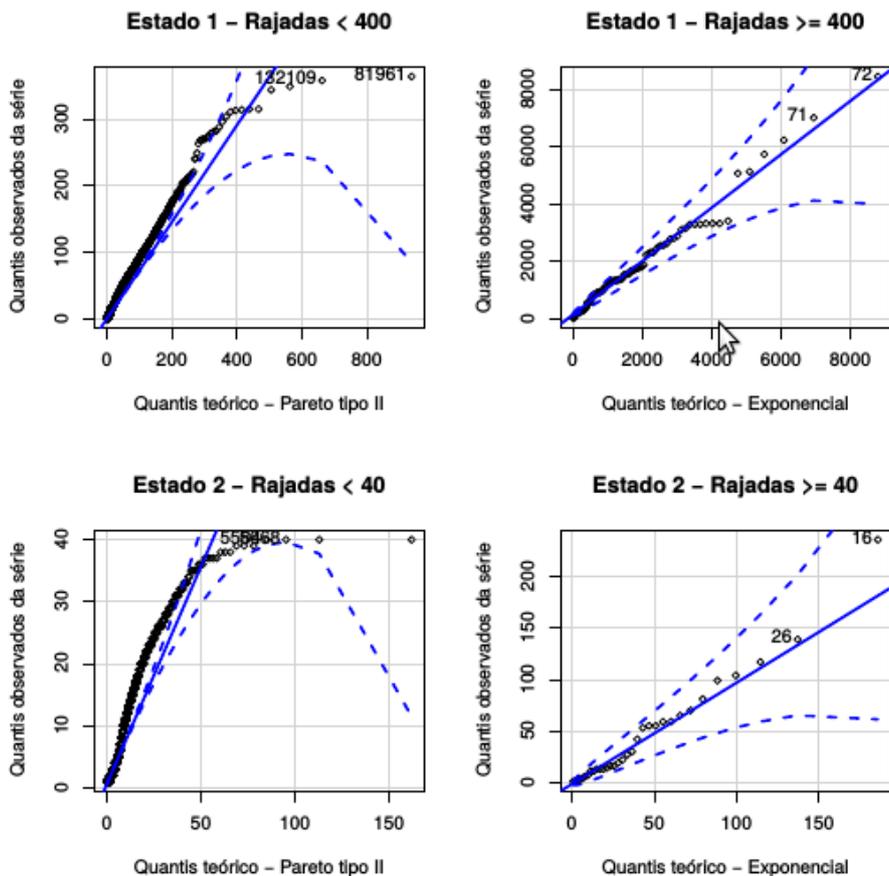


Figura 5.5: Gráficos de QQ-Plot para as distribuições de Pareto tipo II e Exponencial dos estados 1 e 2.

Tabela 5.3: Valores dos parâmetros de escala, forma e média para distribuições dos estados 1 e 2 (SILVA; SANTOS; PEDROSO, 2020).

Estado	Função de Probabilidade	Parâmetros estimados
1	Pareto Tipo II Exponencial	$\alpha = 3,208319$ e $\lambda = 12,323558$ $\mu = 1682,613$
2	Pareto Tipo II Exponencial	$\alpha = 3,18529$ e $\lambda = 6,32771$ $\mu = 44,36364$
3	Pareto Tipo II	$\alpha = 3,419363$ e $\lambda = 7,237939$
4	Pareto Tipo II	$\alpha = 2,072002$ e $\lambda = 4,942283$

5.2 Avaliação do Modelo Proposto

Para verificar a aderência da distribuição baseada no modelo proposto das equações parametrizadas da Tabela 5.3, foi simulado uma série de comprimentos das rajadas de perda para cada estado. Em seguida, foi comparado a média e o desvio padrão desses comprimentos. Os resultados desta comparação são apresentados na Tabela 5.4,

onde o comprimento médio das rajadas de cada estado obtido no treinamento HMM4 é aproximadamente igual ao comprimento simulado se considerado o desvio padrão. Como esperado, o maior comprimento médio das perdas é apresentado no estado 1 com 5,67, seguido dos estados 4, 3 e 1 com médias 4,66, 3,03 e 3,00 respectivamente. O maior comprimento médio das rajadas também é observado nos dados simulados a partir das distribuições de cada estado.

Tabela 5.4: Comparativo da média e desvio padrão do comprimento das rajadas de perdas obtidas no modelo do HMM4 e do modelo simulado.

Estado	HMM4	Modelo simulado
1	5,67 ± 34,44	5,80 ± 28,64
2	3,00 ± 5,75	2,99 ± 6,01
3	3,03 ± 4,70	2,97 ± 4,48
4	4,66 ± 21,04	4,56 ± 11,89

O QQ-Plot dos quantis de comprimento das rajadas dos estados do HMM4 e resultados simulados são mostrados na Figura 5.6. A figura mostra que as distribuições propostas para modelar o comprimento das rajadas podem representar com relativa precisão o comprimento das rajadas reais da amostra de dados reais. Ainda, para o estado 4 é necessário um estudo aprofundado sobre porque alguns pontos não puderam ser totalmente representados para a distribuição proposta no modelo. Contudo, sugere-se que estes pontos tenham como origem a existência de alguns spikes menores. Como é possível perceber há pequenos desvios da linha diagonal para os estados 1, 2 e 3, contudo, esta representação visual do QQ-Plot é mais sensível do que muitos testes existentes.

A Figura 5.7 apresenta o QQ-Plot da comparação entre a amostra original completa do comprimento das rajadas, onde não é considerado as séries de dados por estado, e a série simulada de dados baseado no modelo resultante da matriz de probabilidade de transição **A** gerado pelo HMM4 e as funções de probabilidade descritas no modelo. Esse QQ-Plot mostra que a matriz de transição **A** obtida do treinamento do HMM4 e as distribuições de cada estado conseguiram representar adequadamente o comprimento das rajadas de perdas obtidos na amostra real.

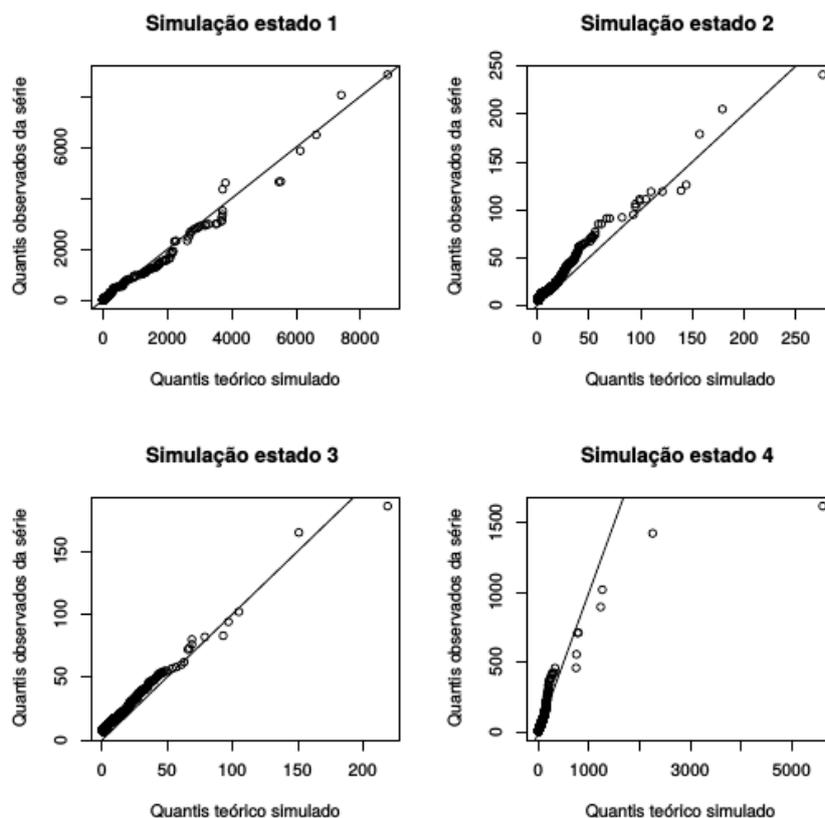


Figura 5.6: QQ-Plot dos quantis das séries obtidas do HMM4 e os simulados por estado.

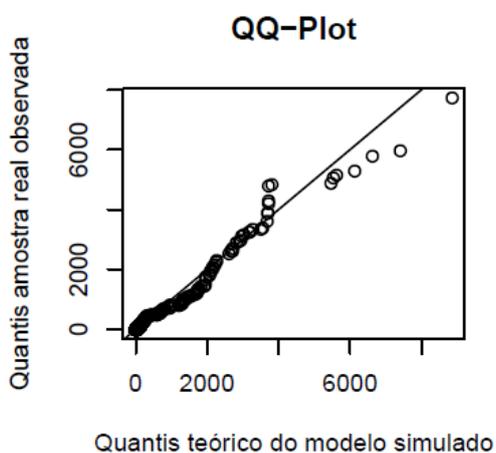


Figura 5.7: QQ-Plot da simulação da matriz de transição **A** com a amostra observada.

Os resultados apresentados anteriormente mostram que os dados simulados a partir da matriz de transição gerada do treinamento do HMM4 e das distribuições de probabilidade parametrizadas adequadamente são aderentes aos dados obtidos na rede com amostra real. Embora esses resultados mostrem que o comprimento das rajadas tenham sido modeladas adequadamente, é necessário verificar se novas séries

de perdas de pacotes, SNR e ocupação do canal conseguem ser representadas pelo HMM4 já treinado. A Tabela 5.5 apresenta o PLR e o desvio padrão de cada estado para as seis amostras utilizadas na validação do modelo. Nas amostras 2 e 5 não foi identificado o estado 3 porque os valores de SNR e ocupação do canal não pertenciam ao terceiro estado quando utilizado o algoritmo de Viterbi para comparar com o HMM4 previamente treinado. Ainda, nas amostras 1 e 6 foi possível identificar o estado 4, porém a perda de pacotes, neste caso, foi de 0%. Isso ocorre devido a sensibilidade do sistema em considerar um estado não ideal para um determinado nível de SNR e ocupação do canal, podendo ser mais comum em níveis próximos aos limites de um estado e outro. Pode-se assumir que essa condição ocorre devido à matriz de probabilidade de transição que sugere que o próximo estado possível é o 4. Embora o algoritmo de Viterbi apresente pequenas variações quando a acurácia na identificação do estado correto, na maioria dos casos, pode-se verificar que o algoritmo conseguiu identificar corretamente os estados, onde o percentual de perda de pacotes de cada estado é aproximado ao percentual de perdas quando comparado à Tabela 5.2 referente ao PLR da amostra inicial antes do treinamento do HMM4.

