



COPPE/UFRJ

REDES SEM FIO DE LONGA DISTÂNCIA: NOVAS CONTRIBUIÇÕES PARA
A JUSTIÇA EM NÍVEL DE USUÁRIO E PARA A QUALIDADE DE SERVIÇO
EM FLUXOS DE VÍDEO ESCALÁVEL

Guilherme Dutra Gonzaga Jaime

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientadores: Edmundo Albuquerque de
Souza e Silva
Rosa Maria Meri Leão
José Roberto Boisson de
Marca

Rio de Janeiro
Novembro de 2010

REDES SEM FIO DE LONGA DISTÂNCIA: NOVAS CONTRIBUIÇÕES PARA
A JUSTIÇA EM NÍVEL DE USUÁRIO E PARA A QUALIDADE DE SERVIÇO
EM FLUXOS DE VÍDEO ESCALÁVEL

Guilherme Dutra Gonzaga Jaime

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Edmundo Albuquerque de Souza e Silva, Ph.D.

Prof. Rosa Maria Meri Leão, Dr.

Prof. José Roberto Boisson de Marca, Ph.D.

Prof. Valmir Carneiro Barbosa, Ph.D.

Prof. José Ferreira de Rezende, Dr.

Prof. Antônio Alfredo Ferreira Loureiro, Ph.D.

RIO DE JANEIRO, RJ – BRASIL

NOVEMBRO DE 2010

Jaime, Guilherme Dutra Gonzaga

Redes sem fio de longa distância: Novas contribuições para a justiça em nível de usuário e para a qualidade de serviço em fluxos de vídeo escalável/Guilherme Dutra Gonzaga Jaime. – Rio de Janeiro: UFRJ/COPPE, 2010.

XX, 148 p.: il.; 29, 7cm.

Orientadores: Edmundo Albuquerque de Souza e Silva

Rosa Maria Meri Leão

José Roberto Boisson de Marca

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2010.

Referências Bibliográficas: p. 133 – 148.

1. Avaliação de desempenho. 2. Redes sem fio. 3. Aplicações de fluxo de vídeo. 4. Perda de pacotes. 5. Qualidade de serviço. 6. PSNR. 7. Vazão. 8. Cadeias de Markov. I. Silva, Edmundo Albuquerque de Souza e *et al.*. *et al.*. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

*Aos meus pais, avós, irmãos e à
Larissa. Vocês são as pedras
angulares de minha vida.*

Agradecimentos

Esta tese de doutorado é o resultado de um projeto de vida que levou 10 anos, 7 meses e 24 dias. Desde 13 de março de 2000, quando cheguei ao LAND, tive a sorte, o prazer e a honra de ser recebido em um ambiente extremamente fértil e colaborativo. Por isso agradeço aos professores Paulo André Santanna Perez e Sibélius Lellis por terem apostado em mim ao me recomendarem aos professores Edmundo e Rosa para o processo de seleção do mestrado da COPPE/UFRJ. Essas cartas de recomendação foram as portas de entrada para este projeto de uma década que resultou em diversos artigos publicados nas principais revistas e conferências da área de Ciência de Computação do Brasil e do mundo.

Ao chegar à UFRJ encontrei colegas verdadeiramente brilhantes. O convívio com estas pessoas me permitiu absorver parte de seu conhecimento, me tornando uma pessoa bem mais experiente e segura. Considero que este foi um fator super importante em minha formação. Por isso, agradeço aos antigos e novos colegas de laboratório e amigos Flávio, Bernardo, Allyson, Fernando, Sadoc, Magnos, Drika, Kelvin, Ana Paula, Bruno Ribeiro, Guto, Boechat, Carolzinha, João Victor, Hugo, Ebenezer, Fabrício, Ede, Watanabe, GDom, Gaspare, Zhou Xiang e Thothadri.

Ao Flavio agradeço especialmente pela ampla passagem de conhecimento em relação ao Linux e à ferramenta Tangram-II. Sem eles, minhas colaborações ao laboratório e minha própria tese não teriam sido os mesmos. Também agradeço especialmente ao Gaspare pela colaboração na parte de experimentações com vídeos escaláveis.

Aos professores Edmundo, Rosa e Boisson deixo meus sinceros e já saudosos agradecimentos. Desde o início, eles sempre prezaram pelo caráter, pela correteza, pela qualidade e, também, pela dedicação e trabalho duro. Espero que esta tese retribua minimamente o fato de terem investido recursos preciosos como tempo, estrutura

laboratorial e confiança. Foi um prazer imensurável participar da equipe LAND como aluno, como administrador do laboratório, como pesquisador e como uma pessoa que compartilha o mesmo alicerce de vocês: o objetivo principal de, através da disseminação da educação de excelência e de trabalho árduo e correto, fazer nossa parte para tornar nosso querido país um lugar cada vez melhor, colaborando de fato para a nossa sociedade.

No centro de toda a estrutura que forma nosso grupo de pesquisa está a Carol, que sempre foi muito além de suas atribuições para fazer o possível e o impossível em prol dos alunos que são como filhos adotivos. Não haveria espaço nesta seção de agradecimentos para listar todo tipo de colaboração importante desempenhadas pela Carol. De maneira análoga, não seria possível agradecer o suficiente. Me limito, portanto, a agradecer carinhosamente a ela. Desejo que continue sempre fazendo este trabalho, que é muitíssimo importante para todos os que estão à sua volta. Como qualquer mãe que sofre ao ver seus filhos partindo para o mundo, a Carol sempre sofre quando mais um aluno vai embora. Felizmente eu ficarei no Rio de Janeiro e frequentarei o LAND por um bom tempo, e poderei sempre revê-la e trazer mais *souveneurs* de países e cidades exóticas, enriquecendo a coleção guardada em sua sala.

O grupo de pessoas citadas nestes agradecimentos fez e faz parte de um laboratório (LAND) motivado e bem equipado, mantido pela UFRJ e por verbas de bolsas e projetos financiados por instituições como o CNPq, a Capes, o CPqD, a FAPERJ, a RNP e a FINEP. Em particular, minha bolsa de mestrado foi financiada pela Capes, o projeto SBTvD, em que participei como pesquisador entre o mestrado e o doutorado, foi financiado pela FINEP, e minha bolsa de doutorado foi obtida através do CNPq. Com isso, agradeço a todas as estas instituições e ao governo brasileiro por ter me oferecido educação de excelência, comparável aos melhores centros de tecnologia do mundo.

Por fim, gostaria destacar que considero que este projeto de pouco mais de 10 anos está terminando com sucesso maior do que eu imaginaria possível. O ambiente descrito durante estes agradecimentos nunca ofereceu resistências ou impedimentos ao meu desenvolvimento. É muito gratificante terminar esta jornada tendo a consciência de que seguí crescendo aceleradamente até onde meus próprios limites

permitiram. Mais do que isso, todas as pessoas citadas nesta seção: professores, a Carol, e colegas alunos, colaboraram para que minhas limitações fossem levadas muito adiante.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

REDES SEM FIO DE LONGA DISTÂNCIA: NOVAS CONTRIBUIÇÕES PARA A JUSTIÇA EM NÍVEL DE USUÁRIO E PARA A QUALIDADE DE SERVIÇO EM FLUXOS DE VÍDEO ESCALÁVEL

Guilherme Dutra Gonzaga Jaime

Novembro/2010

Orientadores: Edmundo Albuquerque de Souza e Silva

Rosa Maria Meri Leão

José Roberto Boisson de Marca

Programa: Engenharia de Sistemas e Computação

Esta tese tem como objetivo a avaliação de desempenho do acesso sem fio à Internet para o uso de aplicações web e de vídeo escalável. O texto trata das principais contribuições alcançadas por este trabalho, que são: (i) o desenvolvimento de um novo modelo de simulação de uma célula EVDO com o qual avaliamos justiça usando métricas como a vazão e o atraso, sob influência de diferentes cenários de mobilidade. Além disso, propomos uma solução simples para melhorar a justiça entre os usuários deste tipo de rede sem fio; e, (ii) a proposta de modelos matemáticos que nos permitem estudar, sob um novo ponto de vista, o processo aleatório de perda em filas *FIFO/Droptail*. Com base nos resultados obtidos, propomos uma nova técnica fim-a-fim de transmissão de fluxos de vídeo escalável. Mostramos com essa técnica que, mesmo sem o emprego de complexos algoritmos de escalonamento e descarte nos roteadores, é possível priorizar os pacotes mais importantes para a qualidade do vídeo, sem o uso de métodos tradicionais de priorização. Através de modelos de simulação desenvolvidos durante esta tese, avaliamos o ganho de qualidade atingido quando nossa técnica é usada, considerando perfis de vídeo bem conhecidos.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

LONG DISTANCE WIRELESS NETWORKS: ON THE USER-LEVEL
FAIRNESS AND QUALITY OF SERVICE OF SCALABLE VIDEO
STREAMING

Guilherme Dutra Gonzaga Jaime

November/2010

Advisors: Edmundo Albuquerque de Souza e Silva

Rosa Maria Meri Leão

José Roberto Boisson de Marca

Department: Systems Engineering and Computer Science

This thesis focuses on Internet wireless access performance evaluation considering web browsing and scalable video streaming. The text describes the main contributions achieved by this work, which are: (i) the development of a new EVDO cell simulation model capable of assessing performance metrics such as throughput, delay and user-level fairness. We then propose a simple way to overcome the main quality of service weaknesses pointed by our simulation experiments; (ii) the proposal of mathematical models that allow us to assess the random loss process of FIFO/Droptail queues, with a new point of view. Based on what was learned, we propose a new end-to-end scalable video streaming technique. We show that, even without the use of more complex scheduling and dropping algorithms, it is possible to achieve a significant level of prioritization for the most important packets, when compared to the conventional technique used to transmit a packet burst. We also develop simulation models to analyze the video quality gain when our technique is in place, considering well-known video profiles.

Sumário

Lista de Figuras	xiii
Lista de Tabelas	xvi
1 Introdução	1
1.1 Visão Geral	1
1.2 Motivação e Objetivos da Tese	4
1.2.1 Um estudo do efeito das limitações da camada física sobre a QoS	5
1.2.2 Proposta de um novo método para transmissão de fluxos de vídeo escalável sobre redes sem fio	6
1.3 Organização desta tese	10
2 Fundamentos teóricos	12
2.1 O que diferencia as redes sem fio das redes cabeadas tradicionais? . . .	13
2.1.1 Introdução	13
2.1.2 Camada Física: impedimentos e o desvanecimento lento	14
2.1.3 Camada Física: desvanecimento rápido	16
2.1.4 Técnicas de adaptação aos impedimentos do meio físico	17
2.2 Visão geral do padrão EVDO rev.0	21
2.3 Cadeias de Markov	26
2.4 Uma breve apresentação sobre codificação de vídeo escalável	28
2.5 Conceitos de Justiça	31
2.6 Conceitos de Justiça	32
3 Trabalhos Relacionados	34

3.1	EVDO rev.0: Desempenho e justiça do padrão de acesso EVDO rev.0	35
3.1.1	Algoritmos de escalonamento para redes sem fio de taxa variável	35
3.1.2	Modelagem e análise de desempenho do padrão EVDO	40
3.2	Vídeo escalável: Escalonamento, descarte e perda	43
3.2.1	Adaptação da taxa no servidor	43
3.2.2	Escalonamento e descarte de pacotes com prioridade aplicados a fluxo de vídeo escalável	44
3.2.3	Discussão	48
4	Modelo de uma célula EVDO, Experimentos e Resultados	49
4.1	Modelo de uma célula EVDO	49
4.1.1	Visão geral do Modelo	49
4.1.2	Camada Física	51
4.1.3	Camada de enlace	59
4.1.4	Modelo dos Usuários	61
4.1.5	Modelo de Mobilidade	63
4.2	Experimentos e resultados	65
4.2.1	Cenários de experimentos e resultados obtidos para o modelo EVDO	66
4.2.2	Experimentos realizados e resultados obtidos	68
4.2.3	Discussão: Há como melhorar a qualidade de serviço em relação à justiça (<i>goodput</i> e atraso)?	74
4.2.4	Experimentos: Melhorando a justiça em nível de usuário	77
5	Uma proposta para priorização de tráfego de vídeo escalável	81
5.1	Enquadramento de pacotes e definição de nível de rajada	82
5.2	Proposta de um modelo analítico para avaliação de perda intra-rajada em fluxos de vídeo	84
5.2.1	Descrição do modelo	85
5.2.2	Definição das medidas de interesse	88
5.2.3	Obtendo solução para o modelo	90
5.2.4	Parâmetros considerados	93
5.2.5	Resultados e discussão	94

5.3	Modelo de simulação A: vídeos reais	100
5.3.1	Descrição do modelo	100
5.3.2	Definição das medidas de interesse	106
5.3.3	Parâmetros considerados	107
5.3.4	Resultados e discussão	111
5.4	Modelo de simulação B: vídeos reais, traces de uma rede EVDO . . .	121
5.4.1	Descrição do modelo	121
5.4.2	Parâmetros considerados	122
5.4.3	Resultados e discussão	124
6	Conclusões e trabalhos futuros	128
6.1	Sobre a justiça em redes sem fio de longa distância	128
6.2	Sobre a qualidade de serviço de fluxos de vídeos escaláveis em redes sem fio	130
	Referências Bibliográficas	133

Lista de Figuras

1.1	A qualidade do canal varia em diferentes escalas de tempo	2
1.2	Cenário geral considerado nesta fase do trabalho	7
2.1	O desvanecimento rápido flutua em torno de uma média local dada pelo desvanecimento lento	14
2.2	Reflexão, espalhamento e difração	15
2.3	(a) Canal direto; (b) Estrutura do canal direto	22
2.4	(a) Canal reverso; (b) Estrutura do canal reverso	25
2.5	Quando a taxa de um enlace sem fio cai abaixo do requerido pelo vídeo, o <i>buffer</i> do cliente é consumido, podendo cair a zero, o que gera problemas de continuidade (pausas)	29
2.6	Codificação escalável vs. codificação não escalável	30
2.7	Exemplo da estrutura de quadros de vídeos codificados usando escalabilidade FGS	31
4.1	Objetos implementados	50
4.2	Modelo de geometrias entre células	54
4.3	Visão Geral do modelo dos usuários	61
4.4	Visão geral do modelo <i>Random Waypoint</i>	64
4.5	<i>Goodput</i> médio em função da população	69
4.6	Atraso médio em função da população	70
4.7	<i>Goodput</i> em função da população por anel de distância	71
4.8	Atraso em função da população e do anel de distância em relação à ERB	72
4.9	<i>Goodput</i> médio percebido pelo usuário em função da população considerada (usuários móveis e estáticos vs. estáticos)	73

4.10	Atraso médio observado em função da população considerada (usuários móveis e estáticos vs. estáticos)	74
4.11	Atraso médio (60 usuários) para os cenários 1 e 2, respectivamente . .	75
4.12	<i>Goodput</i> médio (60 usuários) para os cenários 1 e 2, respectivamente .	76
4.13	Atraso médio em função da distância para cada cenário considerado .	77
4.14	<i>Goodput</i> médio em função da distância para cada cenário considerado	78
4.15	<i>Goodput</i> em função de α	79
4.16	Caminho amostral de uma simulação: <i>goodput</i> de um usuário em função de sua distância à ERB	79
5.1	(a) Esquema de enquadramento proposto; (b) esquema de trans- missão de pacotes proposto	82
5.2	Três exemplos ilustrando nossa definição de nível de rajada e a relação entre λ e β	83
5.3	Cenário geral considerado neste capítulo	84
5.4	Modelo analítico proposto: visão geral	86
5.5	Três exemplos ilustrando nossa definição de nível de rajada e a relação entre λ e β	87
5.6	Representação da cadeia de Markov do modelo analítico	89
5.7	Exemplo de modelo para $n=4$ e $q=2$	90
5.8	(a) Fração de perda do tráfego de vídeo, $q=25$; (b) Distribuição da perda intra-rajada, para $\rho = 0.99$ e $q=25$	95
5.9	(a) Fração de perda total (F) (normalizado), para valores menores de b e $q=25$; (b) Distribuição da perda intra-rajada ($\rho = 0.99$), para valores menores de b e $q=25$	96
5.10	($\rho = 0.99$): Fração de perda intra-rajada (F^i) em função do nível de rajada b	98
5.11	($\rho = 1.105$): Fração de perda intra-rajada (F^i) em função do nível de rajada b	99
5.12	($\rho = 0.89$): Fração de perda intra-rajada (F^i) em função do nível de rajada b	100
5.13	Visão geral do modelo de simulação A	101
5.14	(a) Simulação: enquadramento; (b) transmissão de pacotes	102

5.15	Estrutura dos experimentos de simulação: desde (I) a codificação até (VII) o cálculo do PSNR ao final da transmissão	105
5.16	Pacotes por quadro para 4 perfis de vídeo escalável	109
5.17	Pacotes por quadro para o perfil de vídeo escalável Star Wars IV . . .	110
5.18	Perfil Football: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$	111
5.19	Perfil Foreman: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$	112
5.20	Perfil Mobile: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$	113
5.21	Perfil Crew: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$	113
5.22	Perfil Star Wars IV: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$	114
5.23	Dependência intra-quadro (a) e inter-quadro (b)	115
5.24	Perfil Football: histograma do PSNR quadro-a-quadro	116
5.25	Perfil Crew: histograma do PSNR quadro-a-quadro	118
5.26	Foreman: histograma do PSNR quadro-a-quadro	119
5.27	Perfil Mobile: histograma do PSNR quadro-a-quadro	120
5.28	StarWars IV: histograma do PSNR quadro-a-quadro	121
5.29	Visão geral do modelo de simulação B	122
5.30	Perfil Football ($q=25$, $\rho \approx 0.99$): (a) Sem Disputa, (b) Com Disputa .	125
5.31	Perfil Crew ($q=25$, $\rho \approx 0.99$): (a) Sem Disputa, (b) Com Disputa . .	125
5.32	Perfil Mobile ($q=25$, $\rho \approx 0.98$): (a) Sem Disputa, (b) Com Disputa .	126
5.33	Perfil Foreman ($q=25$, $\rho \approx 0.98$): (a) Sem Disputa, (b) Com Disputa	127