Tabela 5.5: Comparativo da média e do desvio padrão dos percentuais de perdas de pacotes obtidos das amostras de validação.

Estado	Amostra 1		Amostra 2		Amostra 3	
	PLR	Desvio padrão	PLR	Desvio padrão	PLR	Desvio padrão
1	61,26%	18,8%	14,81%	28,17%	43,65%	19,68%
2	0,57%	4,95%	0,59%	4,58%	0,4%	4,05%
3	3,73%	14,16%	n/a	n/a	0,86%	3,32%
4	0%	0%	9,10%	20,18%	10,37%	18,99%
Estado	Amostra 4		Amostra 5		Amostra 6	
	PLR	Desvio padrão	PLR	Desvio padrão	PLR	Desvio padrão
1	53,77%	19,03%	76,21%	17,01%	57,85%	21,54%
2	0,61%	5,04%	0,56%	3,92%	0,58%	4,81%
3	1,46%	4,40%	n/a	n/a	1,41%	8,77%
4	24,58%	28,30%	30,64%	30,50%	0%	0%

5.3 Comparação com Trabalhos Relacionados

Primeiro, foi comparado o comprimento das rajadas de perda observado na amostra real com o comprimento das rajadas de perda gerados pelo modelo de GE. Para parametrizar o modelo GE, utilizou-se a sequência de chegadas e perdas também da amostra principal observada. No modelo GE, a probabilidade de transição entre os estados Ruim e Bom são $p = 0,0393$ e $q = 0,1862$, respectivamente para $p = P(R|B)$ e $q = P(B|R)$. A Figure 5.8 apresenta a distribuição complementar em escala logarítmica (LLCD, *Log-log Complementary Distribution*) da série observada do comprimento das rajadas, seguido da série simulada do modelo proposto e por último a série de rajadas gerada pelo modelo de GE. Como observado, a LLCD resultante do modelo simulado apresenta comportamento de cauda pesada que foi modelada pela distribuição de Pareto Tipo II. Ainda, esses dados simulados apresentam características similares com a amostra observada. A LLCD gerada dos dados do modelo de GE apresenta comportamento similar a da distribuição exponencial, o que retifica que o modelo de GE não consegue representar a presença de cauda pesada como nos dados observados.

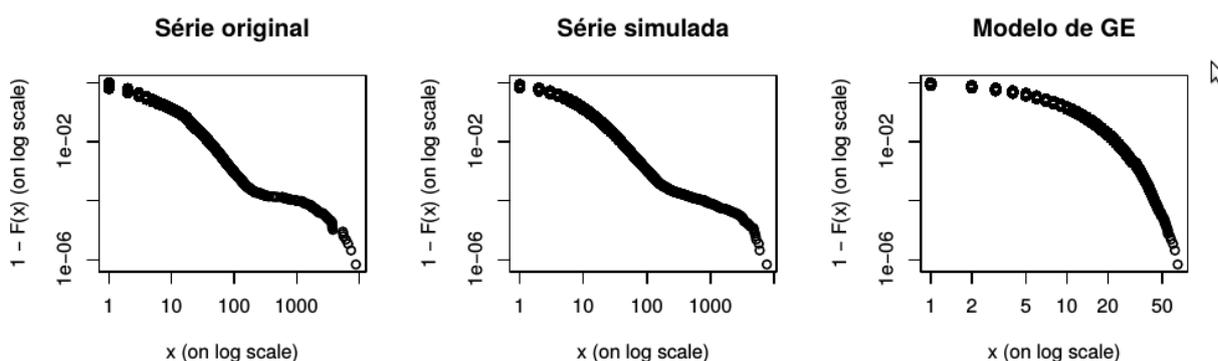


Figura 5.8: LLCD das amostras observada, simulada e do modelo de GE.

A ACF do modelo de GE mostrou também que não há presença de dependência temporal entre o comprimento das rajadas de perdas, como apresentado na Figura 5.9, o que confirma a incapacidade desse modelo representar o comportamento real em redes Wi-Fi. Ainda, devido a transição entre os estados Bom e Ruim do modelo de

GE, o comprimento máximo das rajadas obtido na simulação foi de 65, bem diferente dos vários comprimentos acima desse limite na amostra observada na rede e também simulada no modelo proposto.

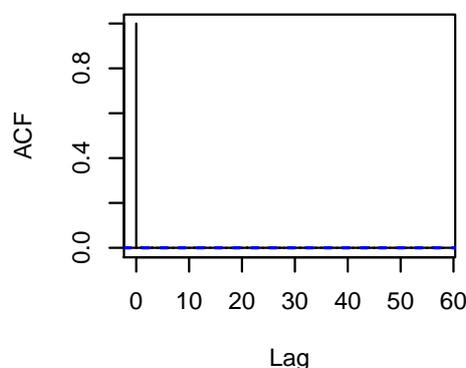


Figura 5.9: ACF do modelo de GE.

Em seguida são apresentados os resultados da simulação de trabalhos disponíveis na literatura. As características e parametrização de cada modelo utilizado para comparação com o modelo proposto é apresentado a seguir:

- Conforme proposto por Arauz e Krishnamurthy (2003), as amostras foram separadas por quatro níveis de SNR de 10dB a 75dB. Cada nível representa um estado do modelo e cada um é modelado como uma cadeia de Markov de dois estados, onde o primeiro intervalo de SNR (10-26dB) é o pior estado e o quarto intervalo (60-75dB) é o melhor. Os estados intermediários são de SNR 27dB à 42dB e 43dB à 59dB. A divisão em quatro níveis de SNR foi definido empiricamente devido ao número de estados utilizados no treinamento do HMM4 proposto no modelo deste trabalho.
- O segundo modelo sugerido por Carvalho, Angeja e Navarro (2005) propõe que o comprimento das rajadas de perda seja modelado com uma distribuição de série geométrica com $P[X = x] = \frac{k\theta^n}{n}$ onde $n = 0,1,2,\dots$, no qual $k = -\frac{1}{\ln(1-\theta)}$. A estimativa do parâmetro de $\theta = 0,936848$ foi estabelecida a partir da aplicação da Estimativa de Máxima Verossimilhança (MLE, *maximum likelihood estimation*) (BÖHNING, 1983).

- O modelo GE com sub-estados proposto por Feng, Liu e Ji (2014) sugere uma estrutura baseada em um estado Bom formado por conjunto de quatro estados adjacentes onde as probabilidades de transição desses estados para o estado Ruim são $p_1 = 0,009749$, $p_2 = 0,004928$, $p_3 = 0,002926$, e $p_4 = 0,001855$, como visto na Figura 3.7. A transição do estado Ruim para o primeiro sub-estado do estado Bom é $q = 0,153737$.
- O modelo baseado em um mecanismo híbrido do modelo de GE e uma distribuição de cauda pesada proposta por Russ e Haghani (2009a) foi simulado. O comprimento das rajadas de perda ≤ 3 foi simulado com um modelo GE onde $p = 0,6240$ e $q = 0,6216$, e para o comprimento das rajadas > 3 foi simulado com uma distribuição de Pareto Tipo II onde os parâmetros de forma e escala são iguais a 2,5192 e 19,7564, respectivamente.
- O modelo de Cadeia de Markov de Estados Finitos (FSMC) proposto por Sanneck e Carle (1999) foi simulado com 4, 5, 7, 10 e 20 estados, onde a probabilidade de transição foi estimada com base na amostra real coletada da rede WiFi. Os valores parametrizados deste modelo estão disponíveis no Apêndice A.

A Tabela 5.6 apresenta a comparação entre os resultados absolutos das características do modelo proposto com os trabalhos relacionados. Na tabela são apresentados o comprimento médio das rajadas de perdas, o desvio padrão e o máximo comprimento obtido em cada trabalho testado. Por fim é apresentado o MSE obtido da comparação do resultado simulado com a amostra original obtida na rede avaliada. Para apresentar a melhor precisão do modelo proposto, é mostrado na primeira coluna os valores medidos na amostra observada na rede WiFi. Como resultado, o modelo proposto apresenta comprimento médio de 5,52 e a amostra real com média de 5,37. O modelo de GE apresentou comprimento médio das rajadas mais próximo dos valores observados na amostra observada. Ainda, observa-se que outros trabalhos como (RUSS; HAGHANI, 2009) apresentou uma média igual a 8,11 e (SANNECK; CARLE, 1999) com médias maiores diretamente relacionada ao aumento no número de esta-

dos. Trabalhos publicados sugerem o aumento do número de estados dos modelos baseados em cadeias de Markov, contudo esse resultado mostrou que ao acrescentar mais estados, além de aumentar a complexidade computacional do modelo (bem como a sua parametrização), pode não apresentar melhores resultados.

Tabela 5.6: Comparativo entre o modelo proposto e trabalhos relacionados.

Característica	Amostra coletada	GE Model	Modelo de 2-estados e SNR	Série geométrica	GE com sub-estados	GE e cauda pesada	FSMC com 4 estados	FSMC com 5 estados	FSMC com 7 estados	FSMC com 10 estados	FSMC com 20 estados	Amostra real de validação
Média do comprimento das rajadas de perda	5,37	5,36	4,88	5,37	6,52	8,11	16,76	22,52	34,85	54,19	150,5	5,52
Comprimento máximo das rajadas de perda	8.853	65	54	158	96	11.074	204	273	489	721	2.286	7.728
Desvio padrão do comprimento das rajadas de perda	31,68	4,83	4,52	7,53	6,01	24,08	15,78	21,45	33,80	53,39	156,667	29,75
MSE	0	0,352	0,319	0,352	0,352	2,907	0,352	0,352	0,351	0,349	0,318	0,0141

O processo de treinamento do modelo proposto em HMM sugere a utilização do conhecido algoritmo de Baum-Welch, que realiza a estimação dos parâmetros do modelo. Ao comparar com modelos baseados em Cadeias de Markov como o de GE que apresenta alta complexidade na estimação de seus parâmetros (HASSLINGER; HOHLFELD, 2008), o modelo proposto representa menor esforço computacional. O aumento no número de estados como nos modelos baseado no FSMC aumentam ainda mais a complexidade do processo de estimação dos parâmetros, contudo podendo não apresentar ganhos significativos na representação das perdas.