Lista de Tabelas

2.1	EVDO Rev.0: Modulação adaptativa em função da qualidade do canal	19
2.2	802.16b: Modulação adaptativa em função da qualidade do canal . . .	19
2.3	Taxas de <i>bit</i> , tamanhos de pacotes e duração dos <i>slots</i> do enlace reverso	25
2.4	Taxas de <i>bit</i> do enlace reverso é função da qualidade percebida pela ERB	26
3.1	Resumo dos algoritmos apresentados na Seção 3.1: Algoritmos de Escalonamento para redes sem fio de taxa variável	39
3.2	Resumo dos algoritmos apresentados na Seção 3.2 - Modelagem e Análise de desempenho do padrão EVDO	42
3.3	Resumo dos trabalhos apresentados na Seção 3.2.2: Escalonamento e descarte de pacotes com prioridade aplicados a fluxo de vídeo escalável	47
4.1	Parâmetros para o modelo de Okumura Hata, cenário: Urbano Denso	52
4.2	Valores para Arb em função da taxa no canal reverso	53
4.3	Valores para $\frac{E_c}{N_t}$ em função da taxa no canal reverso	53
4.4	Valor da SINR necessária para transmissão em cada uma das taxas possíveis no enlace direto para uma PER de 1%.	58
4.5	Probabilidade de transição para a taxa adjacente caso o controle de congestionamento não esteja (<i>p</i>) ou esteja (<i>q</i>) ativo	59
4.6	Parâmetros para o modelo de tráfego <i>web</i>	63
4.7	Valores configurados para os experimentos realizados	67
4.8	σ_g^2 , σ_d^2 , d_g e d_d quantificando a justiça em nível de usuário	80
5.1	Parâmetros considerados para solução do modelo analítico	94
5.2	Análítico vs. simulação <i>A</i> : resumo das diferenças	102

5.3	Parâmetros de Quantização (QP) usados em todos os perfis de vídeo .	108
5.4	Resumo dos perfis de vídeo	110
5.5	Parâmetros de Quantização (QP) usados e a taxa média resultante para cada perfil de vídeo	123

Glossário

- AC - *Access Category.*
- ADSL - *Asymmetric Digital Subscriber Line.*
- BPSK - *Binary Phase Shift Keying.*
- BSC - *Base Station Controller.*
- BTS - *Base Transceiver Station.*
- CBR - *Constant Bit Rate.*
- CDMA - *Code Division Multiple Access.*
- CQI - *Channel Quality Indicator.*
- CSMA - *Carrier Sense Multiple Access.*
- dB - *Decibel.*
- DRC - *Data Rate Control.*
- ERB - *Estação Rádio base.*
- EVDO - *Evolution Data Optimized (Only).*
- EVDV - *Evolution Data and Voice.*
- FEC - *Forward Error Correction.*
- FEDD - *Feasible Earliest Due Date.*
- FGS - *Fine Granular Scalability.*
- FIFO - *First in First Out.*
- FTP - *File Transfer Protocol.*
- Gbps - *Giga bits por segundo.*
- GOP - *Group of Pictures.*
- HDR - *High Data Rate.*
- HMM - *Hidden Markov Model.*

HSDPA - *High-Speed Downlink Packet Access.*

HTTP - *Hypertext Transfer Protocol.*

IP - *Internet Protocol.*

ISDN - *Integrated Services Digital Network.*

ITU - *International Telecommunication Union.*

Kbps - *Kilo bits por segundo.*

Mbps - *Mega bits por segundo.*

MGS - *Medium Grained Scalability.*

M-LWDF - *Modified Largest Weighted Delay First.*

M-PSK - *M-ary Phase Shift Keying.*

M-QAM - *M-ary Quadrature amplitude modulation.*

NS - *Network Simulation.*

PFS - *Proportional Fair Scheduling.*

PSNR - *Peak Signal-to-Noise Ratio.*

QAM - *Quadrature amplitude modulation.*

QoS - *Quality of service.*

QoE - *Quality of experience.*

QP - *Quantization Parameter.*

QPSK - *Quadrature Phase Shift Keying.*

QBD - *Quasi-Birth-Death.*

RF - *Rádio Freqüência.*

R-PEDD - *Proactive Earliest Due Date.*

RRI - *Reverse Rate Indicator.*

RTRA - *Real-Time Retry-Limit Adaptation.*

RTT - *Rount Trip Time.*

RWP - *Random Waypoint.*

SINR - *Signal to Interference and Noise Ratio.*

SNR - *Signal to Noise Ratio.*

SSI - *Signal Strength Indication .*

TA - *Terminal de acesso.*

TDMA - *Time Division Multiple Access.*

TgiP - *Transferência Gradual de Bits de Paridade .*

TTP - *Total-To-Pilot transmit power ratio.*

UDP - *User Datagram Protocol.*

UEPS - *Urgency and Efficiency-Based Packet Scheduling.*

v.a. - *Variável aleatória .*

WAP - *Wireless Application Protocol .*

WFQ-PF - *Weighted Fair Queuing-Proportional Fair.*

Wimax - *Worldwide Interoperability for Microwave Access .*

WLAN - *Wireless Local Area Network .*

Capítulo 1

Introdução

1.1 Visão Geral

Nas últimas duas décadas, temos observado um aumento na capacidade de comunicação em sistemas móveis. Um dos fatores chaves para tal evolução foi o desenvolvimento de tecnologias que possibilitaram a adoção da técnica de comutação de pacotes em sistemas onde a comutação de circuitos predominava, como as redes celulares, por exemplo.

Este fenômeno colaborou para considerável diminuição dos custos do acesso à Internet através de redes sem fio. A possibilidade de acesso móvel à Internet a taxas da ordem de Mbps e a preços acessíveis motivou forte migração dos usuários para as tecnologias sem fio. O mercado rapidamente observou a viabilidade econômica destas redes e passou-se a investir cada vez mais em pesquisas para o desenvolvimento tanto de melhoramentos quanto de novas tecnologias. Assim, esta tendência de crescimento tem se consolidado a passos cada vez maiores.

Todo este crescimento não seria possível sem a pesquisa básica na área de redes sem fio. Estas redes trazem inúmeros desafios devido ao conjunto severo de restrições impostas por sua camada física.

Além de a atenuação do sinal ser consideravelmente maior nas redes sem fio, existe uma série de problemas específicos destas redes e que não surgem nas redes cabeadas. A *difração*, o *espalhamento* e a *reflexão* são exemplos de fenômenos que exigem novas soluções. Para um usuário, estes fenômenos se manifestam através de variações da qualidade do canal que ocorrem em diferentes escalas de tempo (ex:

curto, longo), como representado na Figura 1.1.

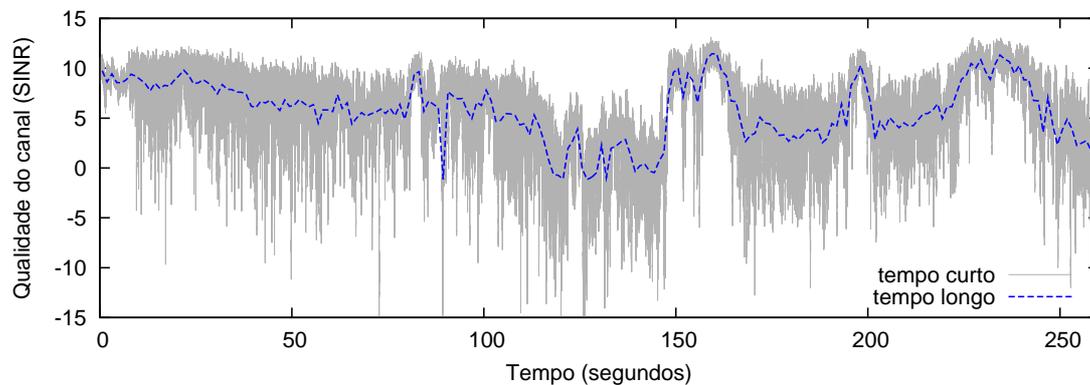


Figura 1.1: A qualidade do canal varia em diferentes escalas de tempo

A variação de qualidade ocorre pois o sinal encontra, por exemplo, obstáculos no caminho entre a fonte e o destino.

Em geral, o conjunto de soluções escolhidas para lidar com variações na qualidade do sinal em canais sem fio busca adaptar o transmissor às flutuações percebidas de modo a minimizar o efeito das interferências no sinal.

Um exemplo importante de adaptação é a modulação adaptativa usada para representar os *bits* no meio físico. Modulações capazes de transmitir dados a taxas mais altas são usadas quando a qualidade do canal está alta. Entretanto, se a qualidade do canal degrada, então a taxa de erros cresce. Nestes casos o sistema passa a usar modulações mais robustas à condições adversas. Estas modulações, todavia, transmitem dados a taxas bem inferiores.

Outro exemplo é o controle de potência que é usado para adaptar a potência de transmissão de modo a usar a menor energia tal que o sinal tenha força suficiente para alcançar o receptor com qualidade necessária para a decodificação com taxa de erro abaixo de um limiar estabelecido (ex: 1%). Isso é importante tanto para manter a interferência sob controle quanto para diminuir, à medida do possível, o consumo de bateria.

Quando a qualidade do canal cai abaixo de um certo limiar, qualquer esquema de modulação irá sofrer com os erros de transmissão independentemente da presença do controle de potência. A solução natural para este impasse é a utilização de múltiplos canais que tenham algum grau de independência quanto à qualidade do meio. Esta

solução é comumente chamada de diversidade.

Na camada física, a diversidade é normalmente obtida: sobre o tempo, quando os *bits* são espalhados no tempo; sobre a frequência, quando são usados diferentes canais de frequência; ou sobre o espaço, quando são usadas várias antenas suficientemente espaçadas. Assim, a diversidade permite que diferentes partes dos símbolos emitidos encontrem diferentes condições em seus respectivos canais, de modo a que haja maior probabilidade de se encontrar pelo menos um canal com qualidade suficiente para a transmissão.

A camada de enlace também procura se adaptar às variações na camada física. Esta adaptação ocorre, geralmente, na forma de algoritmos de escalonamento que procuram dar alguma prioridade a transmissões que encontram os canais em estados mais favoráveis. Técnicas como essa trazem um novo nível de diversidade: a multi-usuário.

Embora a diversidade aumente significativamente o desempenho de redes sem fio, ela não impede que um nó, ainda sim, possa sofrer de decaimento da qualidade do canal por diversos períodos de tempo. Quando isso acontece camadas superiores percebem efeitos colaterais.

Na camada de transporte, o TCP (*Transmission Control Protocol*) pode entender que há congestionamento no núcleo da rede enquanto, na verdade, as perdas são causadas por problemas de transmissão na rede sem fio que liga o usuário ao núcleo da rede. Diversos trabalhos na literatura [1, 2, 3] propõem variações do TCP que contornam este.

Uma aplicação que usa o TCP percebe queda brusca na vazão de sua sessão devido à adaptação do TCP (redução da janela) quando o protocolo interpreta a perda na rede sem fio como sendo um sinal de congestionamento.

Efeitos negativos também ocorrem quando, na camada de transporte, usa-se o serviço UDP (*User Datagram Protocol*) (ex: aplicações de áudio ou vídeo). Os efeitos do decaimento da qualidade do sinal na camada física, neste caso, se manifestarão na forma de problemas na qualidade de serviço, como a perda de pacotes e o esvaziamento do *buffer* no cliente. Conseqüentemente, o usuário é prejudicado ao notar ruídos no som, artefatos no vídeo e pausas forçadas do serviço para recarregamento do *playout buffer*. Este conjunto de problemas afeta negativamente a qualidade de

serviço (QoS - *Quality of Service*).

Muitos modelos de desempenho tem sido projetados para analisar os problemas abordados acima. O objetivo final dos estudos consiste, normalmente, em tentar manter a qualidade de serviço observada pelos usuários mesmo havendo flutuações nas características da camada física.

Modelos analíticos [4, 5, 6] são tipicamente usados para atacar questões específicas da camada física e da camada de enlace. Contudo, trabalhos recentes como, por exemplo, Wee Lum Tan *et al.*[7], indicam que é inviável o uso de modelos analíticos para representar as camadas de uma tecnologia sem fio com alto grau de detalhamento, devido à complexidade destas redes. Por esta razão, boa parte dos trabalhos que estudam o desempenho alcançado por sistemas sem fio, ou que analisam o efeito de flutuações em diferentes camadas sobre o desempenho de sistemas sem fio, tipicamente são realizados através de modelos de simulação.

1.2 Motivação e Objetivos da Tese

Entre os tipos de redes sem fio disponíveis, este trabalho está focado no grupo de padrões que chamamos de redes sem fio de longa distância, ou seja, redes de acesso à Internet que são evoluções ou possuem estrutura similar às redes de telefonia celular. Estas redes possuem alcance de sinal significativamente maior do que redes locais sem fio (ex: IEEE 802.11 [8]).

Nos últimos anos, o foco da pesquisa tem sido a capacidade de serviço destas redes. Entretanto, na maioria dos casos, os trabalhos envolvem unicamente a camada física. Quando as camadas físicas e superiores são estudadas conjuntamente, as superiores são muito simplificadas.

Para representar as características de um canal sem fio, é comum o uso de modelos estocásticos. Estes modelos assumem que diferentes comportamentos do canal ocorrem com diferentes probabilidades e escalas de tempo distintas. Diversos trabalhos [9, 10, 11] disponíveis na literatura realizaram medições em canais sem fio, e a partir dos resultados obtidos, foram propostos modelos estocásticos e seus respectivos parâmetros. Através do uso de modelos que reproduzem as características mais relevantes de um sistema, é possível, por exemplo, estudar o desempenho esperado e

analisar como mudanças em parâmetros ou no protocolo afetam seu funcionamento.

1.2.1 Um estudo do efeito das limitações da camada física sobre a QoS

Nesta parte de nosso trabalho, estudamos o efeito das limitações da camada física sob o ponto de vista dos usuários usando uma abordagem baseada no detalhamento de um número maior de características e de camadas do que é tradicionalmente encontrado na literatura.

Um dos objetivos deste trabalho é propor um modelo matemático que nos permita responder a seguinte questão: As adaptações da camada física e da camada de enlace empregadas em redes sem fio de longa distância são capazes de manter a qualidade de serviço e a justiça a um nível satisfatório?

Como metodologia para o desenvolvimentos de nossos modelos e de nossa discussão, optamos por estudar um dos padrões disponíveis (EVDO Rev.0). Embora nossos experimentos e observações sejam feitos com o foco neste padrão, o conjunto de conceitos apresentados no Capítulo 2 ajudará a entender que qualquer tecnologia atual de redes sem fio de longa distância sofre de limitações análogas.

É possível dividir esta parte do trabalho em seis estágios de desenvolvimento, que trazem contribuições distintas:

- primeiramente propomos um modelo de simulação que representa uma célula e que detalha características das camadas físicas e de enlace.
- em seguida implementamos e agregamos, a este sistema, modelos de usuários estáticos com o intuito de gerar carga suficiente para estressar o sistema.
- a terceira parte destina-se a resolver o modelo para determinar o número de usuários que uma célula pode servir mantendo uma qualidade mínima, do ponto de vista do usuário, que será determinada em função de um conjunto de parâmetros. As contribuições obtidas até este ponto foram publicadas no Simpósio Brasileiro de Redes de Computadores, em 2006 [12].
- a quarta parte consiste em investigar a justiça, ou seja, a diferença de qualidade de serviço entre usuários, em função de sua localização em relação ao

centro da célula. As contribuições trazidas por esta parte de nosso trabalho foram publicadas no *IEEE Personal Indoor Mobile Radio Communications Conference* (PIMRC), ainda em 2006 [13].

- na quinta parte, implementamos mobilidade [14] aos usuários do modelo proposto para avaliar seu efeito sobre a justiça. As contribuições trazidas por esta etapa foram publicadas no *IEEE Wireless Communications and Networking Conference* (WCNC), em 2008 [15].
- em seguida, na sexta parte, reunimos as limitações encontradas, do ponto de vista da qualidade de serviço e justiça, comparando nossa abordagem com a tradicionalmente encontrada na literatura e propomos uma solução simples para contornar tais problemas. Esta fase de nosso trabalho resultou em um artigo no periódico *IEEE Transactions on Broadcasting: Special Issue in Multimedia Broadcasting* [16], publicado em 2008.

1.2.2 Proposta de um novo método para transmissão de fluxos de vídeo escalável sobre redes sem fio

Serviços de fluxo de vídeo sobre redes sem fio estão se tornando cada vez mais populares. Isso ocorre graças tanto às pesquisas crescentes na áreas de redes sem fio, como destacado nas seções 1.1 e 1.2.1, quanto aos mais recentes padrões de codificação de vídeo como o MPEG-4[17] e sua versão escalável [18, 19], que permite degradação gradativa da qualidade obtida mediante à presença de condições desfavoráveis e variações temporais de enlaces sem fio.

Existem diversas questões de desempenho que devem ser enfrentadas para alcançar um nível razoável de qualidade de fluxos de vídeo sobre canais sem fio. De fato, durante a transmissão destes tipos de mídia, variações na taxa efetiva de transmissão não podem ser ignoradas, já que o canal pode permanecer por períodos de tempo em condições ruins, resultando em taxas de transmissão muito baixas.

Neste cenário, fluxos tradicionais (não escaláveis) de vídeo sofrem níveis severos de perda e, potencialmente, esvaziamento do *playout buffer* do cliente, resultando na degradação inaceitável na QoS. O uso de vídeo escalável emerge como uma alternativa para mitigar os efeitos das fortes variações na capacidade do canal.

Basicamente, o uso de vídeo escalável permite a seleção de uma ampla gama de opções de taxas de transmissão para um único fluxo de vídeo codificado [20, 21]. Este é um forte atrativo especialmente para cenários de transmissão sem fio, pois a aplicação pode adaptar a taxa de transmissão para a capacidade disponível naquele intervalo de tempo, no caminho entre a fonte e o destino.

Considere o cenário geral apresentado na Figura 1.2, onde um usuário consome um fluxo de vídeo emitido pelo servidor. O enlace sem fio é o último salto no caminho servidor cliente, e constitui o gargalo deste caminho. Por exemplo, a taxa r disponível no enlace sem fio é menor que a taxa (r') disponível entre o servidor e o roteador do enlace sem fio ($r' < r$). Três soluções têm sido propostas em estudos recentes disponíveis na literatura [22, 23, 24].

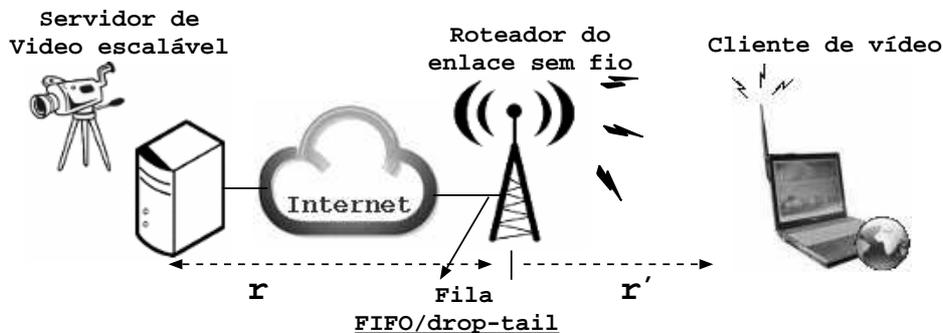


Figura 1.2: Cenário geral considerado nesta fase do trabalho

Uma das soluções consiste em alterar, em tempo de transmissão, a taxa de envio do vídeo, a partir do servidor, com o objetivo de adaptar à taxa disponível no enlace sem fio ao qual o cliente está conectado, visando melhorar a QoS. Uma desvantagem [24] desta abordagem é a baixa precisão da adaptação, que é resultado da combinação da variação rápida da qualidade do sinal no enlace sem fio (vide Figura 1.1), com o fato de que a informação sobre a qualidade do canal sempre chega defasada ao servidor, devido ao atraso da Internet. Assim, flutuações imprevisíveis (*fast-fading*) ou causadas pelo compartilhamento do enlace sem fio afetam consideravelmente a QoS.

A segunda solução diz respeito à política de escalonamento de pacotes, que tem como ideia básica a priorização [25, 26] do serviço, no enlace sem fio. Nestes métodos, pacotes mais importantes para a qualidade final do vídeo são servidos com maior prioridade, em detrimento dos pacotes menos importantes.

Já na terceira solução, é adotada uma política de descarte de pacotes da fila do último salto (o enlace sem fio) que tem como objetivo descartar pacotes de menor importância [27, 26] para a qualidade final do vídeo, enquanto pacotes mais importantes, são mantidos na fila.

As ideias trazidas pelos trabalhos pertencentes aos grupos da segunda e terceira soluções apresentadas acima têm como principal objetivo a diminuição do efeito da perda, ocorrida na fila do enlace sem fio, sobre a QoS. Estes métodos possuem, contudo, desvantagens como a necessidade intrínseca de alteração em padrões e em equipamentos, além da alta complexidade computacional, já que é inevitável processar o conteúdo de pacotes para analisar camadas superiores e, então, determinar seu nível de importância.

Na parte deste trabalho desenvolvida no Capítulo 5, o principal objetivo é propor um método de transmissão e avaliar sua viabilidade para mitigar o efeito de flutuações, ocorridas no último canal (sem fio), sobre a QoS.

Entretanto, diferentemente dos trabalhos disponíveis na literatura, buscamos este objetivo através de um método que:

- não requeira alterações em padrões e em equipamentos de redes sem fio e
- seja computacionalmente eficiente.

Nosso método é simples, porém robusto, e tira proveito tanto das propriedades de vídeos escaláveis quanto do processo randômico das perdas observado na fila do roteador do enlace sem fio.

Através do ajuste do intervalo entre transmissões dos pacotes, no servidor de vídeo, e do uso de uma determinada ordenação dos pacotes emitidos, mostramos que o padrão de perda observado pode ser usado para obter uma QoS superior, sem que haja necessidade de alterar a política de escalonamento e/ou de descarte de pacotes no roteador do enlace sem fio. Em outras palavras, mostramos que a QoS pode ser alterada através da variação do espaçamento durante a transmissão dos pacotes, isto é, variando-se o nível de rajada intra-quadro, varia-se a QoS. Observa-se, ainda, que o espaçamento normalmente empregado não é o que resulta na melhor QoS.

É bem conhecido que o processo de perda de pacotes é sensível ao nível de rajada da fonte de tráfego e que a taxa de perda aumenta com o nível de rajada. Na Seção 3.2, apresentaremos trabalhos relacionados que realizaram estudos desta característica. O que é contra-intuitivo, e ainda não foi abordado por trabalhos relacionados, é o fato de que o uso de níveis leves de rajada combinado com uma ordenação correta de pacotes de vídeos escaláveis possa favorecer a QoS.

Do nosso conhecimento este é o primeiro trabalho que avalia o processo de perda de pacotes individuais em uma rajada, de acordo com sua ordenação. Estudamos o nível de perda observado para o primeiro pacote das rajadas, para o segundo pacote das rajadas, e assim em diante. Denominamos este estudo de análise do processo de perda *intra-rajada*. Esta ideia se baseia no fato de que fluxos escaláveis de vídeo contém pacotes com diferentes níveis de importância para a QoS. O uso do que observamos acima com relação ao processo de perda para trazer melhorias na QoS percebida pelos usuários quando vídeos escaláveis são usados na transmissão também é, de acordo com nossos estudos, uma ideia original. Nossas contribuições, nesta parte da tese, são:

- um modelo analítico que obtém a fração de perda de pacotes individuais de um fluxo de vídeo escalável sobre uma fila FIFO (*First In First Out*), onde os pacotes são transmitidos em rajadas de n pacotes. Os resultados fundamentais obtidos pelo estudo realizado através deste modelo trouxeram contribuições publicadas no SBRC 2010 [28];
- validação das principais observações obtidas do modelo analítico. Para isso propomos e implementamos um modelo de simulação que utiliza cenas de vídeos reais. Uma versão estendida do artigo [28], contendo resultados e contribuições preliminares do referido modelo de simulação foi aceito para publicação no IEEE Globecom 2010 [29];
- um estudo, feito através de um arcabouço desenvolvido durante este trabalho, sobre relação entre o processo de perda observado, e a qualidade objetiva (PSNR - Peak Signal-to-Noise Ratio [30]) do vídeo recebido pelo cliente. Com isso, somos capazes de quantificar o ganho alcançado, de acordo com os parâmetros considerados;

- inclusão de novas funcionalidades ao modelo de simulação do protocolo estudado, com objetivo de validar o comportamento observado para um cenário realista, onde o serviço é realizado de acordo com *traces* reais de uma rede sem fio (EVDO) de acesso à Internet. Enfatizamos que o modelo EVDO é bastante detalhado, e as características do meio físico são incluídas no modelo. As contribuições obtidas até então estão sendo organizadas para serem submetidas a um periódico internacional.

1.3 Organização desta tese

A seguir apresentamos brevemente a maneira como esta tese está organizada.

No Capítulo 2 apresentamos o conjunto principal de conceitos que será a base para o entendimento deste trabalho. Dentre estes conceitos destacam-se a apresentação mais detalhada de fenômenos da camada física como a atenuação, o *slow-fading* e o *fast-fading*; uma breve apresentação das principais técnicas empregadas nas camadas física e de enlace para contornar tais limitações; uma introdução sobre cadeias de Markov e uma visão geral sobre codificação escalável de vídeo.

Um apanhado geral das pesquisas relacionadas à qualidade de serviço e a justiça alcançada pela tecnologia de redes de computadores sem fio à qual escolhemos para basear nosso trabalho (EVDO Rev.0) é apresentado no Capítulo 3. Neste mesmo capítulo, tratamos de trabalhos relacionados ao nosso novo método de transmissão de vídeo escalável, cujo principal objetivo é minimizar o efeito das perdas sobre a QoS. Note que optamos por organizar o Capítulo 3 por subáreas de pesquisa de modo a que possamos situar melhor as diferentes contribuições desta tese em relação aos trabalhos relacionados.

No Capítulo 4 descrevemos o modelo de simulação para uma célula EVDO, detalhando quais soluções foram escolhidas para cada característica reproduzida. Em seguida definimos os cenários de experimentos e apresentamos e discutimos os resultados obtidos.

Em seguida, no Capítulo 5, apresentamos nossa proposta de um novo método para transmissão de fluxo de vídeo escalável. Para isso, apresentamos o modelo analítico e os modelos de simulação usados para fundamentar e avaliar o método em

questão. O Capítulo 5 é finalizado com a descrição, para cada modelo, dos cenários de experimentos considerados e com a exposição e discussão dos resultados obtidos.

Por fim, no Capítulo 6, apresentamos as considerações finais sobre esta tese e as propostas de alguns temas para trabalhos futuros.

Capítulo 2

Fundamentos teóricos

Este capítulo apresenta, de forma sucinta, os principais fundamentos teóricos que foram importantes para o desenvolvimento desta tese. A estrutura do capítulo é indicada a seguir.

Na Seção 2.1 apresentamos as principais limitações de redes sem fio, bem como as técnicas mais conhecidas para contornar estas limitações. Isso é importante pois as contribuições trazidas por esta tese consideram cenários de redes sem fio.

A seção 2.2 é destinada a apresentar um dos principais padrões de redes sem fio de longas distâncias para acesso à Internet: a tecnologia EVDO rev.0. Estes conceitos são úteis, pois na primeira fase de nosso trabalho, apresentada no Capítulo 4, construímos um modelo de simulação de para uma célula EVDO rev.0 para realizar nossos estudos.

Na segunda parte desta tese propomos um modelo markoviano para fundamentar um o estudo da técnica de transmissão de fluxos de vídeo escalável que propomos. Portanto, apresentamos o conceito de processos markovianos na Seção 2.3. Além disso, na Seção 2.4 apresentamos uma breve visão sobre codificação de vídeo escalável.

2.1 O que diferencia as redes sem fio das redes cabeadas tradicionais?

2.1.1 Introdução

Esta seção traz uma visão geral das principais limitações/impedimentos [31, 32] à transmissão de dados da camada física de redes sem fio, bem como dos principais métodos usados para contornar tais características. Isto nos dará o conhecimento necessário para a compreensão tanto dos modelos matemáticos construídos em nosso trabalho quanto das discussões e análises realizadas durante nossos experimentos.

É possível classificar os impedimentos inerentes à camada física de redes sem fio em duas classes [33]. Primeiramente temos os fenômenos de escala de tempo longa, em segundo lugar, temos os fenômenos de escala de curta escala de tempo.

Antes de entender melhor cada uma das duas classes de impedimentos da camada física, é importante compreender como a qualidade do canal observada pelo receptor varia no tempo.

Para isso, suponha um receptor que se movimenta em linha reta e velocidade constante enquanto se afasta do transmissor. A Figura 2.1 ilustra a variação da potência de um sinal observado por este receptor. À medida que o receptor se afasta do transmissor (aumento da distância), a potência do sinal recebido diminui devido à propagação do sinal no meio. Na Figura 2.1, a linha **I** representa a perda por propagação.

Durante a movimentação do receptor, é possível que grandes obstáculos se interponham entre ele e o transmissor. Isso faz com que o sinal percebido pelo receptor apresente alterações de duração relativamente longas, como representado pela curva **II** na Figura 2.1. Esta curva representa o desvanecimento lento (tempo longo) que será abordado na Seção 2.1.2.

Outro grupo de fenômenos afeta o sinal recebido em escalas mais curtas, gerando o desvanecimento rápido (de tempo curto), como indicado pela curva **III** da Figura 2.1. Este tipo de variação da qualidade do sinal será abordado na Seção 2.1.3.

Por fim, na seção 2.1.4, discutimos sucintamente a adaptatividade, nas camadas físicas e de enlace, de redes sem fio em relação a estes fenômenos. É justamente esta necessidade maior de adaptação que diferencia as redes sem fio das redes cabeadas

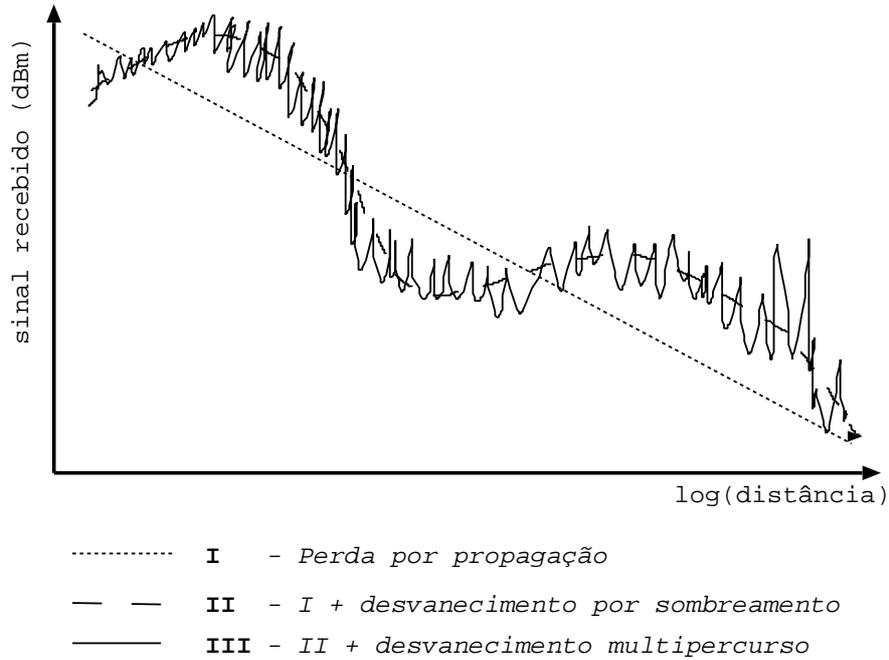


Figura 2.1: O desvanecimento rápido flutua em torno de uma média local dada pelo desvanecimento lento

convencionais.

2.1.2 Camada Física: impedimentos e o desvanecimento lento

A atenuação do sinal pela distância, também chamada de perda por propagação (do inglês *path-loss*) [31, 32, 33] é a redução da potência de um sinal eletromagnético que viaja entre um transmissor e um receptor à medida que se propaga pelo espaço que os separa.

De maneira geral, a perda do sinal devido à distância que separa dois pontos é descrita pela relação

$$P_r = \frac{P_t}{d^n}, \quad (2.1)$$

onde P_r é a potência do sinal recebido, P_t é a potência do sinal transmitido, d é a distância que separa o transmissor do receptor e n é um valor ajustado de acordo com o ambiente de propagação do sinal.

Diversos trabalhos disponíveis na literatura [34, 35, 10, 11, 9, 36] sugerem valores para n de acordo com o ambiente considerado, bem como propõem a adição, à

equação 2.1, de componentes para melhorar a precisão da modelagem da atenuação pela distância.

Há fenômenos que afetam a propagação do sinal em sistemas de comunicação. Entre eles os principais são: a penetração, a reflexão, a difração e o espalhamento. Estes fenômenos, ilustrados na Figura 2.2, que ocorrem quando o sinal encontra obstáculos físicos, alteram o padrão da propagação do sinal. A ocorrência destes fenômenos depende da relação entre o comprimento da onda do sinal em questão (λ) e as dimensões e densidade dos obstáculos a serem ultrapassados.

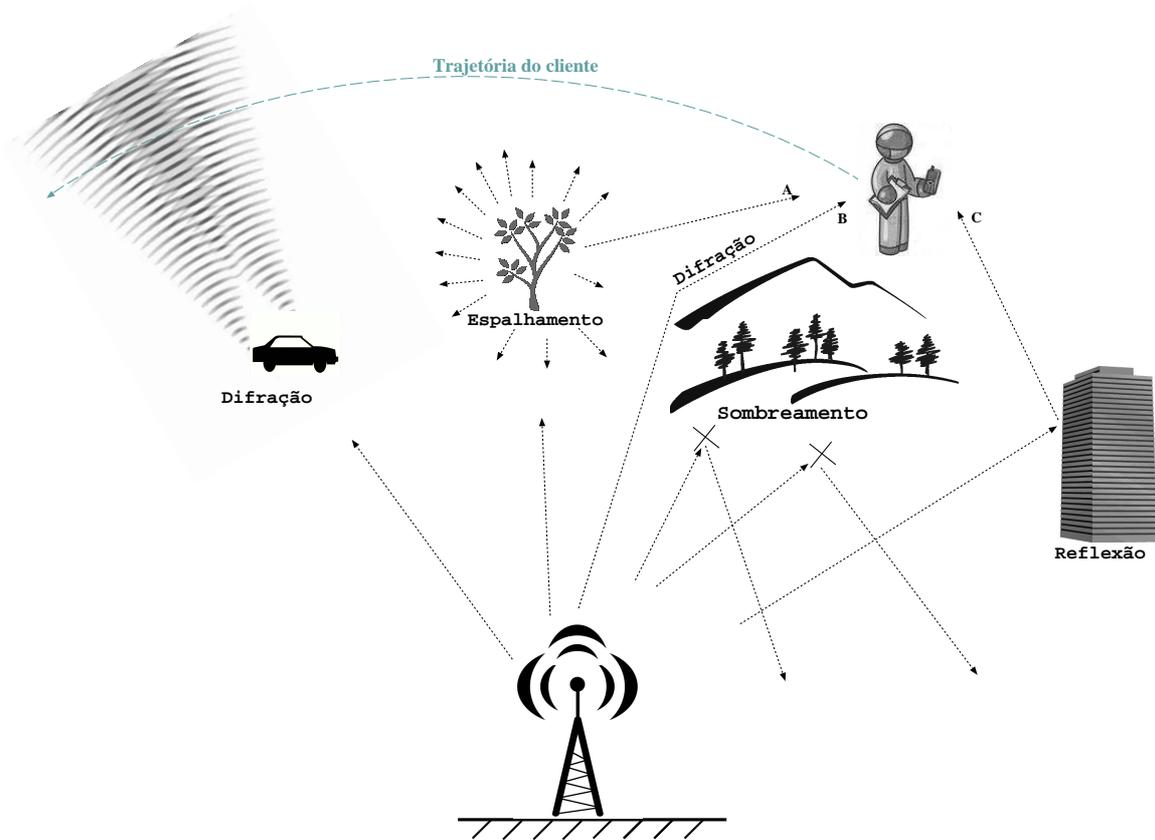


Figura 2.2: Reflexão, espalhamento e difração

A reflexão ocorre quando é encontrado um obstáculo de dimensões bem maiores que λ , fazendo com que um receptor que se localize atrás do obstáculo fique coberto por uma *sombra* e não consiga perceber o sinal. Isso é chamado de sombreamento (do inglês *shadowing*). Na Figura 2.2, o usuário localizado atrás do conjunto de montanhas enfrenta uma situação de sombreamento.