Ainda na Tabela 5.6, os modelos que consideram o uso de distribuições de cauda pesada como o proposto nesta tese e também em (RUSS; HAGHANI, 2009) conseguem representar comprimentos máximos próximo ao da amostra observada, o que mostra que esse tipo de abordagem baseada em distribuições de cauda pesada apresenta melhores resultados na caracterização das séries de dados. O MSE dos comprimentos de rajadas de perdas do modelo proposto é igual a 0,0141 quando comparado aos valores da série de amostra observada e as amostras de validação utilizadas. Este valor baixo, mostra que ao aplicar um conjunto de dados não utilizado no treinamento do HMM, o modelo consegue apresentar valores de comprimento de rajadas próximos ao da amostra observada. Dentre os demais modelos, o trabalho de Russ e Haghani (2009a) apresentou um MSE de 2,907 no qual a modelagem é feita a partir do modelo híbrido com o GE e distribuição de cauda pesada. Contudo, neste trabalho a parametrização é feita através de amostras reais para obtenção dos parâmetros de p e q do modelo de GE, e os valores de escala e forma da distribuição de Pareto Tipo II. O modelo proposto nesta tese pode ser utilizado independente de amostras reais para parametrização, no qual é necessário apenas indicar as condições das camadas física e enlace, respectivamente para a SNR e a ocupação do canal.

A Figura 5.10 apresenta o QQ-Plot do comprimento das rajadas de perdas simulados dos trabalhos relacionados em comparação com os dados da amostra real observada na rede. Observando essas comparações é possível identificar que a maioria

dos modelos testados conseguem representar o comprimento das rajadas de perdas, contudo somente até um determinado comprimento máximo. A partir deste limite, o modelo se torna ineficiente para representar o comportamento real da rede. O modelo híbrido que utiliza o modelo de GE e também uma distribuição de cauda pesada para modelagem da perdas consegue apresentar rajadas superiores ao comprimento de 1.500 no qual também foi identificado na amostra real, contudo ainda não conseguiu representar adequadamente o comportamento total das rajadas quando observado o QQ-Plot. Diferente do modelo proposto que apresentou um QQ-Plot mais aderente aos dados da amostra real observada, como já apresentado anteriormente na Figura 5.7.

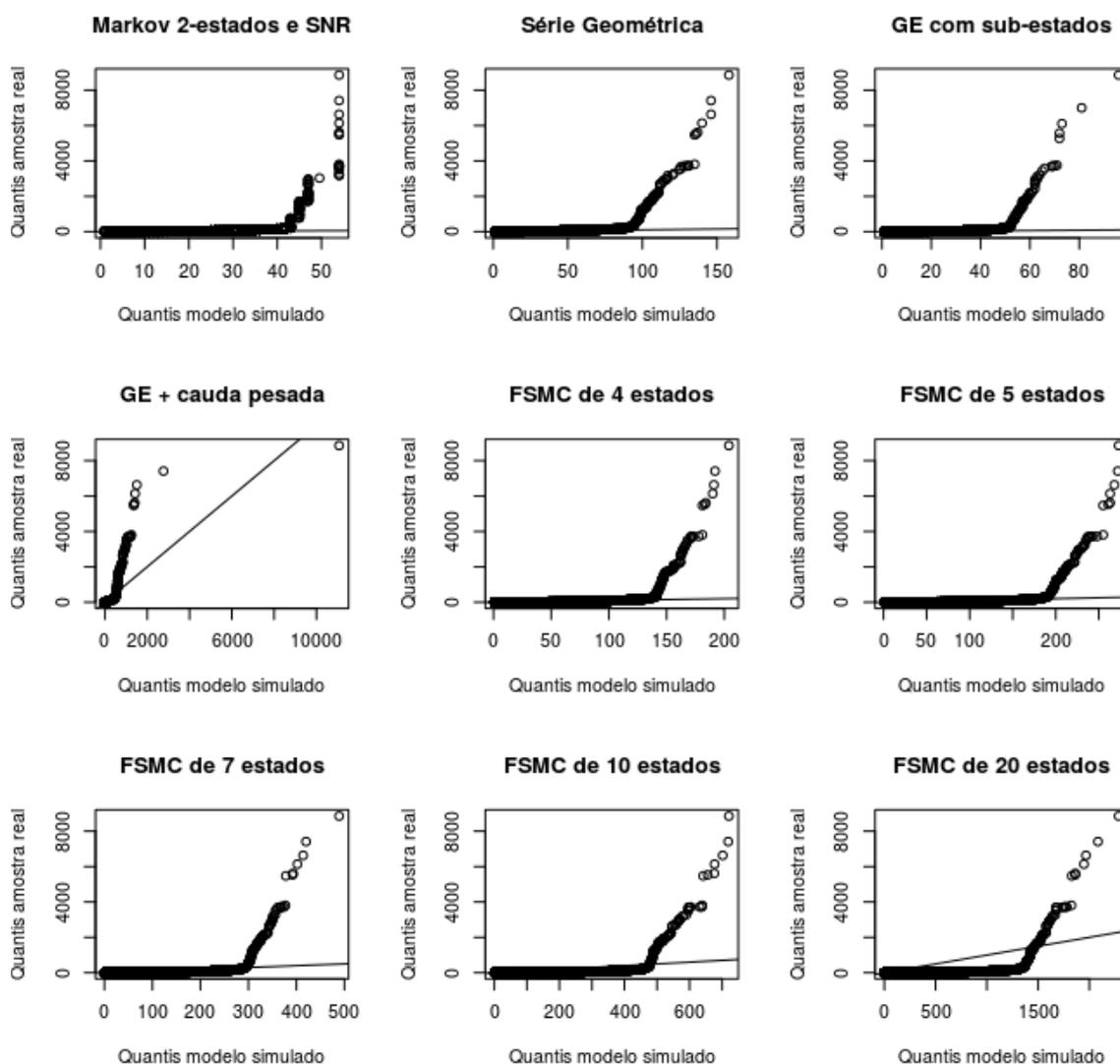


Figura 5.10: QQ-Plot trabalhos relacionados com a amostra real observada.

CAPÍTULO 6

CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho apresentou um novo modelo de perdas de pacotes para rede WiFi a partir da aplicação de modelos ocultos de Markov (HMM) considerando simultaneamente características de camada física e de enlace de uma rede real. O modelo proposto modela o comprimento das rajadas de perdas de pacotes tipicamente existentes em redes WiFi baseado no protocolo IEEE 802.11b/g/n.

O trabalho apresenta uma revisão do protocolo 802.11 e as principais origens de ocorrência de perdas de pacotes em rede sem fio. As perdas na maioria das vezes estão associadas à qualidade do canal e sinal de transmissão, coexistência com outras redes, interferências e fenômenos de multicaminho. Devido a existência de terminais ocultos e também a disputa pelo canal de transmissão por diversos dispositivos podem ocorrer colisões e conseqüentemente as perdas. A falta ou excesso de memória pode também ocasionar perdas nos equipamentos de comutação e perda de pacotes.

Os modelos de perdas de pacotes são, na maioria da vezes, desenvolvidos a partir de cadeias de Markov. O principal modelo de perdas de pacotes conhecido é o modelo de Gilbert-Elliot, no qual diversas pesquisas mostram ser ineficiente para representar o comportamento real em redes WiFi. Demais trabalhos, sugerem o uso de modelos com múltiplos estados ou a aplicação de distribuições de cauda pesada para a modelagem das perdas. Observa-se que os modelos existentes consideram séries de dados reais para a parametrização dos modelos ou utilizam de forma independente características de camada física e/ou enlace.

Testes de correlação mostraram que a SNR e a ocupação do canal de transmissão tem impacto direto na perda de pacotes. A utilização de técnicas de clusterização permitiu identificar 4 grupos de similaridade das séries de SNR e ocupação. Ao aplicar

esses dados como parâmetros de entrada no treinamento do modelo oculto de Markov, foi possível identificar que o HMM conseguiu classificar as observações de SNR e ocupação do canal em quatro estados. Para cada estado, o comprimento das rajadas de perdas foi modelado com a utilização da distribuição de Pareto Tipo II em conjunto com a distribuição exponencial. O modelo proposto conseguiu representar corretamente o comprimento das rajadas de perdas, além de conseguir capturar dependência temporal entre as rajadas. O modelo se mostrou melhor que os trabalhos relacionados, no qual estes não conseguiram capturar dependência temporal entre as rajadas e também a função de distribuição do comprimento das rajadas de perdas.

Uma das grandes vantagens do modelo proposto apresentado é a parametrização, que é resultado do treinamento da HMM a partir de um algoritmo bastante conhecido.

Como trabalhos futuros, sugere-se a aplicação deste modelo em cenários específicos para aplicações como *streaming* de vídeo, voz sobre IP ou web. Ainda, sugere-se testar a eficiência do modelo proposto em protocolos WiFi de geração 5 e 6, respectivamente IEEE 802.11ac e 802.11ax.

APÊNDICE A

PARAMETRIZAÇÃO MODELO FSMC

Considere o modelo proposto por Sanneck e Carle (1999) para a aplicação de cadeias de Markov de número finito de estados, como mostrado na Figura 3.5. O estado 1 (E_1) é livre de erro e os estados seguintes apresentam perda, onde o estado 2 (E_2) representa uma perda após uma chegada, o estado 3 representa duas perdas consecutivas após uma chegada, o estado 4 representa três perdas consecutivas após uma chegada e assim por diante até o estado máximo (E_{MAX}) que representa $MAX - 1$ perdas consecutivas.