Como mostrado na Figura 2.1, o sombreamento gera alterações na força do sinal percebido pelo receptor. Estas alterações aparecem como flutuações lentas na atenuação indicada pela curva I, gerando o comportamento indicado pela curva II. Geralmente estas flutuações são modeladas [37, 9, 33] através de uma distribuição log-normal de média 0, representando uma perturbação em volta do valor indicado pela Equação 2.1.

2.1.3 Camada Física: desvanecimento rápido

Quando as dimensões do obstáculo são apenas ligeiramente maiores que λ , então ocorre a difração (do inglês *diffraction*) nas frentes de onda que ultrapassam as bordas do obstáculo, fazendo com que um receptor consiga perceber o sinal mesmo estando fora da linha de visão do emissor, como indicado pelo sinal atrás do veículo na Figura 2.2. Neste caso, a potência do sinal percebido diminui rapidamente à medida que o receptor se aproxima do obstáculo.

Já o espalhamento ocorre, por exemplo, quando as ondas de rádio incidem sobre obstáculos de dimensão menor ou igual a λ . Como ilustrado pelo sinal que incide sobre árvore da Figura 2.2, o resultado deste fenômeno é que o sinal é espalhado em todas as direções.

Mesmo quando há um único transmissor, a combinação de fenômenos como a reflexão, o espalhamento e a difração resultam na chegada de múltiplos sinais ao receptor, o que é chamado de desvanecimento multipercurso. Na Figura 2.2, o desvanecimento multipercurso é representado pela soma dos componentes A, B e C que incidem sobre o cliente.

Neste exemplo, uma frente do sinal refletiu (C) em um edifício, outra frente sofreu difração (B) e a terceira frente se dispersou ao incidir sobre a árvore (A). O efeito combinado dos vários sinais chega ao usuário.

Como os obstáculos estão em posições suficientemente distintas, a frente do sinal que passou por um dos obstáculos chega adiantada ou atrasada em relação às frentes que passaram por outros. Assim, o desvanecimento multipercurso pode ser definido como uma decomposição do sinal original após passar por obstáculos e sofrer efeitos como a difração, a reflexão e o espalhamento, resultando na chegada de diversas frentes de onda, deslocadas no tempo, ao receptor.

A resultante de um canal em desvanecimento multipercurso é prejudicial, pois parte da potência utilizada na emissão do sinal no transmissor atua como uma componente do ruído ou da interferência observado no receptor, devido à defasagem entre os sinais. Com isso, à medida que o usuário se move através de pequenas distâncias a potência do sinal recebido varia bem mais rapidamente, quando comparada com as flutuações geradas pelo desvanecimento lento. Assim, o desvanecimento multipercurso é classificado como desvanecimento rápido.

Quando não há linha de visão entre o emissor e o receptor, o sinal recebido resulta exclusivamente do desvanecimento multipercurso. Quando isso ocorre, o desvanecimento multipercurso também é chamado de *Rayleigh fading*, pois o sinal resultante pode ser descrito por uma variável aleatória com distribuição Rayleigh [37].

Por outro lado, quando há linha de visão entre os dois pontos, o sinal recebido diretamente do emissor torna-se um componente dominante do sinal total recebido. Neste caso, a variação da potência do sinal recebido é chamado de *Rician Fading*, pois pode ser descrito como uma variável aleatória com distribuição Rician [37].

2.1.4 Técnicas de adaptação aos impedimentos do meio físico

Modulação adaptativa

De maneira geral, a qualidade de um canal de comunicação pode ser escrita como uma relação entre a potência do sinal recebido (P_r) e o ruído ($P_{ruído}$) observado, como mostrado na Equação 2.2.

$$SNR = \frac{P_r}{P_{ruído}} \quad (2.2)$$

Além disso, segundo o teorema de Shannon e Hartley [33, 31, 32], o limite de uma comunicação confiável, em termos de taxa de erro de bits, de um canal depende da banda e do SNR:

$$r < b \log_2 \left(1 + \frac{P_r}{P_{ruído}} \right), \quad (2.3)$$

onde r é a taxa em bits por segundo e b é a largura de banda do canal (Hz).

Como descrito na Equação 2.1, P_r é inversamente proporcional à distância que separa dois pontos. Com isso, é trivial compreender que a Equação 2.3 pode ser reescrita em função da distância. Como conclusão, quanto maior for a distância entre dois pontos, menor será a taxa de bits alcançável de modo a que a taxa de erros de bits fique sob controle.

A representação de *bits* no meio físico através de sinais eletromagnéticos, chamada modulação digital, consegue alcançar taxas comparáveis aos limites ditados pelo teorema de Shannon e Hartley. A modulação digital possibilita a estimação da probabilidade esperada de erros de acordo com a modulação utilizada [38].

Tomemos como exemplos duas modulações amplamente utilizadas em redes sem fio, o PSK (Phase Shift Keying) e o QAM (Quadrature Amplitude Modulation).

No caso da modulação M-PSK, onde representa-se M bits por símbolo da camada física, a probabilidade de erro de *bit* pode ser aproximadamente escrita como [38]

$$P_{e,MPSK} \cong \frac{2}{\max(\log_2 M, 2)} \sum_{i=1}^{\max(M, 4.1)} Q \left(\sqrt{\frac{E_b}{N_0}} 2 \log_2 M \operatorname{sen} \frac{(2i-1)\pi}{M} \right) \quad (2.4)$$

No caso da modulação M-QAM, onde também representa-se M bits por símbolo da camada física, esta probabilidade pode ser aproximadamente escrita como [38]

$$P_{e,MQAM} \cong \frac{4}{\log_2 M} \left(1 - \frac{1}{\sqrt{M}} \right) \sum_{i=1}^{\sqrt{M}/2} Q \left((2i-1) \sqrt{\frac{E_b}{N_0} \frac{3E_b \log_2 M}{(M-1)}} \right) \quad (2.5)$$

onde E_b/N_0 é o SNR normalizado por *bit* e

$$Q(x) = \left(\frac{1}{\sqrt{2}} \right) \int_x^\infty e^{-t^2/2} dt$$

Logo, como pode ser visto nas equações 2.4 e 2.5, quanto maior for o número de bits por símbolo, maior será a probabilidade de erro de *bit*.

De fato, diversos padrões de redes sem fio de longa distância como o EVDO, o WIMAX [39], o HSPDA [40] e o WCDMA [41] usam esquemas adaptativos de modulação [42]. De maneira geral, estes padrões definem o BER (*Bit Error Rate*) alvo (ex: 1%), e calculam qual o SNR necessário para que seja escolhida a modulação de maior taxa, entre um grupo pré-determinado, de modo a que o BER seja satisfeito.

Taxa de bits (kb/s)	Pacotes comprimento (<i>slots</i>)	FEC Taxa	Modulação Esquema	Mínimo SINR (dB)
38.4	16	1/4	QPSK	-11.5
76.8	8	1/4	QPSK	-9.2
102.6	6	1/4	QPSK	-6.5
153.6	4	1/4	QPSK	-3.5
204.8	3	1/4	QPSK	-3.5
307.2	2	1/4	QPSK	-0.6
614.4	1	1/4	QPSK	-0.5
921.6	1	3/8	QPSK	2.2
1228.8	1	1/2	QPSK	3.9
1843.2	1	1/2	8-PSK	8
2457.6	1	1/2	16-QAM	10.3

Tabela 2.1: EVDO Rev.0: Modulação adaptativa em função da qualidade do canal

Para o caso do EVDO Rev.0 as opções de modulação vão desde PSK com $M = 4$ (QPSK) a 38.4kbps até QAM com $M = 16$ (16-QAM) a 2457.6kbps, como mostrado na Tabela 2.1.

Outro exemplo é o WIMAX (802.16d), que usa modulações que vão desde PSK com $M = 2$ (BPSK) a 1.89Mbps até QAM com $M = 64$ (64QAM) a 18.36Mbps, como mostrado na Tabela 2.2

Taxa de pico em 5MHz(Mb/s)	FEC Taxa	Modulação Esquema	Mínimo SNR (dB)
1.89	1/2	BPSK	6.4
3.95	1/2	QPSK	9.4
6.00	3/4	QPSK	11.2
8.06	1/2	16QAM	16.4
12.18	3/4	16QAM	18.2
16.30	2/3	64QAM	22.7
18.36	3/4	64QAM	24.4

Tabela 2.2: 802.16b: Modulação adaptativa em função da qualidade do canal

De maneira análoga, o HSDPA *release* 5 adota modulações desde QPSK a taxas de 1.2Mbps até 16QAM a 14.4Mbps.

Diversidade na camada física

Além da modulação adaptativa, outra técnica amplamente usada para contornar os impedimentos da camada física em rede sem fio é a diversidade.

Em geral, o objetivo da diversidade é diminuir a probabilidade de erro de um sinal transmitido. Para isso, usa-se diversidade no tempo, na frequência ou no espaço.

Na diversidade de tempo múltiplas versões do mesmo sinal são transmitidos em diferentes instantes de tempo. Estas versões podem ser cópias de um mesmo sinal, como também pode-se usar uma combinação de diversidade do tempo e alguma técnicas de *interleaving* e FEC (*Forward Error Correction*).

As diferentes versões do sinal devem ser suficientemente espaçadas no tempo, para que haja alta probabilidade de encontrarem diferentes condições do canal. Este intervalo de tempo (τ) é calculado em função do comprimento de onda (λ) usado na transmissão, e quanto maior o comprimento de onda, maior deve ser τ .

A diversidade no tempo pode ser desinteressante quando τ é grande, pois espera-se mais para transmitir as versões do sinal, o que pode resultar em atrasos. Para estes casos outra opção, como a diversidade de frequência pode ser usada.

A diversidade de frequência é interessante quando o desvanecimento afeta apenas parte das componentes de frequência de um sinal (lembre-se, os fenômenos de camada física dependem de λ). O sinal pode ser transmitido através de vários canais de frequência, ao invés de apenas um, ou pode ser espalhado por uma largura espectral mais ampla.

Outra técnica amplamente usada é a diversidade espacial. Neste caso aplica-se o uso de múltiplas antenas suficientemente espaçadas ($\lambda/2$) de modo a que o desvanecimento multipercurso seja percebido de maneira diferente por cada antena. O receptor dinamicamente seleciona qual antena usar para captação do sinal, escolhendo a que ofereça melhor relação sinal/ruído, de modo a minimizar a taxa de erros.

Diversidade na camada de enlace

Até agora comentamos sobre as técnicas (diversidade) usadas por cada terminal de acesso para aumentar a probabilidade de que o sinal transmitido seja recebido corretamente, bem como sobre o uso de modulação adaptativa para alterar a taxa

de transmissão de acordo com as condições do canal.

Mesmo com a aplicação de todas as técnicas descritas até agora, um terminal de acesso pode encontrar intervalos de tempo com grande variabilidade de qualidade do sinal (alta ou baixa). Considere, agora, o cenário onde vários terminais de acesso estão presentes. Os diferentes terminais perceberão diferentes condições de qualidade do canal em instantes de tempo distintos. Deste modo, quanto maior for o número de nós presentes, maior será a probabilidade de encontrar, em um dado instante de tempo, um nó que esteja percebendo um pico de qualidade em seu canal. Este efeito tem sido comumente chamado de diversidade multi-usuários[43, 44].

A diversidade multi-usuários pode ser explorada por algoritmos de escalonamento de pacotes para, por exemplo, aumentar a vazão agregada do sistema através de atribuição de maior prioridade aos usuários que momentaneamente possuem melhores condições no meio físico.

De fato, a partir da ideia trazida por Knopp e Humblet [43], diversas propostas de algoritmos de escalonamento de pacotes, na camada de enlace, passaram a usar a informação do estado na camada física (ex: SNR) como métrica para decisão de escolha do nó a ser servido. Alguns autores passaram a chamar essa técnica de escalonamento *cross-layer* [45, 46, 47, 48].

2.2 Visão geral do padrão EVDO rev.0

A tecnologia celular EVDO faz parte da família de técnicas CDMA2000 1x otimizada para serviços de dados por pacote. O enlace direto (ou *downlink*) desta técnica, também conhecida por padrão IS-856 [49], permite uma taxa de transmissão de pico de 2.4576 Mbps.

O enlace direto (na direção ERB - Estação Rádio Base para o TA - Terminal de Acesso, como ilustrado na Figura 2.3 (a)) emprega o esquema de Multiplexação por Divisão no Tempo (TDM) e, portanto, em cada janela de tempo (*slot*), com duração igual a 1.67 ms, é feita a transmissão para um e apenas um usuário, sempre utilizando a potência máxima da ERB.

Já o sinal piloto permite uma demodulação coerente no enlace direto e também é usado em outras funções como, por exemplo, na estimativa da qualidade do canal.

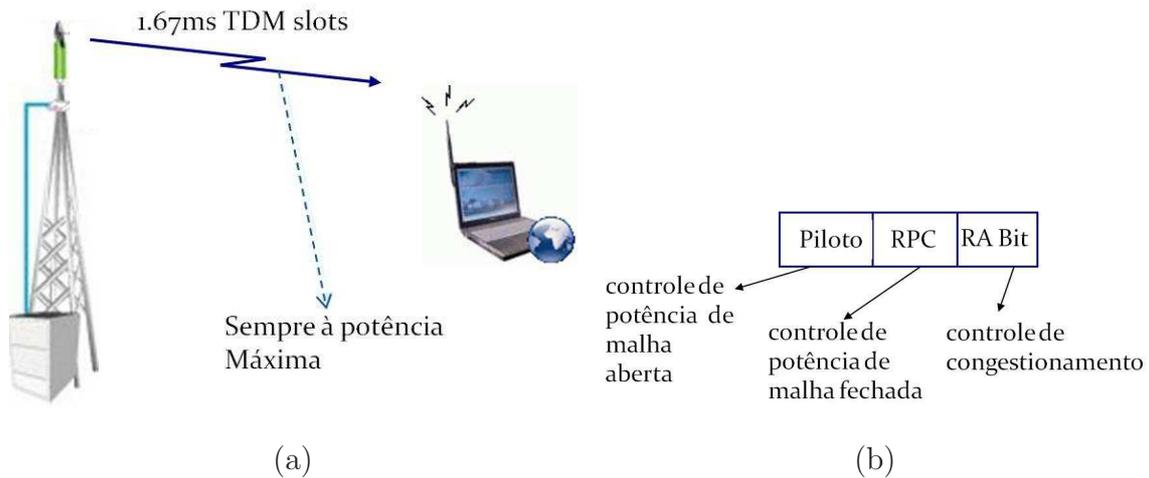


Figura 2.3: (a) Canal direto; (b) Estrutura do canal direto

O sinal no enlace direto já multiplexado é espalhado pelo método de sequência direta (DS), com uma taxa de *chips* de 1.2288 Mbps, ocupando uma faixa de frequência com largura de 1.25MHz, compatível, portanto, com os sistemas IS-95 e CDMA2000 1x.

Essa transmissão pode ser feita em diferentes taxas, e a escolha da taxa específica, a cada instante, depende da razão entre Portadora/Interferência (C/I) percebida pelo TA, que passa essa informação para a ERB (ou ponto de acesso) através de um campo de bits (DRC - Data Rate Control) no enlace reverso (Figura 2.4 (b)).

Além dos dados a serem transmitidos, o sinal no enlace reverso inclui um sinal piloto e um sinal de controle de acesso ao meio (MAC), sendo todos esses três sinais multiplexados no tempo. Por sua vez o sinal da camada MAC é composto de dois sinais multiplexados por divisão de código quais sejam o sinal de controle de potência (RPC) e o *bit* de controle de atividade do canal reverso (RA). Esta estrutura é ilustrada na Figura 2.3 (b).

Esse último *bit* é empregado pela ERB para sinalizar aos terminais na sua área de cobertura do estado de congestionamento do enlace reverso. A forma como é feito esse controle de congestionamento será abordada na seção 3.3.4. O canal lógico RPC serve para implementar o controle de potência a um sistema empregando a técnica de múltiplo acesso por divisão de código (CDMA), sendo que os bits empregados para esse fim são enviados a uma frequência de 600 Hz.

Como a grande maioria das técnicas modernas de comunicação sem fio, o EVDO

também faz uso de um esquema adaptativo de modulação/codificação, tanto no enlace direto quanto no enlace reverso. A Tabela 2.1 ilustra as taxas disponíveis no enlace direto, com os respectivos tamanhos dos pacotes, e esquemas de codificação e modulação empregados.

Pode-se observar que para as taxas mais baixas, o tempo para transmissão no enlace direto de uma determinada quantidade de informação ocupa um maior número de *slots* do que quando a taxa é alta, ou seja, ocupa o enlace por um tempo maior. A escolha da taxa de transmissão para um determinado usuário em certo instante depende da razão C/I , ou, equivalentemente, da razão E_b/N_t . A partir da Tabela 2.1 percebe-se que quanto maior a taxa utilizada, menos robusto é o esquema de modulação/codificação empregado.

Portanto as maiores taxas de transmissão só poderão ser usadas em condições muito favoráveis de propagação ou com o sistema com pouco tráfego. A Tabela 2.1 exemplifica os valores da razão (SINR) entre potência da portadora e potência interferente total, incluindo ruído térmico, necessários para cada uma das taxas de transmissão para que seja garantida uma probabilidade de erro de pacote (PER) de 1% no enlace direto.

O parâmetro DRC, como já mencionado, deve ser enviado pelo TA para a ERB, através do enlace reverso, informando a situação do TA naquele momento em termos de SINR indicando, portanto, qual a taxa máxima que ele pode receber informação respeitando o critério de qualidade.

Nota-se que o esquema de funcionamento do enlace direto do EVDO difere daquele dos sistemas TDMA tradicionais. Nestes últimos para um terminal ativo existe uma alocação fixa de um ou mais *slots* dentro de cada quadro. Já no EVDO é empregado um algoritmo de escalonamento para determinar para qual terminal é feita a transmissão em cada janela de tempo.

O algoritmo de escalonamento a ser empregado não faz parte do padrão. No entanto é frequente encontrar propostas de uso de algoritmos fundamentados no princípio de PFS (Proportional Fair Scheduling) [50].

De acordo com esta técnica a transmissão deve ser sempre feita para o TA que apresenta melhores condições momentâneas de SINR, ou seja, maior DRC, respeitada uma certa condição de justiça. Caso não houvesse essa condição de justiça este

critério estaria maximizando a vazão total da célula ou setor.

A definição de justiça neste caso é que qualquer um dos usuários não fique muito tempo sem transmitir fazendo com que a sua taxa média de recepção se reduza a um valor muito baixo e bastante distinto daquele obtido por usuários em condições mais favoráveis. Esse ajuste de justiça não é simples, pois como a transmissão em taxas mais baixas ocupa vários *slots*, a vazão agregada do sistema cai substancialmente se frequentemente for dada prioridade a um usuário em situação de distância ou propagação desfavorável.

O enlace reverso (*uplink*) do EVDO permite também o uso da técnica de transmissão gradual dos bits de paridade (TgiP), que permite o uso de uma estratégia chamada de *Early Termination*. Essa estratégia permite que o pacote seja transmitido com o mínimo de redundância que as condições momentâneas do canal e de interferência permitem.

Existe um intervalo de tempo entre a medição do valor de DRC e quando o pacote é transmitido com essa taxa pela ERB. Com isso, pode haver diferença entre a qualidade estimada e a qualidade do canal no momento da transmissão pelo canal direto. Quando o TA está em movimento a variabilidade da qualidade do canal é maior, e erros na estimativa do canal podem reduzir a eficiência do sistema. O método de *Early Termination* busca corrigir esse problema.

O funcionamento do *Early Termination* se dá da seguinte maneira: o TA envia seus *bits* de dados e segue transmitindo *bits* de paridade/redundância até que receba uma confirmação da ERB.

O enlace reverso, como ilustrado na Figura 2.4 (a), opera no método de acesso CDMA, já bastante conhecido, mas, diferente do CDMA tradicional, permite o uso de diferentes taxas de transmissão. A taxa real de transmissão depende do carregamento no setor/célula determinado pela ERB e também das condições de propagação/distância de cada terminal de acesso.

O ajuste das taxas, em função do carregamento, é feito de forma probabilística, ou seja, ao comando da ERB cada terminal sorteia se ajustará sua taxa no canal reverso (aumento ou decréscimo) de acordo com uma probabilidade que varia de com a taxa que está sendo usada naquele momento.

Para que o terminal possa usufruir de uma taxa maior, é necessário que ele trans-

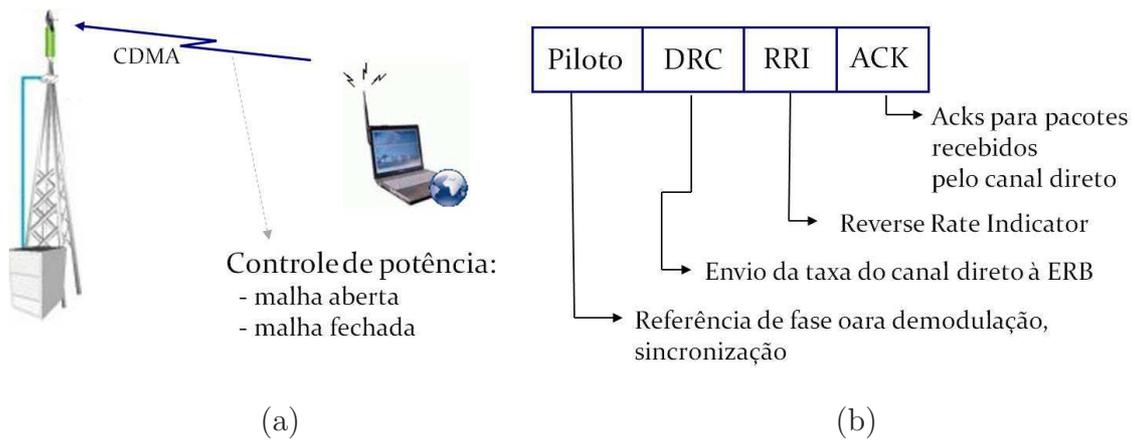


Figura 2.4: (a) Canal reverso; (b) Estrutura do canal reverso

mita a uma potência maior. Portanto se ele estiver muito distante da ERB ele, em geral, ficará restrito a taxas de transmissão mais baixas do que os usuários localizados próximos à ERB, embora o valor exato de cada taxa dependa do carregamento instantâneo da célula/setor.

Taxa de bits (kbps)	Tamanho do pacote (bytes)	Duração do pacote (ms)	Tamanho do pacote (slots)	Modulação
9.6	512	53.33	32	BPSK
19.2	1024	53.33	32	BPSK
38.4	2048	53.33	32	BPSK
76.8	4096	53.33	32	BPSK
153.6	8192	53.33	32	BPSK

Tabela 2.3: Taxas de *bit*, tamanhos de pacotes e duração dos *slots* do enlace reverso

A Tabela 2.3 ilustra as taxas de modulação disponíveis no enlace reverso com as respectivas durações de cada pacote. Note que todos os esquemas usam a mesma modulação e os pacotes têm a mesma duração, independente da taxa escolhida. Tabela 2.3.

Na Tabela 2.4 são fornecidos os ganhos de potência com relação ao canal piloto com que devem ser transmitidos os sinais em cada uma das taxas. Assim como no enlace direto, no enlace reverso terminais mais distantes da ERB também terão mais dificuldades de transmitir usando taxas mais altas devido à limitação de potência do terminal.

Como pode ser visto na Figura 2.4 (b), além do canal de tráfego de dados, o

Taxa de bits (kbps)	Qualidade necessária (SINR em dB)
9.6	3.75
19.2	6.75
38.4	9.75
76.8	13.25
153.6	18.5

Tabela 2.4: Taxas de *bit* do enlace reverso é função da qualidade percebida pela ERB

enlace reverso também inclui um canal piloto, um canal de confirmação (ACK) e o canal MAC. O canal MAC é subdividido em dois outros canais lógicos, a saber: o DRC, cuja aplicação já foi discutida e o RRI (Reverse Rate Indicator) que informa a ERB a taxa que está sendo utilizada no enlace reverso.

A partir do momento que uma conexão é estabelecida entre a ERB e o TA este último transmite continuamente nos canais piloto, DRC e RRI. No entanto o terminal só transmite no canal ACK quando é detectada a transmissão de um pacote de dados a ele destinada no enlace direto. A função do canal ACK é, portanto, informar se um pacote enviado no enlace direto foi recebido corretamente. No enlace reverso, assim como em outros sistemas CDMA, é possível e desejável o uso de *soft handoff*.

2.3 Cadeias de Markov

Tempo discreto

Uma cadeia de Markov [51, 52] de tempo discreto é uma sequência de variáveis aleatórias, $X_n \in \mathbb{N}$, que definem um espaço de estados S e possui matriz de transição $P = p_{ij}$ ($i, j \in S$), tais que as seguintes propriedades procedem:

1. a distribuição inicial é descrita pelo vetor $\pi_i(0)$ ($i \in S$);
2. Para quaisquer $i, j \in S$, e $n \in \mathbb{N}$: $P[X_{n+1} = j | X_n = i] = p_{ij}$. Os elementos da matriz P representam as probabilidades de transição em um passo;

3. Para quaisquer $n \in \mathbb{N}$ e $i_0, i_1, i_2, \dots, i_{n-1}, i_n, j \in S$:

$$P[X_{n+1} = j | X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i_n] = P[X_{n+1} = j | X_n = i_n],$$

ou seja, (X_{n+1}) depende apenas do estado presente (X_n) .⁰ Esta propriedade é conhecida na literatura como propriedade da falta de memória (*memoryless property*) [52], ou propriedade de Markov.

Um estado i possui período t_i quando qualquer retorno ao estado i ocorre em um número de passos múltiplo de t_i . Se $t_i = 1$, então o estado i é dito aperiódico. Se $t_i = 1 \forall i \in S$ então a cadeia de Markov é chamada de aperiódica.

Se qualquer estado de uma cadeia pode ser alcançado, em um número finito de passos, a partir de qualquer outro estado, a cadeia é definida como sendo do tipo irreduzível.

Em uma cadeia do tipo irreduzível, aperiódica e finita, a solução estacionária para a distribuição de probabilidade dos estados é única, não depende do estado inicial e pode ser obtida a partir do sistema de equações lineares $\pi = \pi P$, considerando a condição de normalização dada por [52]¹: $\pi e^T = 1$.

Tempo contínuo

Uma cadeia de Markov de tempo contínuo é um processo estocástico que satisfaz a propriedade de Markov e que os passos que geram transição entre estados ocorrem em tempo aleatório exponencialmente distribuído. A matriz geradora infinitesimal Q define as taxas de transições $q_{i,j}$ ($i, j \in S$), de maneira análoga à matriz P construída para o caso discreto.

Sua solução, em estado estacionário, é obtida através da solução do sistema de equações lineares dado por $\pi Q = 0$, considerando a normalização $\pi e^T = 1$.

Como uma cadeia contínua é análoga [53] a uma cadeia discreta, é possível mapear o caso contínuo em um caso discreto similar. Desta forma pode-se trabalhar com uma ou outra representação de forma a facilitar a modelagem. Usaremos uma ou outra representação da modelagem introduzida no capítulo 5.

¹O vetor e é um vetor com valor 1 em todas suas posições.

2.4 Uma breve apresentação sobre codificação de vídeo escalável

O objetivo desta seção é apresentar a ideia básica de codificação escalável de alta granularidade (Fine Granularity Scalability - FGS) de vídeo, bem como ilustrar como o fluxo de vídeo armazenado pode ser truncado com o propósito de reduzir a taxa necessária para sua transmissão. De acordo com o método de codificação FGS, o resultado da decodificação de um fluxo truncado deve ser comparável à qualidade de um fluxo de taxa inferior.

Em um sistema de vídeo tradicional, o codificador comprime o sinal de entrada de vídeo em um fluxo de taxa de bits que deve ser menor ou igual à capacidade do canal. Neste caso o desafio da codificação do vídeo é formulado como um problema de otimização onde se deseja maximizar a qualidade do vídeo *dada a capacidade do canal*.

Para que essa formulação seja válida, primeiramente o codificador deve conhecer previamente a capacidade do canal e o cliente deve ser capaz de decodificar todos os bits do fluxo com agilidade suficiente para reconstruir o vídeo.

Em redes sem fio de taxa variável, não se conhece a taxa real de transmissão do canal. Isso ocorre pois, como apresentado na Seção 2.1.4, os padrões de transmissão sem fio adaptam a taxa de transmissão de sua camada física com o objetivo de manter a taxa de erro sob controle.

Uma solução trivial seria codificar o vídeo em uma qualidade que resulte em uma taxa média equivalente à da rede sem fio em questão. Porém, além de haver a necessidade de codificar um fluxo de vídeo diferente para cada perfil de taxa, esta solução não garante que o vídeo seja entregue ao receptor à mesma taxa de consumo do vídeo, já que a taxa do enlace sem fio pode variar de ordens de grandeza (por exemplo, veja a Tabela 2.1).

O cenário ilustrado na Figura 2.5 mostra períodos em que a taxa do canal está abaixo da taxa de consumo do vídeo, indicando que o *playout buffer* do cliente está esvaziando. Se esta situação perdura por longos períodos (ex: desvanecimento lento), o *buffer* do cliente potencialmente se esvazia.

Este fenômeno gera dois problemas altamente indesejados pelo usuário:

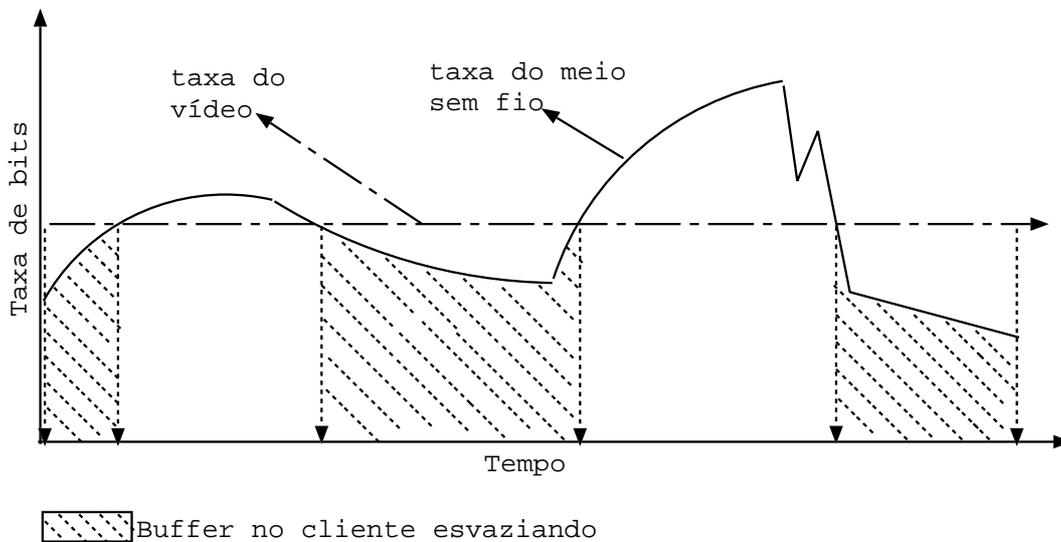


Figura 2.5: Quando a taxa de um enlace sem fio cai abaixo do requerido pelo vídeo, o *buffer* do cliente é consumido, podendo cair a zero, o que gera problemas de continuidade (pausas)

- aumento da taxa de perda de pacotes que é percebido pelo usuário através de incômodos artefatos nas imagens reproduzidas, e
- interrupção da apresentação do vídeo, ou seja, o vídeo pára de ser apresentado até que a taxa de rede volte a subir e *buffer* do cliente volte a acumular quantidade suficiente de bits.

Uma das soluções possíveis para este problema é codificar o vídeo mais de uma vez e gerar dois ou mais fluxos distintos a serem armazenados no servidor, cada um com uma qualidade (taxa) diferente. O servidor poderia trocar o fluxo que está enviando ao cliente de acordo com as flutuações na taxa. Este método de transmissão de vídeo é conhecido como *simulcast* [54].

Um dos problemas desta solução é que há replicação de informação nos discos, o que é ineficiente. Além disso, para que o aumento na taxa de transmissão no enlace sem fio resulte na melhoria do vídeo apresentado ao usuário, ele deve ser significativo o suficiente para alcançar a taxa do próximo fluxo de qualidade superior disponível no servidor.

Para contornar o problema da replicação de informação nos discos, surgiram propostas de codificação escalável em camadas [55, 54]. Estes métodos codificam e

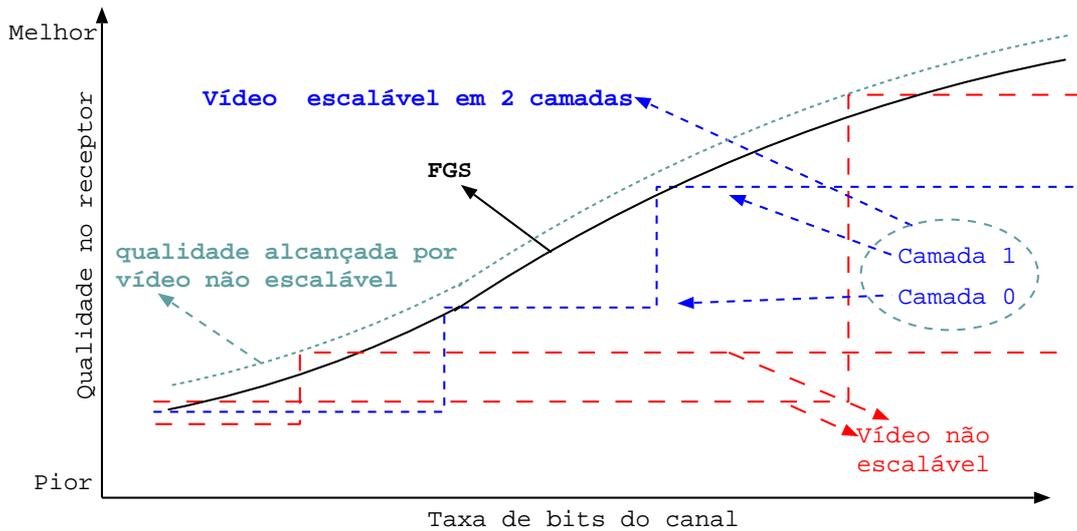


Figura 2.6: Codificação escalável vs. codificação não escalável

armazenam, em um mesmo arquivo no servidor, duas ou mais opções de qualidade/taxa, através do uso de uma camada básica e uma ou mais camadas usadas para melhorar a qualidade da camada básica.

Esta solução ainda possui a limitação de que só há alteração de qualidade quando a alteração na taxa do enlace sem fio alcançar a taxa requerida pela próxima camada disponível.

Por fim, a codificação escalável de granularidade fina (FGS - *Fine Granularity Scalability*) [18, 19, 20] permite que a qualidade de um vídeo enviado ao receptor acompanhe qualquer alteração na taxa disponível, como apresentado na Figura 2.4. Nesta solução, o desafio da codificação do vídeo é formulado como um problema de otimização onde deseja-se maximizar a qualidade do vídeo *dado um intervalo de capacidades do canal*, em contraste com a codificação tradicional que busca maximizar a qualidade do vídeo *dada uma capacidade específica*.