Ainda, $P(E_2|E_1) = q$, $P(E_1|E_1) = 1 - q$, $P(E_1|E_2) = p_1$, $P(E_1|E_3) = p_2$, $P(E_1|E_4) = p_3$ até $P(E_1|E_{MAX}) = p_{MAX-1}$. A parametrização das probabilidades de transição entre os estados para os seguintes número de estado (4, 5, 7, 10 e 20) são apresentadas a seguir:

A.1 FSMC de 4 estados

$$P(E_2|E_1) = q = 0,019458$$

$$P(E_1|E_1) = 1 - q = 0,980542$$

$$P(E_1|E_2) = p_1 = 0,062078$$

$$P(E_1|E_3) = p_2 = 0,030082$$

$$P(E_1|E_4) = p_3 = 0,061577$$

$$P(E_4|E_4) = 1 - p_3 = 0,938423$$

A.2 FSMC de 5 estados

$$P(E_2|E_1) = q = 0,019458$$

$$P(E_1|E_1) = 1 - q = 0,980542$$

$$P(E_1|E_2) = p_1 = 0,062078$$

$$P(E_1|E_3) = p_2 = 0,030082$$

$$P(E_1|E_4) = p_3 = 0,015923$$

$$P(E_1|E_5) = p_4 = 0,045654$$

$$P(E_5|E_5) = 1 - p_4 = 0,954346$$

A.3 FSMC de 7 estados

$$P(E_2|E_1) = q = 0,019458$$

$$P(E_1|E_1) = 1 - q = 0,980542$$

$$P(E_1|E_2) = p_1 = 0,062078$$

$$P(E_1|E_3) = p_2 = 0,030082$$

$$P(E_1|E_4) = p_3 = 0,015923$$

$$P(E_1|E_5) = p_4 = 0,009928$$

$$P(E_1|E_6) = p_5 = 0,006573$$

$$P(E_1|E_7) = p_6 = 0,029153$$

$$P(E_7|E_7) = 1 - p_6 = 0,970847$$

A.4 FSMC de 10 estados

$$P(E_2|E_1) = q = 0,019458$$

$$P(E_1|E_1) = 1 - q = 0,980542$$

$$P(E_1|E_2) = p_1 = 0,062078$$

$$P(E_1|E_3) = p_2 = 0,030082$$

$$P(E_1|E_4) = p_3 = 0,015923$$

$$P(E_1|E_5) = p_4 = 0,009928$$

$$P(E_1|E_6) = p_5 = 0,006573$$

$$P(E_1|E_7) = p_6 = 0,004015$$

$$P(E_1|E_8) = p_7 = 0,003419$$

$$P(E_1|E_9) = p_8 = 0,003167$$

$$P(E_1|E_{10}) = p_9 = 0,018552$$

$$P(E_{10}|E_{10}) = 1 - p_9 = 0,981448$$

A.5 FSMC de 20 estados

$$P(E_2|E_1) = q = 0,019458$$

$$P(E_1|E_1) = 1 - q = 0,980542$$

$$P(E_1|E_2) = p_1 = 0,062078$$

$$P(E_1|E_3) = p_2 = 0,030082$$

$$P(E_1|E_4) = p_3 = 0,015923$$

$$P(E_1|E_5) = p_4 = 0,009928$$

$$P(E_1|E_6) = p_5 = 0,006573$$

$$P(E_1|E_7) = p_6 = 0,004015$$

$$P(E_1|E_8) = p_7 = 0,003419$$

$$P(E_1|E_9) = p_8 = 0,003167$$

$$P(E_1|E_{10}) = p_9 = 0,002122$$

$$P(E_1|E_{11}) = p_{10} = 0,001498$$

$$P(E_1|E_{12}) = p_{11} = 0,001452$$

$$P(E_1|E_{13}) = p_{12} = 0,001400$$

$$P(E_1|E_{14}) = p_{13} = 0,001302$$

$$P(E_1|E_{15}) = p_{14} = 0,001198$$

$$P(E_1|E_{16}) = p_{15} = 0,000999$$

$$P(E_1|E_{17}) = p_{16} = 0,000763$$

$$P(E_1|E_{18}) = p_{17} = 0,000701$$

$$P(E_1|E_{19}) = p_{18} = 0,000752$$

$$P(E_1|E_{20}) = p_{19} = 0,006360$$

$$P(E_{20}|E_{20}) = 1 - p_{19} = 0,993640$$

BIBLIOGRAFIA

ABDELRAHMAN, R. B. M.; MUSTAFA, A. B. A.; OSMAN, A. A. A Comparison Between IEEE 802.11a,b,g,n and ac Standards. **IOSR Journal of Computer Engineering**, v. 17, n. 5, p. 26–29, 2015.

ABUSUBAIH, M. Approach for Discriminating Losses in 802.11 Wireless LANs. **IET Communications**, v. 6, n. 10, p. 1262–1269, Julho 2012.

ADAMS, S.; BELING, P. A.; COGILL, R. Feature Selection for Hidden Markov Models and Hidden Semi-Markov Models. **IEEE Access**, v. 4, p. 1642–1657, 2016.

AL-FUQAHA, A. et al. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. **IEEE Communications Surveys Tutorials**, v. 17, n. 4, p. 2347–2376, Junho 2015.

ALATABANI, L. E.; ABDALLA, A. G. E. FHSS, DSSS, And Hybrid DS/FH Performance Evaluation for VSAT. **International Journal Of Scientific and Technology Research**, v. 4, n. 9, p. 58–62, 2015.

ALLAHDADI, A.; MORLA, R. Anomaly detection and modeling in 802.11 wireless networks. **Journal of Network and Systems Management**, Springer, v. 27, n. 1, p. 3–38, 2019.

ANGRISANI, L. et al. Assessing Coexistence Problems of IEEE 802.11b and IEEE 802.15.4 Wireless Networks Through Cross-layer Measurements. In: **IEEE Instrumentation Measurement Technology Conference IMTC 2007**. Warsaw, Polônia: IEEE, 2007. p. 1–6.

_____. Experimental Study of Coexistence Issues Between IEEE 802.11b and IEEE 802.15.4 Wireless Networks. **IEEE Transactions on Instrumentation and Measurement**, v. 57, n. 8, p. 1514–1523, Agosto 2008.

ARAUZ, J.; KRISHNAMURTHY, P. Markov Modeling of 802.11 Channels. In: **IEEE 58th Vehicular Technology Conference (VTC 2003)**. Orlando, FL, EUA: IEEE, 2003. v. 2.

BAGHAEI, N.; HUNT, R. IEEE 802.11 wireless LAN security performance using multiple clients. In: **12th IEEE International Conference on Networks (ICON 2004)**. Singapura: IEEE, 2004. v. 1, p. 299–303.

BEJARANO, O.; KNIGHTLY, E. W.; PARK, M. IEEE 802.11ac: from Channelization To Multi-user MIMO. **IEEE Communications Magazine**, v. 51, n. 10, p. 84–90, Outubro 2013.

BELLORADO, J. et al. Coexistence of Ultra-wideband Systems with IEEE-802.11 a Wireless LANs. In: **Global Telecommunications Conference (GLOBECOM 2003)**. São Francisco, CA, EUA: IEEE, 2003.

BERLEMANN, L. et al. Coexistence and Interworking of IEEE 802.16 and IEEE 802.11e. In: **63rd Vehicular Technology Conference (VTC 2006)**. Melbourne, Vic., Austrália: IEEE, 2006. v. 1, p. 27–31.

BERMINGHAM, M. L. et al. Application of High-dimensional Feature Selection: Evaluation for Genomic Prediction in Man. **Scientific reports**, Nature Publishing Group, v. 5, p. 1–12, 2015.

BIANCHESI, N. M. P. et al. A Design of Experiments Comparative Study on Clustering Methods. **IEEE Access**, IEEE, v. 7, p. 167726–167738, 2019.

BIANCHI, G.; TINNIRELLO, I. Kalman Filter Estimation of The Number of Competing Terminals in an IEEE 802.11 Network. In: **IEEE Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003)**. São Francisco, CA, EUA: IEEE, 2003. v. 2, p. 844–852.

BIANCHI, G.; TINNIRELLO, I.; SCALIA, L. Understanding 802.11e Contention-based Prioritization Mechanisms and Their Coexistence with Legacy 802.11 Stations. **IEEE Network**, v. 19, n. 4, p. 28–34, Julho 2005.

BLOOM, N.; DAVIS, S. J.; ZHESTKOVA, Y. Covid-19 shifted patent applications toward technologies that support working from home. **University of Chicago, Becker Friedman Institute for Economics Working Paper**, n. 2020-133, p. 1–10, 2021.

BÖHNING, D. Maximum likelihood estimation of the logarithmic series distribution. **Statistische Hefte**, Springer, v. 24, n. 1, p. 121–140, 1983.

BORELLA, M. S. Measurement and interpretation of internet packet loss. **Journal of Communications and Networks**, v. 2, n. 2, p. 93–102, 2000.

BORGIO, M. et al. Analysis of The Hidden Terminal Effect in Multi-rate IEEE 802.11b Networks. In: **7th International Symposium on Wireless, Personal, and Multimedia Communications (WPMC 2004)**. Abano Terme, Padova, Itália: IEEE, 2004. v. 4.

CANDELA, M.; LUCONI, V.; VECCHIO, A. Impact of the COVID-19 pandemic on the Internet latency: A large-scale study. **Computer Networks**, Elsevier, v. 182, p. 107495, 2020.

CARDOSO, K. V.; REZENDE, J. F. D. Accurate Hidden Markov Modeling of Packet Losses in Indoor 802.11 Networks. **IEEE Communications Letters**, v. 13, n. 6, p. 417–419, Junho 2009.

CARVALHO, L.; ANGEJA, J.; NAVARRO, A. A New Packet Loss Model of the IEEE 802.11g Wireless Network for Multimedia Communications. **IEEE Transactions on Consumer Electronics**, v. 51, n. 3, p. 809–814, Agosto 2005.

CEN, S.; COSMAN, P. C.; VOELKER, G. M. End-to-end Differentiation of Congestion and Wireless Losses. **IEEE/ACM Transactions on Networking**, v. 11, n. 5, p. 703–717, Outubro 2003.