O vídeo pode ser exibido a uma taxa qualquer contida no intervalo $[R_{min}, R_{max}]$, onde R_{min} é a menor e R_{max} é a maior taxa disponível no fluxo codificado. A principal limitação desta solução, ilustrada na Figura 2.4 é que para uma mesma taxa R_i a qualidade de um vídeo FGS fica abaixo da qualidade de um vídeo codificado usando as soluções tradicionais.

Os quadros de um fluxo de vídeo FGS possuem a estrutura ilustrada na Figura 2.7. A camada base (BL - *Base Layer*), é codificada através do método não escalável tradicional. A camada de melhoramento (EL - *Enhancement Layer*) é co-

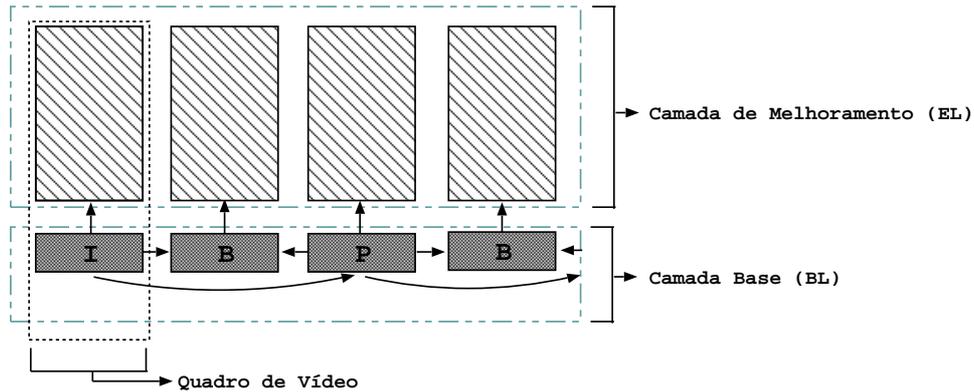


Figura 2.7: Exemplo da estrutura de quadros de vídeos codificados usando escalabilidade FGS

dificada através da computação da diferença entre a imagem original e a imagem resultante da reconstrução da BL. Como consequência, o decodificador (ex: cliente) deve possuir a BL para que seja possível decodificar a EL associada, ou seja, existe uma relação de dependência: a EL depende da BL.

A EL é organizada partindo-se dos *bits* (*bitplane*) menos significativos para os mais significativos. Por isso, a EL pode ser truncada em qualquer ponto, e a qualidade final do quadro é diretamente proporcional ao número de *bits* decodificados com sucesso para cada imagem/quadro.

Em uma aplicação de fluxo de vídeo FGS, a taxa de bits (*bitrate*) variável é alcançada através da emissão, pelo servidor, da BL mais qualquer porção truncada da EL, de acordo, por exemplo, com a taxa disponível no enlace sem fio onde o cliente está conectado. Claramente, se apenas a BL é emitida, a taxa mínima R_{min} é requerida para tocar o vídeo. Analogamente, a taxa máxima R_{max} é obtida quando a BL e todos os *bits* da EL são recebidos pelo cliente.

2.5 Conceitos de Justiça

O termo justiça (do inglês *fairness* [56, 57]), no contexto de redes de computadores, está relacionado com a distribuição dos recursos da rede entre usuários. Quanto mais uniforme for a distribuição de recursos, maior é o nível de justiça alcançado.

Em geral, o estudo da justiça na área de redes em computadores tem como objetivo aprimorar o comportamento de arquiteturas de rede para que estas incorporem

o valioso conceito de distribuição igualitária de recursos.

Uma questão importante, quando se avalia a justiça, é a definição do recurso (ex: métrica) a ser estudado.

No Capítulo 3, serão abordados diversos trabalhos que tratam de justiça em redes sem fio de longa distância. Como discutiremos, a métrica de justiça usada por estes métodos é o número de *slots* de tempo ou o número de oportunidades de transmissão alocado a cada usuário. Porém, como discutiremos no Capítulo 4, o número de *bits* emitidos por unidade de tempo depende da modulação e FEC utilizadas. Deste modo, mesmo que seja alocada a mesma quantidade de *slots* para os usuários, isso não significa que eles alcançarão justiça quando se considera métricas como a vazão e o atraso.

Portanto, nosso trabalho se diferencia dos disponíveis na literatura porque avaliamos a justiça através da análise da vazão (do inglês *goodput*), e do atraso enfrentados por usuários.

Definimos o *goodput* como a vazão de tráfego *web* para os clientes, e o atraso

2.6 Conceitos de Justiça

O termo justiça (do inglês *fairness* [56, 57]), no contexto de redes de computadores, está relacionado com a distribuição dos recursos da rede entre usuários. Quanto mais uniforme for a distribuição de recursos, maior é o nível de justiça alcançado.

Em geral, o estudo da justiça na área de redes de computadores tem como objetivo aprimorar o comportamento de arquiteturas de rede para que estas incorporem o valioso conceito de distribuição igualitária de recursos.

Uma questão importante, quando se avalia a justiça, é a definição do recurso (ex: métrica) a ser considerado.

No Capítulo 3, serão abordados diversos trabalhos que tratam de justiça em redes sem fio de longa distância. Como veremos, a métrica de justiça usada por estes métodos é o número de *slots* de tempo ou o número de oportunidades de transmissão alocado a cada usuário. Porém, como discutiremos no Capítulo 4, o número de *bits* emitidos por unidade de tempo depende da modulação e FEC utilizadas. Deste modo, mesmo que seja alocada a mesma quantidade de *slots* para os usuários,

isso não significa que eles alcançarão justiça quando se consideram métricas como a vazão e o atraso.

Portanto, nosso trabalho se diferencia dos disponíveis na literatura porque avaliamos a justiça através da análise da vazão obtida por cada usuário (do inglês *goodput*), e do atraso enfrentado pelos usuários.

Definimos o *goodput* como a vazão de tráfego *web* para os clientes, e o atraso como sendo o tempo medido entre o momento em que o cliente clica em uma hiperligação e o instante em que a página solicitada está completamente apresentada na tela.

Capítulo 3

Trabalhos Relacionados

Este capítulo apresenta os principais trabalhos relacionados com esta tese. Os primeiros dois conjuntos de trabalhos, apresentados na Seção 3.1, estão relacionados com as contribuições apresentadas no Capítulo 4.

Na Seção 3.1.1, apresentamos as principais propostas de algoritmos de escalonamento *cross-layer* de pacotes que objetivam fornecer algum nível de justiça a tecnologias como a EVDO.

Em seguida, na Seção 3.1.2, comentamos sobre os trabalhos que contribuem com modelos de simulação para avaliação de desempenho do EVDO, e discutimos que a literatura disponível tipicamente estuda o desempenho alcançado através da vazão agregada ou total. Entretanto há uma lacuna no que diz respeito a estudos detalhados sobre a justiça alcançada por este padrão, usando como métrica a vazão em função da posição do usuário na célula.

O terceiro e o quarto grupo de trabalhos, apresentados na Seção 3.2, são importantes para que se entenda o escopo e as contribuições apresentadas no Capítulo 5.

A Seção 3.2.1 trata de trabalhos em que a ideia central consiste em adaptar a taxa de emissão de fluxos de vídeo escalável de acordo com variações na taxa de transmissão do enlace sem fio ao qual o cliente está conectado. Com isso, busca-se diminuir o esvaziamento do *playout buffer* do cliente e, também, a perda de pacotes observada na fila do roteador adjacente ao enlace sem fio. Embora os trabalhos desta subseção não sejam, em primeira análise, diretamente relacionado à proposta do Capítulo 5, o entendimento de suas contribuições e limitações são importantes para o compreender do escopo do nosso trabalho. Isso ocorre, pois o objetivo final

destes trabalhos é o mesmo da proposta do Capítulo 5: melhorar a qualidade do fluxo de vídeo. Nos trabalhos apresentados, a melhoria de qualidade é alcançada através do controle de taxa que influencia positivamente o nível de perda agregada. Por outro lado, nossa proposta consiste em adaptar o nível de rajada de forma a alterar o processo de perda de tal maneira que traga benefícios para a qualidade percebida pelo usuário.

A Seção 3.2.2 trata de trabalhos que propõem o uso de escalonamento com prioridade, ou da combinação de escalonamento e descarte com prioridade, com o objetivo de abrandar os efeitos da perda sobre a QoS de fluxos de vídeo escalável. Este objetivo é alcançado servindo-se prioritariamente pacotes que tem maior impacto positivo sobre a QoS (ex: pacotes BL). Além disso, quando a fila sofre uma sobrecarga, pacotes que causam pouco impacto negativo na QoS são descartados prioritariamente em relação aos demais (ex: pacotes EL contendo os últimos níveis de codificação). Nossa proposta evita duas importantes limitações deste grupo de trabalhos: alta complexidade e a necessidade intrínseca de mudanças na infra-estrutura (roteadores) da Internet.

3.1 EVDO rev.0: Desempenho e justiça do padrão de acesso EVDO rev.0

3.1.1 Algoritmos de escalonamento para redes sem fio de taxa variável

Diversos algoritmos de escalonamento foram propostos para que redes sem fio de longa distância convencionais, ex. celulares, passassem a ser capazes de operar através de comutação de pacotes. Um subconjunto destes algoritmos [58, 59, 60, 61, 62, 63] discute o *Proportional Fair Scheduler* (PFS) e propõe adaptações a ele.

O PFS foi proposto inicialmente em [58]. Este escalonador busca tirar vantagem das variações temporais do canal dando prioridade de transmissão a usuários que enfrentam as melhores condições em suas camadas físicas. O objetivo deste algoritmo é maximizar a vazão do canal direto. Os resultados do trabalho são apresentados na forma de vazão agregada média do canal direto em função do número de usuários

simultâneos no sistema.

Em [59] os autores argumentam que o PFS garante vazão adequada, mas o algoritmo não consegue lidar com questões de justiça que são características de fluxos de pacotes sensíveis ao atraso (e.g. vídeo em tempo real). Com o objetivo de atacar estas questões, eles propõem um escalonador chamado *Weighted Fair Queuing Proportional Fair* (WFQ-PF). Este escalonador é capaz de aumentar levemente a justiça para o caso de o sistema não estar sobrecarregado (e.g. 10 usuários). Os autores assumem que o tamanho do pacote para os usuários com piores condições de canal ocupa apenas um *slot* de tempo. Logo, a métrica para a justiça, neste caso, é o número de *slots* alocados ao usuário. Note que esta suposição é muito limitadora, já que os padrões de redes sem fio de longas distâncias usam modulação adaptativa, ou seja, usuários que enfrentam condições deterioradas de canal usam mais *slots* para transmitir um número inferior de *bits*. No modelo de simulação desenvolvido nesta tese, consideramos detalhadamente o uso de modulação adaptativa. Além disso, estudamos a justiça usando a vazão dos usuários como métrica, uma vez que a qualidade percebida por um usuário é função da vazão por ele obtida.

A adição, ao PFS, de suporte a prioridade e QoS é discutida em [60]. O trabalho conclui que há um balanceamento entre a vazão do sistema e o nível de QoS desejado. Em outras palavras, os parâmetros de QoS escolhidos são alcançados ao custo de decréscimo na vazão média do sistema.

Andrews [62] mostra que o PFS e, também, algumas variações deste algoritmo apresentam instabilidade no limite superior de suas filas sob alguns cenários. Os autores afirmam que esta informação é de interesse já que o PFS foi proposto como o algoritmo de escalonamento para o sistema sem fio de dados de alta velocidade (HDR - High Data Rate).

O trabalho [63] propõe um novo algoritmo de escalonamento com suporte a QoS. Os autores dividem uma única célula em cinco regiões de distância (em relação à estação base) e assumem que todos os usuários em uma mesma região percebem o mesmo SINR, mas cada usuário percebe flutuações instantâneas devido ao desvanecimento. Além disso, assumiu-se que todos os usuários são estacionários e distribuídos uniformemente entre as cinco regiões. A população considerada foi de vinte usuários, sendo quatro por região. Então, foram consideradas quatro diferentes

prioridades no PFS. Adicionalmente, cada usuário de uma determinada região recebe uma destas quatro prioridades disponíveis. Os autores concluem que é possível dar suporte a algum nível de diferenciação de QoS entre usuários ao preço de leve degradação da vazão média do sistema. As principais diferenças entre o trabalho apresentado neste artigo, e esta tese são que: (i) não nos restringimos a avaliar o nível de diferenciação de serviço alcançável pelo PFS, (ii) não dividimos os usuários em subgrupos de qualidade de canal, mas sim modelos separadamente a qualidade do canal de cada usuário e (iii) argumentamos que o nível de diferenciação de serviço alcançado através de ajustes no PFS não é suficiente para refletir em justiça quando se considera métricas como a vazão e o atraso dos usuários.

Uma comparação de vários algoritmos de escalonamento, incluindo o PFS, foi apresentada em [61]. O resultado obtido é a vazão média por setor para diferentes cenários de mobilidade em função do número de usuários no sistema. Além disso é apresentada a função de distribuição cumulativa da vazão por usuário para fins de comparação entre os algoritmos estudados. Através da análise destas métricas os autores concluem que o algoritmo que alcança a maior vazão por setor sofre de maior variação na vazão por usuário. Por outro lado, o algoritmo que obteve menor vazão por setor alcançou maior justiça na vazão individual dos usuários.

Outros algoritmos de escalonamento mais elaborados foram propostos pelos trabalhos [64, 65, 66, 67, 63, 68]. O *Modified Largest Weighted Delay First (M-LWDF)* foi proposto por [64]. Uma das principais conclusões do trabalho indica que é possível atender usuários de tempo real sobre um canal sem fio compartilhado, tal como o CSMA/HDR (e.g. EVDO). O trabalho em questão também mostra que o algoritmo M-LWDF pode ser usado para: (1) maximizar o número de usuários que podem co-existir com a QoS especificada; (2) prover algum nível de diferenciação de QoS (ex: número de *slots* alocados); e (3) fornecer garantia de vazão mínima e isolamento de fluxos.

O trabalho de [65] usou o (PSNR) como métrica de qualidade para avaliação de fluxos de vídeo enviados sobre canais sem fio que usavam um algoritmo de escalonamento de pacotes proposto pelos próprios autores: o *Earliest Deadline First (EDF)*. A principal ideia do trabalho é usar um marcador de tempo para escalar pacotes de vídeo baseado na importância atribuída a cada um desses pacotes.

Como principal resultado, os autores observaram melhoras no PSNR quando o EDF foi usado para escalonar pacotes em comparação com esquemas convencionais de escalonamento.

Outro escalonador adaptativo é proposto em [66]. O algoritmo possui parâmetros que podem ser ajustados para explorar o balanceamento entre a vazão média do sistema e a justiça em termos do número médio de *slots* atribuído a cada usuário e também a sua taxa de transmissão. Os autores mostraram que o preço para alcançar a justiça entre usuários é uma séria degradação na vazão média do sistema. Ressaltamos, novamente, o uso de métricas de justiça que refletem o ponto de vista da rede.

Kong and Teh [67] discutiram o escalonador *Realistic Proactive Earliest Due Date* (R-PEDD+), o qual antecipada e dinamicamente ajusta o prazo de expiração de um pacote de acordo com futuras mudanças nas condições do canal. Este escalonador deriva a informação necessária sobre mudanças no canal a partir de transmissões recentes de pacotes (medições passivas). Foram obtidos a vazão média do sistema e a taxa de pacotes descartados para diferentes condições do canal e diferentes parâmetros do R-PEDD+. Uma das principais conclusões foi que o R-PEDD+ é capaz de superar, em alguns dos casos considerados, o algoritmo (FEDD - *feasible earliest due date*), que garante limite superior para o atraso em redes sem fio de taxa variável.

Por fim, o artigo [68] propôs o algoritmo *Urgency and Efficiency Based Wireless Packet Scheduling* (UEPS) que poderia ser usado, também, para tráfego de tempo real. O UEPS é comparado tanto com o PFS quanto com o *Modified Largest Weighted Delay First* (M-LWDF). Os resultados de experimentos de simulação mostraram que o UEPS é capaz de alcançar melhor vazão que o PFS e o M-LWDF ao mesmo tempo que é capaz de satisfazer as necessidades de QoS de tráfegos de tempo real.

A Tabela 3.1 resume os trabalhos relacionados apresentados nesta seção. Esta tabela nos ajuda observar que as métricas usadas para avaliar a justiça normalmente refletem o ponto de vista da rede.

Nesta tese buscamos uma avaliação diferenciada na qual avaliamos a justiça usando métricas que refletem o ponto de vista do usuário. Para isso usamos o *goodput* e o atraso observados pelo usuário para medir a qualidade de serviço e a

Trabalho	Métricas avaliadas	Principal Resultado/Contribuição
Jalali <i>et al.</i> [58]	Vazão do canal direto	PFS maximiza a vazão
Rhee <i>et al.</i> [59]	Núm. de slots alocados	Discute problema de justiça (slots) do PFS
Kuenyoung <i>et al.</i> [60]	Vazão do canal direto	Adiciona algum nível de QoS ao PFS
Andrews [62]	Tamanho das filas	Discute instabilidade do PFS em relação ao tamanho das filas
Ban [63]	Vazão agregada em cada região	Adiciona algum nível de QoS com leve degradação da vazão
Andrew <i>et al.</i> [64]	Atraso Vazão	Proposta do M-LWDF: suporte a usuários de tempo real
Fan <i>et al.</i> [66]	Núm. de slots alocados	Algoritmo adaptativo que permite ajuste entre justiça e vazão
Kong <i>et al.</i> [67]	Vazão agregada Taxa de perda	Escalonamento baseado em medições de pacotes recém enviados
Ryu <i>et al.</i> [68]	Vazão agregada, atraso	UEPS: melhorias em relação ao M-LWDF

Tabela 3.1: Resumo dos algoritmos apresentados na Seção 3.1: Algoritmos de Escalonamento para redes sem fio de taxa variável

justiça nos cenários considerados. No Capítulo 4 nós discutiremos que a justiça, quando se considera métricas como o número de *slots* alocados por usuário, não necessariamente equivale à justiça do ponto de vista de métricas como a vazão e do atraso. Seremos capazes, inclusive, de quantificar esta diferença. Esta é uma das contribuições de nosso trabalho.