CHANG, N.; RASHIDZADEH, R.; AHMADI, M. Robust Indoor Positioning Using Differential Wi-fi access Points. **IEEE Transactions on Consumer Electronics**, v. 56, n. 3, p. 1860–1867, Agosto 2010.

CHEN, J.; CHAN, S. H. G.; LIEW, S.-C. Mixed-mode WLAN: The Integration of ad Hoc Mode with Wireless LAN Infrastructure. In: **Global Telecommunications Conference (GLOBECOM 2003)**. São Francisco, CA, EUA: IEEE, 2003. v. 1, p. 231–235.

CHEN, L.-J.; HUNG, H.-H. A Two-state Markov-based Wireless Error Model for Bluetooth Networks. **Wireless Personal Communications**, Springer, v. 58, n. 4, p. 657–668, 2011.

CHIASSERINI, C. F.; RAO, R. R. Performance of IEEE 802.11 WLANs in a Bluetooth Environment. In: **IEEE Wireless Communications and Networking Conference. Conference Record**. Chicago, IL, EUA: IEEE, 2000. v. 1, p. 94–99.

_____. Coexistence Mechanisms for Interference Mitigation Between IEEE 802.11 WLANs and Bluetooth. In: **Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies**. Nova York, NY, EUA: IEEE, 2002. v. 2, p. 590–598.

CHOK, N. S. **Pearson's Versus Spearman's and Kendall's Correlation Coefficients for Continuous Data**. Dissertação (Mestrado) — Escola de Saúde Pública, Universidade de Pittsburgh, 2010.

CHUNG, T. D. et al. Simulation of WirelessHART Networked Control System with Packet Dropout. In: **10th Asian Control Conference (ASCC)**. Kota Kinabalu, Malásia: IEEE, 2015. p. 1–6.

CISCO. **Cisco Annual Internet Report (2018-2023) White Paper**. 2020. <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>. [Online; acessado em 30-Julho-2020].

CONTI, A. et al. Bluetooth and IEEE 802.11b Coexistence: analytical Performance Evaluation in Fading Channels. **IEEE Journal on Selected Areas in Communications**, v. 21, n. 2, p. 259–269, Fevereiro 2003.

CORDEIRO, C. D. M.; AGRAWAL, D. P. Employing Dynamic Segmentation for Effective Co-located Coexistence Between Bluetooth and IEEE 802.11 WLANs. In: **Global Telecommunications Conference (GLOBECOM 2002)**. Taipei, Taiwan: IEEE, 2002. v. 1, p. 195–200.

CROVELLA, M. E.; BESTAVROS, A. Self-similarity in World Wide Web Traffic: Evidence and Possible Causes. **IEEE/ACM Transactions on Networking**, v. 5, n. 6, p. 835–846, Dezembro 1997.

CROW, B. P. et al. IEEE 802.11 wireless local area networks. **IEEE Communications Magazine**, v. 35, n. 9, p. 116–126, Setembro 1997.

DUANGSOITHONG, R.; WINDEATT, T. Relevant and redundant feature analysis with ensemble classification. In: **Seventh International Conference on Advances in Pattern Recognition**. Kolkata, Índia: IEEE, 2009. p. 247–250.

DUEL-HALLEN, A. Fading Channel Prediction for Mobile Radio Adaptive Transmission Systems. **Proceedings of the IEEE**, v. 95, n. 12, p. 2299–2313, Dezembro 2007.

ELAARAG, H. Improving TCP Performance over Mobile Networks. **ACM Computing Surveys (CSUR)**, ACM, v. 34, n. 3, p. 357–374, 2002.

ELLIOTT, E. O. Estimates of Error Rates for Codes on Burst-noise Channels. **The Bell System Technical Journal**, v. 42, n. 5, p. 1977–1997, Setembro 1963.

FENG, J.; LIU, Z.; JI, Y. Wireless Channel Loss analysis - a Case Study Using WiFi-Direct. In: **2014 International Wireless Communications and Mobile Computing Conference (IWCMC)**. Nicosia, Cyprus: IEEE, 2014. p. 244–249.

FERNANDES, M. V. D. S. **Modelagem de Erros em Surtos em Sistemas de Comunicações**. Dissertação (Mestrado) — Programa de Pós-Graduação em Engenharia Elétrica (PPGE), Instituto Militar De Engenharia (IME), 2002.

FOSTER, I. et al. End-to-end Quality of Service for High-end Applications. **Computer Communications**, Elsevier, v. 27, n. 14, p. 1375–1388, 2004.

FRITCHMAN, B. A Binary Channel Characterization Using Partitioned Markov Chains. **IEEE Transactions on Information Theory**, v. 13, n. 2, p. 221–227, Abril 1967.

FUKUHARA, T. et al. Channel occupancy estimation method for Wireless LAN. In: **IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications**. Tóquio, Japão: IEEE, 2009. p. 1163–1167.

GANDARILLAS, C. et al. Dynamic transmit-power control for wifi access points based on wireless link occupancy. In: **2014 IEEE Wireless Communications and Networking Conference (WCNC)**. Istanbul, Turquia: IEEE, 2014. p. 1093–1098.

GARETTO, M.; CHIASSERINI, C.-F. Performance analysis of 802.11 WLANs Under Sporadic Traffic. In: **International Conference on Research in Networking**. Waterloo, ON, Canadá: Springer, 2005. p. 1343–1347.

GETTYS, J.; NICHOLS, K. Bufferbloat: Dark buffers in the internet. **Queue**, ACM Nova York, NY, EUA, v. 9, n. 11, p. 40–54, 2011.

GILBERT, E. N. Capacity of a Burst-noise Channel. **The Bell System Technical Journal**, v. 39, n. 5, p. 1253–1265, Setembro 1960.

GIUSTINIANO, D. et al. Measuring transmission opportunities in 802.11 links. **IEEE/ACM Transactions on Networking**, v. 18, n. 5, p. 1516–1529, Outubro 2010.

GOLMIE, N.; CHEVROLLIER, N.; REBALA, O. Bluetooth and WLAN Coexistence: Challenges and Solutions. **IEEE Wireless Communications**, v. 10, n. 6, p. 22–29, Dezembro 2003.

GÓMEZ, D. et al. On The Modeling of a Realistic Wireless Channel By Means of a Hidden Markov Process. In: **2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)**. Barcelona, Espanha: IEEE, 2012. p. 397–402.

GONZALEZ, T. et al. Influence of COVID-19 confinement on students' performance in higher education. **PloS one**, Public Library of Science San Francisco, CA USA, v. 15, n. 10, p. 1–23, 2020.

GUHA, R. K.; SARKAR, S. Characterizing Temporal SNR Variation in 802.11 Networks. **IEEE Transactions on Vehicular Technology**, v. 57, n. 4, p. 2002–2013, Julho 2008.

HAMALAINEN, M.; TESI, R.; IINATTI, J. UWB Coexistence with IEEE 802.11a and UMTS in Modified Saleh-Valenzuela Channel. In: **2004 International Workshop on Ultra Wideband Systems Joint with Conference on Ultra Wideband Systems and Technologies**. Kyoto, Japão: IEEE, 2004. p. 45–49.

HAN, B.; LEE, S. Efficient Packet Error Rate Estimation in Wireless Networks. In: **2007 3rd International Conference on Testbeds and Research Infrastructure for the Development of Networks and Communities**. Atenas, Grécia: IEEE, 2007. p. 1–9.

HAN, J.; PEI, J.; KAMBER, M. **Data Mining: Concepts and Techniques**. Waltham, MA, EUA: Elsevier, 2011.

HAN, S. et al. Coexistence Performance Evaluation of IEEE 802.15.4 Under IEEE 802.11b Interference in Fading Channels. In: **IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications**. Atenas, Grécia: IEEE, 2007. p. 1–5.

HARTWELL, J. A.; FAPOJUWO, A. O. Modeling and Characterization of Frame Loss Process in IEEE 802.11 Wireless Local area Networks. In: **IEEE 60th Vehicular**

Technology Conference (VTC 2004). Los Angeles, CA, EUA: IEEE, 2004. v. 6, p. 4481–4485.

HASAN, S. F. et al. HMM-based modelling of roadside-to-vehicle WLAN communications. In: **7th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP 2010)**. Newcastle upon Tyne, UK: IEEE, 2010. p. 427–431.

HASSAN, M.; KRUNZ, M. M.; MATTA, I. Markov-based Channel Characterization for Tractable Performance analysis in Wireless Packet Networks. **IEEE Transactions on Wireless Communications**, v. 3, n. 3, p. 821–831, Maio 2004.

HASSLINGER, G.; HOHLFELD, O. The Gilbert-Elliott Model for Packet Loss in Real Time Services on the Internet. In: **14th GI/ITG Conference - Measurement, Modelling and Evaluation of Computer and Communication Systems**. Dortmund, Alemanha: IEEE, 2008. p. 1–15.

HAUER, J.-H.; HANDZISKI, V.; WOLISZ, A. Experimental Study of The Impact of WLAN Interference on IEEE 802.15. 4 Body area Networks. **Wireless Sensor Networks**, Springer, p. 17–32, 2009.

HAYAJNEH, T. et al. A Survey of Wireless Technologies Coexistence in WBAN: analysis and open Research Issues. **Wireless Networks**, Springer Science & Business Media, v. 20, n. 8, p. 2165–2199, 2014.

HOWITT, I. IEEE 802.11 and Bluetooth Coexistence analysis Methodology. In: **IEEE 53rd Vehicular Technology Conference (VTC 2001)**. Rhodes, Grécia: IEEE, 2001. v. 2, p. 1114–1118.

HOWITT, I.; MITTER, V.; GUTIERREZ, J. Empirical Study for IEEE 802.11 and Bluetooth Interoperability. In: **IEEE 53rd Vehicular Technology Conference (VTC 2001)**. Rhodes, Grécia: IEEE, 2001. v. 2, p. 1109–1113.

HOWITT, I.; SHUKLA, A. Coexistence Empirical Study and Analytical Model for Low-Rate WPAN and IEEE 802.11b. In: **2008 IEEE Wireless Communications and Networking Conference**. Las Vegas, NV, EUA: IEEE, 2008. p. 900–905.

HUANG, J. et al. Beyond Co-existence: Exploiting WiFi White Space for Zigbee Performance assurance. In: **The 18th IEEE International Conference on Network Protocols**. Kyoto, Japão: IEEE, 2010. p. 305–314.

HUO, H. et al. Coexistence Issues of 2.4GHz Sensor Networks with Other RF Devices at Home. In: **2009 Third International Conference on Sensor Technologies and Applications**. Atenas, Glyfada, Grécia: IEEE, 2009. p. 200–205.

HUO, H.-w. et al. Coexistence of 2.4 GHz Sensor Networks in Home Environment. **The Journal of China Universities of Posts and Telecommunications**, Elsevier, v. 17, n. 1, p. 9–18, 2010.

HWANG, T. et al. OFDM and Its Wireless Applications: A Survey. **IEEE Transactions on Vehicular Technology**, v. 58, n. 4, p. 1673–1694, Maio 2009.

ITU-T. **ITU-T Telecommunication Standardization Sector of ITU. Recommendation E.800 Definitions of Terms Related To Quality of Service**. 2008. <https://www.itu.int/rec/T-REC-E.800-200809-I>. [Online; acessado 10-Junho-2019].

JAMSHAID, K. et al. Deflating link buffers in a wireless mesh network. **Ad Hoc Networks**, Elsevier, v. 16, p. 266–280, 2014.

JUNG, B. H. et al. Interference Mediation for Coexistence of WLAN and ZigBee Networks. In: **IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications**. Cannes, França: IEEE, 2008. p. 1–5.

KAFI, M. A. et al. Congestion control protocols in wireless sensor networks: A survey. **IEEE Communications Surveys Tutorials**, v. 16, n. 3, p. 1369–1390, Março 2014.

KAMERMAN, A.; ERKOCEVIC, N. Microwave oven Interference on Wireless LANs operating in The 2.4 GHz ISM Band. In: **8th IEEE International Symposium on**

Personal, Indoor and Mobile Radio Communications. Helsinki, Finlândia: IEEE, 1997. v. 3, p. 1221–1227.

KANG, M. S. et al. Adaptive Interference-Aware Multi-Channel Clustering Algorithm in a ZigBee Network in The Presence of WLAN Interference. In: **2nd International Symposium on Wireless Pervasive Computing.** San Juan, Porto Rico: IEEE, 2007. p. 200–205.

KARANDE, S. et al. Analysis and Modeling of Errors at The 802.11b Link Layer. In: **International Conference on Multimedia and Expo.** Baltimore, MD, EUA: IEEE, 2003. v. 1, p. 673–674.

KARMAKAR, R.; CHATTOPADHYAY, S.; CHAKRABORTY, S. Dynamic link adaptation in ieee 802.11ac: A distributed learning based approach. In: **IEEE 41st Conference on Local Computer Networks (LCN 2016).** Dubai, Emirados Árabes Unidos: IEEE, 2016. p. 87–94.

_____. Impact of IEEE 802.11n/ac PHY/MAC High Throughput Enhancements on Transport and Application Protocols-A Survey. **IEEE Communications Surveys Tutorials**, v. 19, n. 4, p. 2050–2091, Fourthquarter 2017.

KELLNER, M. I.; MADACHY, R. J.; RAFFO, D. M. Software Process Simulation Modeling: Why? What? How? **Journal of Systems and Software**, Elsevier, v. 46, n. 2-3, p. 91–105, 1999.

KHAYAM, S. A.; RADHA, H. Constant-Complexity Models for Wireless Channels. In: **25th IEEE International Conference on Computer Communications (INFOCOM).** Barcelona, Espanha: IEEE, 2006. p. 1–11.

KHAYAM, S. A. et al. Markov and multifractal wavelet models for wireless MAC-to-MAC channels. **Performance Evaluation**, Elsevier, v. 64, n. 4, p. 298–314, 2007.

KODA, Y. et al. Time Series Measurement of IEEE 802.11ad Signal Power Involving Human Blockage with HMM-Based State Estimation. In: **IEEE 86th Vehicular Technology Conference (VTC 2017)**. Toronto, ON, Canadá: IEEE, 2017. p. 1–5.

KONRAD, A. et al. A markov-based channel model algorithm for wireless networks. **Wireless Networks**, Springer, v. 9, n. 3, p. 189–199, 2003.

KOPKE, A.; WILLIG, A.; KARL, H. Chaotic Maps as Parsimonious Bit Error Models of Wireless Channels. In: **22nd Annual Joint Conference of the IEEE Computer and Communications Societies**. São Francisco, CA, EUA: IEEE, 2003. v. 1, p. 513–523.

KRUNZ, M. M.; KIM, J. G. Fluid analysis of Delay and Packet Discard Performance for QoS Support in Wireless Networks. **IEEE Journal on Selected Areas in Communications**, v. 19, n. 2, p. 384–395, Fevereiro 2001.

KUHN, M.; JOHNSON, K. **Applied predictive modeling**. Nova York, NY, EUA: Springer, 2013.

LANEMAN, J. N.; TSE, D. N. C.; WORNELL, G. W. Cooperative Diversity in Wireless Networks: Efficient Protocols and outage Behavior. **IEEE Transactions on Information Theory**, v. 50, n. 12, p. 3062–3080, Dezembro 2004.

LANSFORD, J.; STEPHENS, A.; NEVO, R. Wi-Fi (802.11b) and Bluetooth: Enabling Coexistence. **IEEE Network**, v. 15, n. 5, p. 20–27, Setembro 2001.

LAUFER, R.; KLEINROCK, L. The Capacity of Wireless CSMA/CA Networks. **IEEE/ACM Transactions on Networking**, v. 24, n. 3, p. 1518–1532, Junho 2016.

LEE, K. K.; CHANSON, S. T. Packet Loss Probability for Real-time Wireless Communications. **IEEE Transactions on Vehicular Technology**, v. 51, n. 6, p. 1569–1575, Novembro 2002.

_____. Packet Loss Probability for Bursty Wireless Real-time Traffic Through Delay Model. **IEEE Transactions on Vehicular Technology**, v. 53, n. 3, p. 929–938, Maio 2004.

LI, Y. et al. A Novel Algorithm for Initial Cluster Center Selection. **IEEE Access**, v. 7, p. 74683–74693, Junho 2019.

LI, Z. et al. Modeling and Analysis of Distortion Caused By Markov-Model Burst Packet Losses in Video Transmission. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 19, n. 7, p. 917–931, Julho 2009.

_____. Modeling of Distortion Caused By Markov-model Burst Packet Losses in Video Transmission. In: **IEEE International Workshop on Multimedia Signal Processing**. Rio de Janeiro, RJ, Brasil: IEEE, 2009b. p. 1–6.

LIANG, C.-J. M. et al. Surviving Wi-Fi Interference in Low Power Zigbee Networks. In: **8th ACM Conference on Embedded Networked Sensor Systems**. Zurique, Suíça: ACM Digital Library, 2010. p. 309–322.

LIU, W. et al. Assessing The Coexistence of Heterogeneous Wireless Technologies with an SDR-Based Signal Emulator: A Case Study of Wi-Fi and Bluetooth. **IEEE Transactions on Wireless Communications**, v. 16, n. 3, p. 1755–1766, Março 2017.

LOU, H.-L. Implementing the viterbi algorithm. **IEEE Signal Processing Magazine**, v. 12, n. 5, p. 42–52, 1995.

LUGLIO, M. et al. Cross-Layer Architecture for a Satellite-Wi-Fi Efficient Handover. **IEEE Transactions on Vehicular Technology**, v. 58, n. 6, p. 2990–3001, 2009.

MATHEW, A. et al. Interference of 802.11b WLAN and Bluetooth: Analysis and Performance Evaluation. **AIRCC Publishing**, p. 140–150, 2010.

MCDUGALL, J.; MILLER, S. Sensitivity of Wireless Network Simulations To a Two-state Markov Model Channel Approximation. In: **Global Telecommunications Conference (GLOBECOM 2003)**. São Francisco, CA, EUA: IEEE, 2003. v. 2, p. 697–701.

MILOS, J. et al. Measurement Setup for Evaluation The Coexistence Between LTE Downlink and WLAN Networks. In: **2016 10th International Symposium on**

Communication Systems, Networks and Digital Signal Processing (CSNDSP).

Praga, República Checa: IEEE, 2016. p. 1–4.

MIRKOVIC, D.; ARMITAGE, G.; BRANCH, P. A Survey of Round Trip Time Prediction Systems. **IEEE Communications Surveys Tutorials**, v. 20, n. 3, p. 1758–1776, thirdquarter 2018.

MIU, A. et al. Divert: Fine-grained Path Selection for Wireless LANs. In: **2nd International Conference on Mobile Systems, Applications, and Services**. Boston, Massachusetts, EUA: ACM Digital Library, 2004. p. 203–216.

MO, Y.; GARONE, E.; SINOPOLI, B. LQG Control with Markovian Packet Loss. In: **European Control Conference (ECC)**. Zurique, Suíça: IEEE, 2013. p. 2380–2385.

MORSY, M. H. S.; SOWAILEM, M. Y. S.; SHALABY, H. M. H. Upper and lower bounds of burst loss probability for a core node in an optical burst switched network with pareto distributed arrivals. In: **17th International Conference on Telecommunications**. Doha, Qatar: IEEE, 2010. p. 523–527.

MUKHERJEE, S.; PENG, X.; GAO, Q. QoS Performances of IEEE 802.11 EDCA and DCF: A Testbed Approach. In: **5th International Conference on Wireless Communications, Networking and Mobile Computing**. Beijing, China: IEEE, 2009. p. 1–5.

MYOUNG, K.-J. et al. IEEE 802.11b Performance analysis in The Presence of IEEE 802.15.4 Interference. **IEICE Transactions on Communications**, IEEE, v. 90, n. 1, p. 176–179, 2007.