3.1.2 Modelagem e análise de desempenho do padrão EVDO

A maioria dos trabalhos de avaliação de desempenho de redes EVDO analisam a vazão média alcançada pelos canais direto e reverso. Alguns dos trabalhos consideram modelos de tráfego muito simples como, por exemplo, número infinito de fontes de dados e modelos simplificados de protocolos como FTP, HTTP e WAP.

Os primeiros dois trabalhos [58], [6] na área de avaliação de desempenho do EVDO analisavam apenas um dos canais (direto ou reverso). Em [58] o canal direto do padrão CDMA2000 1xEVDO foi simulado no intuito de encontrar sua vazão média em função do número de usuários simultâneos no sistema. No modelo implementado, usuários escolhem uniformemente uma nova posição em intervalos de 30s. Em cenários considerando antenas setorizadas e antenas omnidirecionais foram encontradas as vazões de 1.1Mbps e 1.5Mbps, respectivamente.

Esteves [6] estudou a capacidade máxima do canal reverso da rede EVDO usando tanto modelos analíticos quanto de simulação. Seu trabalho mostrou que a capacidade do canal reverso decai com o crescimento do número de usuários devido ao *overhead* dos canais de controle (ex: *pilot*, DRC). Além disso, uma capacidade máxima entre 250 a 600 kbps foi encontrada, dependendo do número de antenas receptoras, do ambiente de interferência e do perfil de mobilidade empregado.

Seguindo estes trabalhos iniciais, um conjunto de artigos [69, 5, 70, 71] estudou ambos os canais direto e reverso com o objetivo de analisar o desempenho do sistema. Estes artigos se interessavam, principalmente, em encontrar a vazão média alcançada.

O trabalho de Black e Gurelli [69] propôs modelos de simulação, em nível de rede, usando modelos de canais ITU (*International Telecommunication Union*) bem conhecidos. Seu objetivo era encontrar a vazão agregada por setor em função do número de usuários no sistema. Entre os resultados, destaca-se a vazão máxima de 1.4Mbps para um cenário misto de usuários de alta e baixa mobilidade.

Bi e Vitebsky [5] caracterizaram a vazão do sistema EVDO em diferentes ambientes de radiofrequência. O principal resultado obtido foi a vazão média por setor em função de ambos o número de usuários e os cenários de mobilidade considerados. Além disso, eles propuseram um método para calcular a população máxima supor-

tada pelo sistema de modo a manter maximizada a vazão média por setor. Assim, eles encontraram o valor máximo de 1.28Mbps de vazão média para uma população de 20 usuários.

Mohanty *et al.* [70] analisaram o desempenho, em nível de aplicação, do EVDO. Para isso eles focaram, principalmente, a interação entre o mecanismo de controle de congestionamento do TCP e as variações do canal sem fio. Eles obtiveram a vazão agregada média, por setor, do canal direto em diferentes situações de desvanecimento e considerando diferentes configurações de DRC. Assim eles mostram que, em um cenário com 4 usuários FTP e 16 usuários *web*, o tempo médio de resposta é comparável a um canal ISDN de 128kbps, dedicado, a cada usuário. Finalmente, são apresentadas evidências de que, quando o sistema EVDO é implementado com parâmetros apropriados, a perda na vazão devido à interações entre o TCP e a variabilidade do canal é mínima.

Huang *et al.* [71] estudaram a vazão agregada em diferentes condições de desvanecimento do canal, de configurações do DRC e de níveis de potência de canais de controle. Sob as condições consideradas, a vazão agregada do canal direto variou de 500 a 600kbps.

Os trabalhos [72, 73, 74] também se concentraram no desempenho do IS-856, mas estudaram sensibilidade de ajuste de parâmetros que diferem dos trabalhos acima. Esteves *et al.* [72] estudou o desempenho do EVDO considerando apenas usuários fixos e obteve: (1) a vazão máxima alcançável pelo sistema (que foi de 1.57Mbps) e (2) a vazão por usuário em função do *path-loss* considerando dois algoritmos de escalonamento: o PFS e o GOS=1. Este, juntamente com [61, 60], é um dos primeiros trabalhos que comenta sobre o balanceamento entre a capacidade do setor e a justiça. Ressaltamos, mais uma vez, que a justiça é estudada através de ajuste de parâmetros em algoritmos de alocação de *slots*, ou seja, o uso de métricas de justiça que refletem o ponto de vista da rede.

Trabalho	Métricas avaliadas	Principal Resultado/Contribuição
Esteves <i>et al.</i> [6]	Capacidade do canal reverso	Capacidade de 1.1 a 1.5Mbps
Black <i>et al.</i> [69]	Vazão agregada por setor	Vazão de 1.4Mbps
Bi <i>et al.</i> [5]	- Vazão agregada - 98º percentil da vazão dos usuários	Vazão de 1.28Mbps sob pop. de 20 usuários
Mohanty <i>et al.</i> [70]	Vazão agregada por setor	Estudo da interação entre o EVDO e o TCP
Huang <i>et al.</i> [71]	Vazão agregada sob dif. desvanecimentos	Vazão de 500 a 600Kbps sob o cenário considerado
Esteves <i>et al.</i> [72]	Vazão por setor	Discussão da relação entre vazão do setor e a justiça. Vazão de 1.57Mbps
Xiao <i>et al.</i> [74]	Vazão média por setor	Vazão de 500Kbps para usuários <i>web</i>

Tabela 3.2: Resumo dos algoritmos apresentados na Seção 3.2 - Modelagem e Análise de desempenho do padrão EVDO

Xiao *et al.* [74] usou um simulador em nível de *chip* da especificação original do canal direto do EVDO com o objetivo de comparar seu desempenho com outras versões do sistema, como o CDMA-1x e o 1xEV-DV. O principal resultado é a vazão agregada média do canal direto para algumas aplicações como FTP, WAP e *web*. Eles concluíram que a vazão média por setor é de 500kbps para tráfego *web* e 1.5Mbps para tráfego FTP. Adicionalmente, a eficiência espectral - quantidade de informação transmitida em uma faixa de frequência - do EVDO e do EVDV se mostrou bem maior que a do CDMA2000-1x, segundo os resultados apresentados.

Como sumarizado pela Tabela 3.2, os trabalhos referenciados nesta seção não analisam em detalhe a vazão e o atraso alcançados por usuários em diferentes pontos arbitrários da célula. Esta análise, juntamente com uma implementação mais

detalhada do sistema, e a está entre as contribuições dessa tese.

3.2 Vídeo escalável: Escalonamento, descarte e perda

Esta seção faz uma revisão geral dos trabalhos relacionados às contribuições do Capítulo 5.

3.2.1 Adaptação da taxa no servidor

A ideia central, para este grupo, consiste em adaptar a taxa de emissão de fluxos de vídeos escaláveis de acordo com variações no enlace sem fio ao qual o cliente está conectado. Com isso, busca-se diminuir o esvaziamento do *playout buffer* do cliente e, também, a perda de pacotes observada na fila do roteador adjacente ao enlace sem fio.

Em geral, as propostas deste subconjunto de trabalhos combina o uso de fluxos escaláveis de vídeo, como o MPEG4-FGS [18, 75] e o H.264/SVC [76], com técnicas de adaptação da taxa na camada de aplicação [77, 22, 78, 23, 79]. Uma desvantagem dessa abordagem é sua precisão limitada [24] em relação previsão da variação da taxa no enlace. O compartilhamento do enlace sem fio e o desvanecimento de tempo curto geram flutuações que não podem ser previstas e, portanto, podem causar níveis elevados de perda, o que afeta consideravelmente [80, 81, 82, 83] a QoS.

A princípio pode parecer simples contornar esse impasse. Bastaria adotar uma adaptação mais conservadora onde a taxa de envio do vídeo seria potencialmente menor do que a taxa disponível no canal. Entretanto é bem conhecido [84, 85, 86, 75, 87] que a qualidade resultante do vídeo é diretamente proporcional à taxa do fluxo que é transmitido com sucesso, logo a proposta pode causar uma queda desnecessária na qualidade do vídeo obtido pelo usuário.

3.2.2 Escalonamento e descarte de pacotes com prioridade aplicados a fluxo de vídeo escalável

Como explicado na seção anterior, a adaptação da taxa de envio no servidor não é, por si só, suficiente para alcançar a melhor QoS quando se trata de fluxo de vídeos escaláveis. Isso ocorre pois mesmo com o uso das ideias propostas na literatura há períodos de tempo em que flutuações imprevisíveis no enlace sem fio resultam em importante degradação da QoS.

Com isso, o uso de escalonamento com prioridade ou da combinação de escalonamento e descarte com prioridade foram propostos por outro grupo de trabalhos disponíveis na literatura. O objetivo principal, neste caso, é abrandar os efeitos da perda sobre a QoS de fluxos de vídeo escalável. Este objetivo é alcançado através do serviço prioritário de pacotes que tem maior impacto sobre a QoS (ex: pacotes BL). Além disso, quando a fila sofre uma sobrecarga, pacotes que causam menor impacto positivo sobre a QoS são descartados prioritariamente (ex: pacotes EL contendo os últimos níveis de melhoramentos).

Abaixo descrevemos sucintamente os principais trabalhos relacionados à esta ideia. Ao final desta seção destacamos as principais diferenças entre os trabalhos apresentados e a proposta apresentada no Capítulo 5.

O uso de descarte diferenciado de pacotes (UPP - *Unequal Packet Protection*) com aplicação a fluxos escaláveis de vídeo foi proposto por Van der Schaar *et al.* [88, 27]. O trabalho considera cenários de redes ATM não confiáveis e propõe uma heurística onde são considerados três níveis de prioridade entre quadros: alta prioridade para pacotes referentes a quadros I, média prioridade para pacotes que correspondem a quadros P e baixa prioridade para pacotes B. Além disso, pacotes contendo BL possuem maior prioridade que pacotes EL. Usando o simulador NS, os experimentos realizados comparam o PSNR do serviço de melhor esforço com o uso de UPP. Os resultados mostraram melhorias de $\approx 24dB$ para $\approx 30dB$ no PSNR médio de fluxos FGS, dependendo da taxa média do vídeo.

Com ideias análogas a [88, 27], Liebl *et al.* [89] realizaram experimentos de simulação considerando uma célula de uma rede HSDPA 3G. No cenário considerado, flutuações na taxa de serviço são causadas tanto pelas variações no enlace sem fio quanto pela disputa por recursos, supondo 4 usuários. Quando o UPP foi usado, o

PSNR quadro-a-quadro dos períodos de congestionamento passou de $\approx 3dB$ para $\approx 14dB$.

O trabalho de [24] propõe que a adaptação da taxa do vídeo seja feita no último roteador, ligado a um enlace sem fio. O cenário considerado supõe um canal Gilbert-Elliot [90]. A ideia consiste na análise do conteúdo de pacotes de vídeo com o objetivo de descartar porções de menor impacto à QoS, com o objetivo de adaptar a taxa do fluxo ao estado do enlace sem fio. Um algoritmo de otimização é usado para escolher a porção do vídeo a ser descartado. Com isso, o PSNR quadro-a-quadro alcançado é superior ao das propostas de UPP, já que estas são baseadas em heurísticas.

Pliakas *et al.* [91] realizaram experimentos de simulação onde demonstram que a combinação do uso de fluxos de vídeos escaláveis com a correta configuração de serviços 802.11e [92]/Diffserv [93] resulta em melhor QoS. Para isso, realizaram experimentos onde comparam o PSNR de transmissões de vídeo usando serviço de melhor esforço combinado com Diffserv. Os resultados mostram que o uso de Diffserv diminui a degradação do PSNR nos momentos de severo congestionamento provocado pela inserção de tráfego de fundo.

Um rede 802.11e *ad hoc* é considerada em [25]. A principal contribuição do artigo consiste em um algoritmo detalhado para escolher as prioridades de pacotes para quem sejam alocados a diferentes categorias de acesso (*AC - Access Category*) de acordo com sua importância para QoS. Um cenário de simulação considerando 4 nós, dois emissores e dois receptores, é usado para estudar o ganho atingido quando a priorização é usada. Cada nó emissor gera 4 fluxos de tráfego CBR (*Constant Bit Rate*) que agem como tráfego de fundo mais 1 fluxo do vídeo escalável a ser analisado. A capacidade das filas de cada emissor foi configurada para 50 pacotes. Os resultados obtidos mostraram que a seguinte compensação foi alcançada: por um lado os picos de baixo PSNR ($12 - 16dB$) foram eliminados, mas por outro lado os picos de alto PSNR ($32 - 36dB$) também foram eliminados. O PSNR alcançado oscilou entre 20 e $22dB$.

Uma combinação de escalonamento e descarte priorizados foi proposta por Qi-ong Li *et al.* [55], considerando uma rede WLAN - *Wireless Local Area Network* ad hoc (IEEE 802.11 [94]) com um emissor e um receptor. Adicionalmente, o trabalho propõe mudanças no cálculo do parâmetro RTRA (*Real-Time Retry-limit Adapta-*

tion) do padrão IEEE 802.11 para obter melhorias no PSNR de fluxo de vídeos escaláveis. Os resultados indicaram atenuação dos vales do PSNR calculado, assim como diminuição das pausas causadas por esvaziamento do *playout buffer* do cliente.

Em [26], Zhang *et al.* propuseram um algoritmo de escalonamento mais elaborado, brevemente explicado a seguir. Considera-se que o canal consiste em *slots* de tempo a serem alocados a usuários TDMA (*Time Division Multiple Access* [95]). O algoritmo de alocação de compartimentos de tempo é formulado como um problema de otimização onde supõe-se que as flutuações do canal de cada usuário segue uma distribuição estatística conhecida e que as mudanças de estado entre dois compartimentos adjacentes é independente. O objetivo da formulação considerada é maximizar a soma ponderada do PSNR do conjunto de usuários. Para isso, a escolha do próximo usuário a ser servido leva em consideração a relevância de seu pacote para a QoS, bem como o estado de seu enlace sem fio. Experimentos de simulação indicaram que o PSNR quadro-a-quadro é maior do que o obtido pelos principais algoritmos de escalonamento disponíveis na literatura.

Para facilitar a compreensão das diferenças entre os trabalhos comentados nesta seção e traçarmos um paralelo com nossas contribuições nesta área, um sumário é apresentado na Tabela 3.3.

Os esquemas propostos na literatura demonstram que é possível alcançar boa melhora na QoS através do uso de classificação e priorização de pacotes de fluxos de vídeo escalável. Porém uma desvantagem destes esquemas é a complexidade da necessidade de se processar camadas superiores para que seja possível classificar corretamente cada pacote.

Outra importante desvantagem é que as soluções existentes envolvem modificações em padrões [55], ou são direcionadas a padrões específicos [88, 27, 89, 91, 25]. A literatura carece, portanto, de um método que mais simples, que seja aplicável a diferentes padrões de acesso à Internet e que não incorra em modificações a equipamentos.

Os trabalhos relacionados apresentados nesta subseção, embora estudem algoritmos de priorização de pacotes, analisam o PSNR em função da perda média agregada. De acordo com o estudo levantado nesta tese, não foram feitos estudos

¹médio

²quadro-a-quadro

Ideia	Padrão/Cenário considerado	Métrica avaliada	Principal resultado/contribuição
descarte com prioridade [88, 27]	ATM sem prioridade	PSNR ¹	prioridade resultou em melhora de 24db para 30db
descarte com prioridade [89]	3G (HSDPA)	PSNR ²	prioridade resultou em melhora de 3dB para 14dB (quadros com perda)
adaptação da taxa usando descarte seletivo[24]	canal Gilbert-Elliot	PSNR ²	adaptação proposta gerou melhoras no número de pausas e no PSNR
classificação de pacotes para uso de ACs [91, 25]	802.11e [91] e 802.11e [25] ad hoc	PSNR ²	uso de ACs gerou melhoras no PSNR
escalonamento/descarte com prioridade, e alteração no RTRA [55]	802.11 (WLAN) ad hoc	PSNR ²	adaptações propostas geraram melhoras no número de pausas e no PSNR
otimizar a soma do PSNR através de escalonamento ótimo [26]	TDMA	PSNR ²	algoritmo proposto resulta em melhor PSNR do que os já disponíveis

Tabela 3.3: Resumo dos trabalhos apresentados na Seção 3.2.2: Escalonamento e descarte de pacotes com prioridade aplicados a fluxo de vídeo escalável

sobre o processo de perda e sua relação com os diferentes níveis de importância de cada pacote de vídeo escalável.

3.2.3 Discussão

O processo de perda e sua relação com o espaçamento entre pacotes (*loss process versus packet burstiness*) é uma área bem estudada. Alguns dos trabalhos focam na distribuição de conteúdo multimídia [96], enquanto outro grupo está direcionado a outras aplicações [97, 98]. Em geral, a medida de interesse estudada é a taxa média de perda [96, 97] ou métricas relacionadas ao tamanho da fila [98].

Os efeitos do tamanho de rajadas de perda sobre a distorção quadrática média de sequências de vídeo não escalável foi avaliado, por simulação, em [99]. Entretanto, os autores não consideraram vídeo escalável e, além disso, não estudaram a relação entre o nível de rajada e o processo de perda *intra-rajada*.

A proposta descrita no Capítulo 5 desta tese tenta resolver as limitações dos trabalhos relacionados apresentados nesta seção. Desenvolvemos modelos analíticos e de simulação que evidenciam que é possível obter melhorias na qualidade percebida mesmo sem o uso de algoritmos elaborados.

Isso é alcançado através de uma nova maneira de estudar o processo de perda em filas FIFO. Estudamos a relação entre (i) o espaçamento entre pacotes (nível de rajada) e (ii) o que chamamos de processo de perda *intra-rajada*. Esta abordagem nos permite mostrar que é possível usar inteligentemente o nível de rajada para alcançar sensíveis melhorias na QoS.

Capítulo 4

Modelo de uma célula EVDO, Experimentos e Resultados

Neste capítulo apresentamos um modelo de simulação para avaliação de desempenho e da justiça do padrão EVDO rev.0, bem como uma proposta para contornar limitações na justiça. O modelo e a proposta baseiam-se nos fundamentos teóricos apresentados no Capítulo 2 e algoritmos, recomendações e parâmetros disponíveis nos trabalhos relacionados do Capítulo 3, e divide-se em quatro partes: camada física, camada de enlace, camada dos modelos de usuários e o modelo de mobilidade.