NAKATANI, T. et al. Estimating the Physical Distance Between two Locations with Wi-fi Received Signal Strength Information using Obstacle-aware Approach. **Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies**, v. 2, n. 3, p. 1–26, 2018.

NAKAYAMA, Y.; SEZAKI, K. Per-flow throughput fairness in ring aggregation network with multiple edge routers. **Big Data and Cognitive Computing**, v. 2, n. 3, p. 17, 2018.

NASIPURI, A.; ZHUANG, J.; DAS, S. R. A Multichannel CSMA MAC Protocol for Multihop Wireless Networks. In: **IEEE Wireless Communications and Networking Conference (WCNC)**. New Orleans, LA, EUA: IEEE, 1999. v. 3, p. 1402–1406.

NGUYEN, D. et al. Wireless Broadcast Using Network Coding. **IEEE Transactions on Vehicular Technology**, v. 58, n. 2, p. 914–925, Fevereiro 2009.

NGUYEN, H. X.; ROUGHAN, M. Rigorous Statistical analysis of Internet Loss Measurements. **IEEE/ACM Transactions on Networking (TON)**, IEEE Press, v. 21, n. 3, p. 734–745, 2013.

ONG, E. H. et al. IEEE 802.11ac: Enhancements for Very High Throughput WLANs. In: **2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications**. Toronto, ON, Canadá: IEEE, 2011. p. 849–853.

OPHIR, L.; BITRAN, Y.; SHERMAN, I. Wi-Fi (IEEE 802.11) and Bluetooth Coexistence: Issues and Solutions. In: **IEEE 15th International Symposium on Personal, Indoor and Mobile Radio Communications**. Barcelona, Espanha: IEEE, 2004. v. 2, p. 847–852.

PARK, J.-A. et al. Experiments on Radio Interference Between Wireless LAN and other Radio Devices on a 2.4 GHz ISM Band. In: **The 57th IEEE Semiannual Vehicular Technology Conference (VTC 2003)**. Jeju, Coréia do Sul: IEEE, 2003. v. 3, p. 1798–1801.

PERAHIA, E.; GONG, M. X. Gigabit Wireless LANs: an overview of IEEE 802.11ac and 802.11ad. **ACM SIGMOBILE Mobile Computing and Communications Review**, ACM, v. 15, n. 3, p. 23–33, 2011.

PETROVA, M. et al. Interference Measurements on Performance Degradation Between Colocated IEEE 802.11g/n and IEEE 802.15.4 Networks. In: **Sixth International Conference on Networking (ICN'07)**. Martinique, França: IEEE, 2007.

POKHREL, S. R. et al. TCP Performance over Wi-Fi: Joint Impact of Buffer and Channel Losses. **IEEE Transactions on Mobile Computing**, v. 15, n. 5, p. 1279–1291, Maio 2016.

POLLIN, S. et al. Harmful coexistence between 802.15.4 and 802.11: A measurement-based study. In: **3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2008)**. Singapura: IEEE, 2008. p. 1–6.

PRABHAKARAN, P.; SANKAR, R. Impact of Realistic Mobility Models on Wireless Networks Performance. In: **2006 IEEE International Conference on Wireless and Mobile Computing, Networking and Communications**. Montreal, Que., Canadá: IEEE, 2006. p. 329–334.

PRASAD, A. R. et al. Performance Evaluation, System Design and Network Deployment of IEEE 802.11. **Wireless Personal Communications**, Springer, v. 19, n. 1, p. 57–79, 2001.

QIU, L. et al. Troubleshooting Wireless Mesh Networks. **ACM SIGCOMM Computer Communication Review**, ACM, v. 36, n. 5, p. 17–28, 2006.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Áustria, 2019. Disponível em: <<https://www.R-project.org/>>.

RABINER, L.; JUANG, B. An Introduction To Hidden Markov Models. **IEEE ASSP Magazine**, v. 3, n. 1, p. 4–16, Janeiro 1986.

RAJAB, S. A.; BALID, W.; REFAI, H. H. Comprehensive study of spectrum occupancy for 802.11b/g/n homogeneous networks. In: **2015 IEEE International**

Instrumentation and Measurement Technology Conference. Pisa, Itália: IEEE, 2015. p. 1741–1746.

RAYANCHU, S. et al. Diagnosing Wireless Packet Losses in 802.11: Separating Collision from Weak Signal. In: **27th Conference on Computer Communications (INFOCOM 2008)**. Fênix, AZ, EUA: IEEE, 2008. p. 735–743.

RIVERA-LARA, E. J. et al. Analysis of the Relationship between QoS and SNR for an 802.11g WLAN. In: **2008 International Conference on Communication Theory, Reliability, and Quality of Service**. Bucarest, Romênia: IEEE, 2008. p. 103–107.

ROHLING, L. J. et al. Evidências da Falha do Modelo de Gilbert-Elliott em Redes sem Fio. In: **XXXIV Simpósio Brasileiro de Telecomunicações - SBrT2016**. Santarém, PA, Brasil: Sociedade Brasileira de Telecomunicações, 2016. p. 488–492.

ROY, M.; S., J. H. Performance Analysis of MQAM-OFDM Based WLAN in Presence of Zigbee Interference in AWGN and Rayleigh Fading Channel. In: **Sixth International Conference on Information Technology: New Generations**. Las Vegas, NV, EUA: IEEE, 2009. p. 1178–1183.

ROZNER, E. et al. ER: Efficient Retransmission Scheme for Wireless LANs. In: **Proceedings of The 2007 ACM CoNEXT Conference**. Nova York, NY, EUA: ACM Digital Library, 2007. p. 1–12.

RUBINSTEIN, M. G.; REZENDE, J. F. Qualidade De Serviço Em Redes 802.11. In: **XX Simpósio Brasileiro de Redes De Computadores (SBRC2002)**. Búzios, RJ, Brasil: Sociedade Brasileira de Computação, 2002. p. 1–16.

RUSS, S. H.; HAGHANI, S. 802.11g Packet-loss Behavior at High Sustained Bit Rates in The Home. **IEEE Transactions on Consumer Electronics**, v. 55, n. 2, p. 788–791, Maio 2009.

SALIH, O. S. et al. Hidden Markov Models for Packet-level Errors in Bursty Digital Wireless Channels. In: **2009 Loughborough Antennas Propagation Conference**. Loughborough, UK: IEEE, 2009. p. 385–388.

SANNECK, H. A.; CARLE, G. Framework Model for Packet Loss Metrics Based on Loss Runlengths. In: **Electronic Imaging**. San Jose, CA, EUA: SPIE Digital Library, 1999. p. 177–187.

SARKAR, T. K. et al. A Survey of Various Propagation Models for Mobile Communication. **IEEE Antennas and Propagation Magazine**, v. 45, n. 3, p. 51–82, Junho 2003.

SCHUSTER-BÖCKLER, B.; BATEMAN, A. An Introduction to Hidden Markov Models. **Current Protocols in Bioinformatics**, v. 18, n. 1, p. 1–9, 2007.

SENTHILKUMAR, S. et al. Hidden markov model based channel selection framework for cognitive radio network. **Computers & Electrical Engineering**, Elsevier, v. 65, p. 516–526, 2018.

SEQUEIRA, L. et al. The Influence of the Buffer Size in Packet Loss for Competing Multimedia and Bursty Traffic. In: **International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)**. Toronto, ON, Canadá: IEEE, 2013. p. 134–141.

SGORA, A.; VERGADOS, D. J.; VERGADOS, D. D. A Survey of TDMA Scheduling Schemes in Wireless Multihop Networks. **ACM Computing Surveys (CSUR)**, ACM, v. 47, n. 3, p. 53, 2015.

SHIN, S. Y. et al. Lecture Notes in Computer Science: Packet Error Rate analysis of IEEE 802.15.4 Under IEEE 802.11b Interference. **Wired/Wireless Internet Communications**, Springer, p. 618–618, 2005.

_____. Packet Error Rate Analysis of ZigBee Under WLAN and Bluetooth Interferences. **IEEE Transactions on Wireless Communications**, v. 6, n. 8, p. 2825–2830, Agosto 2007.

SHIN, S. Y.; PARK, H. S.; KWON, W. H. Mutual Interference analysis of IEEE 802.15.4 and IEEE 802.11b. **Computer Networks**, Elsevier, v. 51, n. 12, p. 3338–3353, 2007.

_____. Packet Error Rate analysis of IEEE 802.15.4 Under Saturated IEEE 802.11b Network Interference. **IEICE Transactions on Communications**, v. 90, n. 10, p. 2961–2963, 2007.

SHOWAIL, A.; JAMSHAD, K.; SHIHADA, B. An empirical evaluation of bufferbloat in IEEE 802.11n wireless networks. In: **IEEE Wireless Communications and Networking Conference (WCNC)**. Istanbul, Turquia: IEEE, 2014. p. 3088–3093.

SHUAIB, K. et al. Co-existence of Zigbee and WLAN, A Performance Study. In: **Wireless Telecommunications Symposium**. Pomona, CA, EUA: IEEE, 2006. p. 1–6.

SIKORA, A.; GROZA, V. F. Coexistence of IEEE802.15.4 with other Systems in The 2.4 GHz-ISM-Band. In: **IEEE Instrumentation and Measurement Technology Conference Proceedings**. Ottawa, Ont., Canadá: IEEE, 2005. v. 3, p. 1786–1791.

SILVA, C. A. G. d.; PEDROSO, C. M. MAC-Layer Packet Loss Models for Wi-Fi Networks: A Survey. **IEEE Access**, v. 7, p. 180512–180531, 2019.

SILVA, C. A. G. d.; SANTOS, E. L. d.; OSINSKI, C. Wireless Industrial Networks under Interference Conditions based on IEEE 802.15.4. In: **2018 13th IEEE International Conference on Industry Applications (INDUSCON)**. São Paulo, SP, Brasil: IEEE, 2018. p. 169–173.

SILVA, C. A. G. da; SANTOS, E. L. dos; PEDROSO, C. M. Caracterização de Perda de Pacotes em Redes Wi-Fi com Modelos Ocultos de Markov em Camadas 1 e 2. In: **XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBrT2020**. Florianópolis, SC, Brasil: SBrT, 2020. p. 1–5.

SILVEIRA, F.; SILVA, E. d. S. e. Predicting Packet Loss Statistics with Hidden Markov Models for FEC Control. **Computer Networks**, Elsevier, v. 56, n. 2, p. 628–641, 2012.

SINGH, V.; SHARMA, R.; TOMAR, M. S. An Analytical Study of Interference Problem Between ZigBee and WI-FI. In: **2013 International Conference on Communication Systems and Network Technologies**. Gwalior, Índia: IEEE, 2013. p. 257–261.

STALLINGS, W. IEEE 802.11: Wireless LANs from a to n. **IT Professional**, v. 6, n. 5, p. 32–37, 2004.

STUDER, C.; FATEH, S.; SEETHALER, D. A 757Mb/s 1.5 Mm² 90nm CMOS Soft-input Soft-output MIMO Detector for IEEE 802.11n. In: **Proceedings of European Solid-State Circuits Conference (ESSCIRC)**. Sevilha, Espanha: IEEE, 2010. p. 530–533.

TANG, C.; MCKINLEY, P. K. Modeling multicast packet losses in wireless lans. In: **6th ACM International Workshop on Modeling analysis and Simulation of Wireless and Mobile Systems**. San Diego, CA, EUA: ACM Digital Library, 2003. p. 130–133.

TANG, M. F. et al. Packet loss rate of an optical burst switch with nonlinear optical loop mirrors. **IEICE Electronics Express**, The Institute of Electronics, Information and Communication Engineers, v. 3, n. 11, p. 243–248, 2006.

TAO, T.; LU, J.; CHUANG, J. Hierarchical Markov Model for Burst Error analysis in Wireless Communications. In: **53rd Vehicular Technology Conference (VTC 2001)**. Rhodes, Grécia: IEEE, 2001. v. 4, p. 2843–2847.

TINNIRELLO, I.; BIANCHI, G. Interference Estimation in IEEE 802.11 Networks. **IEEE Control Systems**, IEEE, v. 30, n. 2, p. 30–43, 2010.

TOLEDO, T.; KATZ, R. State dependence in lane-changing models. **Transportation research record**, SAGE Publications Sage, Los Angeles, CA, EUA, v. 2124, n. 1, p. 81–88, 2009.

TRAMARIN, F. et al. On The Use of IEEE 802.11n for Industrial Communications. **IEEE Transactions on Industrial Informatics**, v. 12, n. 5, p. 1877–1886, Outubro 2016.

TSERTOU, A.; LAURENSEN, D. I. Revisiting the Hidden Terminal Problem in a CSMA/CA Wireless Network. **IEEE Transactions on Mobile Computing**, v. 7, n. 7, p. 817–831, Julho 2008.

URMELA, S.; NANDHINI, M. Collective Dendrogram Clustering with Collaborative Filtering for Distributed Data Mining on Electronic Health Records. In: **Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)**. Coimbatore, Índia: IEEE, 2017. p. 1–5.

VARSHNEY, U. The Status and Future of 802.11-based WLANs. **Computer**, v. 36, n. 6, p. 102–105, Junho 2003.

VISSER, I.; SPEEKENBRINK, M. et al. depmixs4: an r package for hidden markov models. **Journal of Statistical Software**, v. 36, n. 7, p. 1–21, 2010.

VITERBI, A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. **IEEE Transactions on Information Theory**, v. 13, n. 2, p. 260–269, 1967.

WANG, C.-X.; XU, W. Packet-level Error Models for Digital Wireless Channels. In: **IEEE International Conference on Communications (ICC 2005)**. Seul, Coréia do Sul: IEEE, 2005. v. 4, p. 2184–2189.

_____. A New Class of Generative Models for Burst-Error Characterization in Digital Wireless Channels. **IEEE Transactions on Communications**, v. 55, n. 3, p. 453–462, Março 2007.

WANG, H. S.; MOAYERI, N. Finite-state Markov Channel-a Useful Model for Radio Communication Channels. **IEEE Transactions on Vehicular Technology**, v. 44, n. 1, p. 163–171, Fevereiro 1995.

WANG, X.; CAI, L. Interference Analysis of Co-Existing Wireless Body Area Networks. In: **IEEE Global Telecommunications Conference (GLOBECOM 2011)**. Houston, TX, EUA: IEEE, 2011. p. 1–5.

WILLIG, A. A New Class of Packet- and Bit-level Models for Wireless Channels. In: **13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications**. Lisboa, Portugal: IEEE, 2002. v. 5, p. 2434–2440.

WILLIG, A. et al. Measurements of a Wireless Link in an Industrial Environment Using an IEEE 802.11-compliant Physical Layer. **IEEE Transactions on Industrial Electronics**, v. 49, n. 6, p. 1265–1282, Dezembro 2002.

WINTER, J. M. et al. Coexistence Issues in Wireless Networks for Factory automation. In: **12th IEEE International Conference on Industrial Informatics (INDIN)**. Porto Alegre, RS, Brasil: IEEE, 2014. p. 370–375.

_____. Wireless Coexistence and Spectrum Sensing in Industrial Internet of Things: An Experimental Study. **International Journal of Distributed Sensor Networks**, v. 11, n. 11, p. 1–12, 2015.

WIRTZ, H. et al. High-performance, energy-efficient mobile wireless networking in 802.11 infrastructure mode. In: **IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems**. Filadélfia, PA, EUA: IEEE, 2014. p. 291–299.

WOLTER, K. et al. Ph-distributed Fault Models for Mobile Communication. In: **The Winter Simulation Conference (WSC)**. Berlin, Alemanha: IEEE, 2012. p. 429.

WU, J. et al. Bandwidth-Efficient Multipath Transport Protocol for Quality-Guaranteed Real-Time Video Over Heterogeneous Wireless Networks. **IEEE Transactions on Communications**, v. 64, n. 6, p. 2477–2493, 2016.

XIAO, J. et al. A Survey on Wireless Indoor Localization from the Device Perspective. **ACM Computing Surveys (CSUR)**, ACM, v. 49, n. 2, p. 25, 2016.

XIE, L. et al. Performance Analysis of IEEE 802.11 DCF in Multi-hop Ad Hoc Networks. In: **Eighth IEEE/ACIS International Conference on Computer and Information Science**. Shanghai, China: IEEE, 2009. p. 222–227.

YADAV, S.; SINGH, D. A Survey on Congestion Control Mechanism in Multi-hop Wireless Network. In: **3rd International Conference on Computing for Sustainable Global Development (INDIACom)**. Nova Delhi, Índia: IEEE, 2016. p. 683–688.

YAJNIK, M. et al. Measurement and modeling of the temporal dependence in packet loss. In: **18th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)**. Nova York, NY, EUA: IEEE, 1999. v. 1, p. 345–352.

YANG, D.; XU, Y.; GIDLUND, M. Wireless Coexistence Between IEEE 802.11 and IEEE 802.15.4 based Networks: A Survey. **International Journal of Distributed Sensor Networks**, v. 7, n. 1, p. 1–17, 2011.

YANG, P. et al. Taming Cross-Technology Interference for Wi-Fi and ZigBee Coexistence Networks. **IEEE Transactions on Mobile Computing**, v. 15, n. 4, p. 1009–1021, Abril 2016.

YOON, D. G. et al. Packet Error Rate Analysis of IEEE 802.11b Under IEEE 802.15.4 Interference. In: **IEEE 63rd Vehicular Technology Conference**. Melbourne, Vic., Austrália: IEEE, 2006. v. 3, p. 1186–1190.

_____. Performance analysis of IEEE 802.11b Under Multiple IEEE 802.15.4 Interferences. In: **Wired/Wireless Internet Communications**. Coimbra, Portugal: Springer, 2007. p. 213–222.

YOUNESIAN, E. et al. Packet-loss Modeling for Multi-radio Wireless Sensor Networks. In: **IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)**. Larnaca, Chipre: IEEE, 2014. p. 673–678.

YU, Y.; MILLER, S. L. A Four-State Markov Frame Error Model for the Wireless Physical Layer. In: **IEEE Wireless Communications and Networking Conference**. Kowloon, China: IEEE, 2007. p. 2053–2057.

YUAN, W.; WANG, X.; LINNARTZ, J. P. M. G. A Coexistence Model of IEEE 802.15.4 and IEEE 802.11b/g. In: **14th IEEE Symposium on Communications and Vehicular Technology in The Benelux**. Delft, Holanda: IEEE, 2007. p. 1–5.

YUN, J. et al. A Channel Switching Scheme for Avoiding Interference of Between IEEE 802.15.4 and Other Networks. In: **International Multi-symposiums on Computer and Computational Sciences**. Shanghai, China: IEEE, 2008. p. 136–139.

ZACHARIAS, S. et al. Coexistence Measurements and analysis of IEEE 802.15.4 with Wi-Fi and Bluetooth for Vehicle Networks. In: **12th International Conference on ITS Telecommunications**. Taipei, Taiwan: IEEE, 2012. p. 785–790.

ZARIKOFF, B. W.; LEITH, D. J. Measuring Pulsed Interference in 802.11 Links. **IEEE/ACM Transactions on Networking**, v. 21, n. 2, p. 509–521, Abril 2013.

ZHANG, D. et al. Survey on Coexistence of Heterogeneous Wireless Networks in 2.4 GHz and TV White Spaces. **International Journal of Distributed Sensor Networks**, v. 13, n. 4, p. 1–47, 2017.

ZHANG, X.; SHIN, K. G. A Case for the Coexistence of Heterogeneous Wireless Networks. In: **The 3rd ACM Workshop on Wireless of The Students, By The Students, for the Students**. Las Vegas, Nevada, EUA: ACM Digital Library, 2011a. p. 1–4.

_____. Enabling Coexistence of Heterogeneous Wireless Systems: Case for ZigBee and WiFi. In: **The 12th ACM International Symposium on Mobile Ad Hoc Networking and Computing**. Paris, França: ACM Digital Library, 2011b. p. 1–6.

ZHU, H. et al. A Survey of Quality of Service in IEEE 802.11 Networks. **IEEE Wireless Communications**, v. 11, n. 4, p. 6–14, 2004.