4.1 Modelo de uma célula EVDO

4.1.1 Visão geral do Modelo

O modelo proposto consiste de uma população de usuários utilizando o sistema EVDO para acesso à Internet. A implementação foi realizada através da criação de três objetos: ERB, Internet e Usuário *web*, como pode ser visto na Figura 4.1.

O objeto *Usuário web* representa um usuário que acessa páginas *web* através de um navegador; o objeto *Internet* representa os servidores *web* disponíveis na grande rede; e o objeto *Estação Base* reproduz o comportamento de uma estação rádio-base segundo o padrão CDMA2000 1xEVDO.

Um diferencial do modelo proposto é que são levadas em consideração todas as características mais importantes das camadas físicas e de enlace, como os contro-

les de potência, o decaimento do sinal devido à propagação, a interferência intra e inter-celular, o controle de congestionamento, o *slow fading*, o algoritmo de escalonamento de pacotes e a mobilidade de usuários. A implementação deste conjunto de características em um único modelo é uma das contribuições deste trabalho, e resultou em um modelo, de código aberto, em nível de rede.

As seguintes considerações dão suporte ao modelo proposto:



Figura 4.1: Objetos implementados

- consideramos cenários com usuários estáticos, bem como cenários com um subconjunto dos usuários estáticos e os usuários restantes móveis. O modelo de mobilidade foi implementado com base em [14].
- cada usuário gera carga ao sistema de acordo com o modelo de tráfego associado, independentemente da carga gerada por outros usuários.
- os usuários pertencem a uma única ERB que utiliza antena omnidirecional e, portanto, não há *soft-handoff*. Note que embora o usuário não possa sair da célula em questão, o modelo leva em consideração a interferência inter-celular.
- embora o algoritmo *Proportional Fair Scheduler* não faça parte da especificação do sistema EVDO, ele é amplamente recomendado como o mais adequado para uso neste sistema. Portanto, usamos este algoritmo para o escalonamento de pacotes.
- a interferência do canal direto (*downlink*) é computada através do método geométrico.

- assumimos que a interferência inter-celular é de 40% do valor da interferência intra-celular.
- como a maioria dos terminais se move a baixas velocidades ou são estáticos, este cenário traz, por consequência, uma importante simplificação para a modelagem do canal direto no que diz respeito ao *Early Termination* (transmissão incremental dos *bits* de paridade) [100]. A variação nas condições do canal é sensivelmente menor, o que faz com que a estimativa do DRC feita pelo cliente, momentos antes de a ERB transmitir para o mesmo, não se torne obsoleta (o canal varia pouco) sendo, por isso, mais precisa. Logo não há, neste caso, ganho significativo na implementação do *Early Termination* e por isso não o implementamos em nosso modelo.

Nas seções que seguem, apresentaremos as características das camadas física e de enlace de dados implementadas no modelo. A subseção 4.1.5 apresenta o modelo de mobilidade implementado. Por fim, a seção 4.1.4 introduz nosso modelo de usuários.

4.1.2 Camada Física

Nesta camada representamos os seguintes mecanismos: o decaimento do sinal devido a propagação, controle de potência de malha aberta, controle de potência de malha fechada, a interferência intra-celular, interferência inter-celular, o desvanecimento e o controle de congestionamento. Para isso levamos em consideração canais de controle como o piloto, o DRC e o RA (Reverse Activity) *bit*.

Modelo de Propagação

O modelo de propagação é baseado em [10] e [11]. Em nosso trabalho, nós consideramos apenas o cenário urbano denso, e a descrição deste cenário, em dB, é dada pela Equação 4.1, considerando que a frequência a ser utilizada é de 450MHz. Os parâmetros representados nesta equação foram configurados de acordo com a Tabela 4.1.

$$a(h_r) = 3,2 (\log(11,75) \cdot h_r)^2 - 4,17 \quad 1 \leq h_r \leq 10m \quad (4.1)$$

Parâmetro	Valor
frequência da portadora	450MHz
altura do terminal	1.5m
altura da ERB (H_r)	40m
perda de penetração	10dB (0dB) para usuários móveis
desvanecimento (<i>flat fading</i> ¹)	v.a. lognormal, $\mu = 0dB$ e $\sigma = 8dB$

Tabela 4.1: Parâmetros para o modelo de Okumura Hata, cenário: Urbano Denso

Controle de Potência

No controle de potência de malha aberta, o cliente estima a potência que deve ser usada para iniciar a primeira transmissão segundo a seguinte equação:

$$P_{piloto} = P_{OH} + P_{pl} + P_{rs} + D \quad (4.2)$$

onde P_{piloto} é a potência que será utilizada para a transmissão do sinal piloto para o primeiro pacote, P_{OH} é a perda de propagação de acordo com o modelo de Hata-Okumura, P_{pl} é a perda de penetração, P_{rs} é a sensibilidade da ERB e D é o desvanecimento lento do sinal.

O valor encontrado para P_{piloto} é comparado com o valor da potência máxima (P_{max} - potência com a qual o sinal foi transmitido no emissor) subtraído do ganho para se alcançar a taxa desejada afim de se determinar se o cliente tem potência suficiente para transmitir nessa taxa, ou seja:

$$P_{piloto} > P_{max} - \delta_r \quad (4.3)$$

onde δ_r é o ganho necessário para se transmitir à taxa r .

Se a potência P_{piloto} que deve ser usada pelo cliente for maior do que a potência máxima disponível, δ_r é reduzido até que haja potência para transmissão ou até que δ_r alcance seu valor mínimo.

¹O desvanecimento incide com mesma intensidade sobre todas as componentes de frequência

Uma vez que a potência inicial de transmissão do cliente foi estimada, o controle de potência de malha fechada passa a operar. Este controle é utilizado para fazer um ajuste fino na potência do sinal piloto usada pelo cliente. O objetivo deste ajuste é acompanhar pequenas variações no canal.

É importante notar que a ERB monitora a potência recebida de todos os clientes. Esta monitoração é feita da seguinte maneira:

$$\begin{aligned}
 I_t &= FI \cdot \sum_i 10^{\frac{P_r(i)}{10}} \\
 P_n &= 10 \cdot \log_{10} (10^{-10.41} + I_t) \\
 A_{R_b} &= 10 \cdot \log_{10} \left[1 + 10 \cdot \left(\frac{DCG}{10} \right) \right] \\
 P_{rd} &= E_c/N_t[RRI(i)] + A_{R_b}[RRI(i)] + P_n \quad \forall \text{ cliente } i \quad (4.4)
 \end{aligned}$$

onde I_t é a interferência total, $P_r(i)$ é a potência recebida pelo cliente i , P_n é a potência do ruído de fundo, A_{R_b} corresponde ao ganho do canal de dados e tem valor diferente para cada taxa, $RRI(i)$ é a taxa utilizada pelo cliente i no enlace reverso, FI é o fator de interferência (causado pelas outras células) e P_{rd} é a potência desejada. No padrão atual, A_{R_b} e E_c/N_t são valores constantes, como mostrado nas tabelas 4.2 e 4.3 respectivamente.

Taxa (Kbps)	0	9.6	19.2	38.4	76.8	153.6
Dados	1.76	5.88	7.94	10.39	13.55	18.59
Ack	6.03	6.03	6.03	6.03	6.03	6.03
Dados+Ack	6.03	8.05	9.42	11.29	14.00	18.74

Tabela 4.2: Valores para A_{R_b} em função da taxa no canal reverso

Taxa (Kbps)	0	9.6	19.2	38.4	76.8	153.6
$\frac{E_c}{N_t}$	-21.0	-21.0	-21.5	-21.5	-22.0	-22.5

Tabela 4.3: Valores para $\frac{E_c}{N_t}$ em função da taxa no canal reverso

A ERB calcula I_t e compara com a potência recebida especificamente do cliente i . Se a potência desejada (P_{rd}) for maior do que a potência recebida pelo cliente

i ($P_r(i)$), então o cliente i deve aumentar a sua potência em $1dB$, caso contrário o cliente deve reduzir a sua potência em $1dB$. Este comando é enviado por um canal específico para o controle de potência.

Estimativa do DRC

Outra característica importante representada em nosso modelo é a estimativa do DRC, realizada pelos clientes. Para que a estimativa do DRC seja feita, é fundamental que se conheça a interferência no canal direto. Nesta subseção, começamos pelo modelo de interferência do canal direto e seguimos apresentando a estimativa do DRC.

A Figura 4.2 ilustra como as estações base de outras células interferem na transmissão para o usuário (canal direto). Nesta figura, são consideradas duas situações: (a) o usuário encontra-se na reta que liga o centro da célula a um vértice e (b) o usuário encontra-se na reta perpendicular à aresta que passa pelo centro da célula. Estas duas situações serão chamadas, respectivamente, de geometrias D e E.

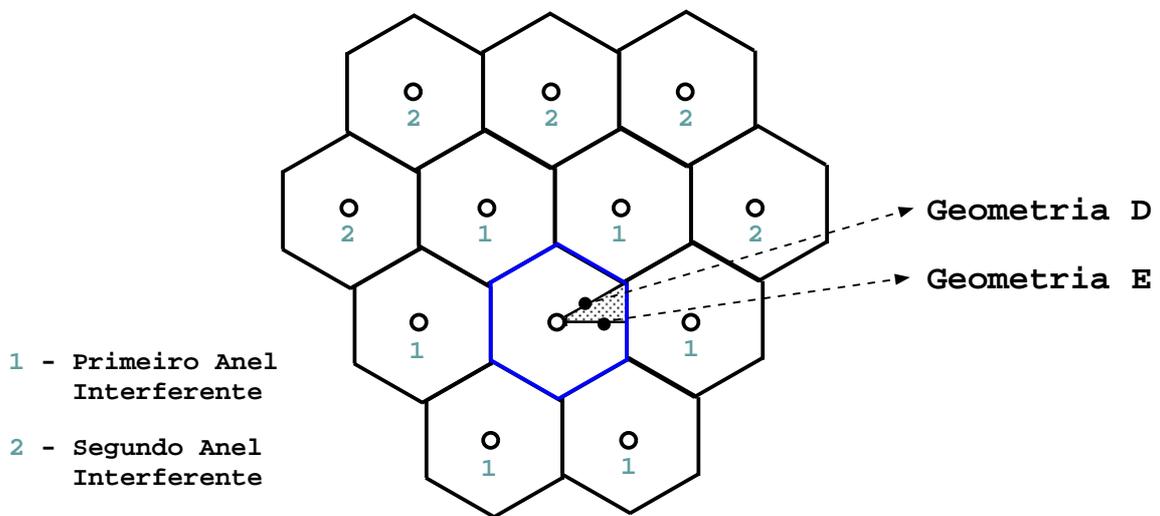


Figura 4.2: Modelo de geometrias entre células

Geometria D:

Seja I_1 a potência da interferência causada pelo primeiro anel interferente e I_2 a potência da interferência causada pelo segundo anel. A interferência total I_{BS} é igual a

$$I_{BS} = I_1 + I_2. \quad (4.5)$$

Pela figura, é possível observar que existe simetria acima e abaixo da reta que liga o centro da célula ao vértice. Neste cenário, I_1 pode ser escrito como:

$$I_1 = 2(A_D + B_D + C_D) \quad (4.6)$$

onde A_D , B_D e C_D representam as potências de interferência das células adjacentes a célula em questão e sendo iguais a, respectivamente, $10^{\frac{A'_D}{10}}$, $10^{\frac{B'_D}{10}}$ e $10^{\frac{C'_D}{10}}$,

$$\begin{aligned} A'_D &= f\left(R\sqrt{3+q^2}\right) \\ B'_D &= f\left(R\sqrt{3+q^2+3q}\right) \\ C'_D &= f\left(R\sqrt{3+q^2-3q}\right) \end{aligned} \quad (4.7)$$

onde R é o raio da célula e q é a distância normalizada (pelo raio) do cliente para a ERB.

Já a interferência total devido ao segundo anel interferente é dada por:

$$I_2 = 2(B_D + C_D + D_D + E_D + F_D) + A_D + G_D \quad (4.8)$$

onde,

$$A_D = 10^{\frac{A'_D}{10}}, B_D = 10^{\frac{B'_D}{10}}, C_D = 10^{\frac{C'_D}{10}}, D_D = 10^{\frac{D'_D}{10}}, E_D = 10^{\frac{E'_D}{10}}, F_D = 10^{\frac{F'_D}{10}} \text{ e } G_D = 10^{\frac{G'_D}{10}} \text{ com}$$

$$\begin{aligned}
A'_D &= f(R(3-q)) \\
B'_D &= f\left(R\sqrt{12+q^2}\right) \\
C'_D &= f\left(R\sqrt{12+q^2-3q}\right) \\
D'_D &= f\left(R\sqrt{9+q^2-3q}\right) \\
E'_D &= f\left(R\sqrt{9+q^2+3q}\right) \\
F'_D &= f\left(R\sqrt{12+q^2+3q}\right) \\
G'_D &= f(R(3+q))
\end{aligned}
\tag{4.9}$$

Geometria E: Similar à geometria D, a geometria E também foi dividida em dois anéis interferentes I_1 e I_2 .

$$I_1 = A_E + B_E + 2C_E + 2D_E \tag{4.10}$$

$$\tag{4.11}$$

com:

$$A_E = 10^{\frac{A'_E}{10}} \quad B_E = 10^{\frac{B'_E}{10}}; \quad C_E = 10^{\frac{C'_E}{10}}; \quad D_E = 10^{\frac{D'_E}{10}};$$

$$\begin{aligned}
A'_E &= f\left[R\left(\sqrt{3}-q\right)\right] \\
B'_E &= f\left[R\left(\sqrt{3}+q\right)\right] \\
C'_E &= f\left(R\sqrt{3+q^2-\sqrt{3}q}\right) \\
D'_E &= f\left(R\sqrt{3+q^2+\sqrt{3}q}\right).
\end{aligned}
\tag{4.12}$$

Já a interferência total devido ao segundo anel interferente é dada por:

$$I_2 = 2(E_E + F_E + G_E + H_E + J_E) + K_E + L_E, \tag{4.13}$$

onde

$$\begin{aligned}
E_E &= 10^{\frac{E'_E}{10}} \\
F_E &= 10^{\frac{F'_E}{10}} \\
G_E &= 10^{\frac{G'_E}{10}} \\
H_E &= 10^{\frac{H'_E}{10}} \\
J_E &= 10^{\frac{J'_E}{10}} \\
K_E &= 10^{\frac{K'_E}{10}} \\
L_E &= 10^{\frac{L'_E}{10}},
\end{aligned}
\tag{4.14}$$

com:

$$\begin{aligned}
K'_E &= f \left[R \left(2\sqrt{3} - q \right) \right] \\
L'_E &= f \left[R \left(2\sqrt{3} + q \right) \right] \\
E'_E &= f \left(R \sqrt{12 + q^2 - 3q\sqrt{3}} \right) \\
F'_E &= f \left(R \sqrt{12 + q^2 + 3q\sqrt{3}} \right) \\
G'_E &= f \left(R \sqrt{9 + q^2} \right) \\
H'_E &= f \left(R \sqrt{9 + q^2 - 3q\sqrt{3}} \right) \\
J'_E &= f \left(R \sqrt{9 + q^2 + 3q\sqrt{3}} \right)
\end{aligned}
\tag{4.15}$$

As demais situações de interferência (I_t), exemplificadas pela área sombreada da Figura 4.2, serão consideradas como combinações lineares das geometrias D (geo_d) e E (geo_e). Note que em nossa abordagem foi considerado que a potência de transmissão e a área de cobertura de todas as células são iguais.

Como apresentado pelo conjunto de Equações 4.16, o valor de DRC escolhido é obtido de acordo com o valor da razão SINR (do inglês *Signal to Interference and Noise Ratio*)($sinr$) entre a potência (f) recebida pelo terminal de sua estação base, e a potência interferente (I_t) calculada previamente por este método adicionado ao

ruído térmico (i_0). Os valores de limiares de SINR para determinação do DRC estão reproduzidos na Tabela 4.4.

$$\begin{aligned}
 I_t &= i_0 + g_f \cdot geo_d + (1.0 - g_f) \cdot geo_e; \\
 f &= p_{erb} - a[h_r] - p_{loss} - p_{desv}; \\
 sinr &= 10 \log_{10} \left(\frac{10^{\frac{f}{10.0}}}{I_t} \right),
 \end{aligned}
 \tag{4.16}$$

onde g_f representa a posição relativa do cliente entre as retas geo_d e geo_e , p_{erb} é a potência recebida da ERB da célula à qual o cliente se localiza, p_{loss} é a perda de potência por penetração e p_{desv} é a perda de potência por desvanecimento.

DRC	1	2	3	4	5	6	7	8	9	10	11	12
Taxa (Kbps)	38	76	153	307	307	614	614	921	1228	1228	1800	2400
SINR (dB)	-11.5	-9.2	-6.5	-3.5	-3.5	-0.6	-0.5	2.2	3.9	4.0	8.0	10.3

Tabela 4.4: Valor da SINR necessária para transmissão em cada uma das taxas possíveis no enlace direto para uma PER de 1%.

Controle de Congestionamento

A ERB monitora a potência total recebida no enlace reverso. Se este valor ultrapassar um limite, a estação avisa os terminais que o controle de congestionamento deve ser ativado. De maneira análoga, se essa potência estiver abaixo desse limite, a estação avisa os terminais que eles devem desabilitar o controle de congestionamento.

Estes comandos são enviadas em todos os *slots* do canal direto. Em ambos os casos as alterações são feitas de forma probabilística, ou seja, cada terminal TA_i , transiciona para a próxima taxa acima com probabilidade $p_{i,r}$ e decresce para a taxa logo abaixo com probabilidade $q_{i,r}$, conforme o comando recebido e conforme a taxa r atual.

Os valores de $p_{i,r}$ e $q_{i,r}$ [49] dependem da taxa de *bits* que o TA está transmitindo naquele instante e estão listados na Tabela 4.5. Porém a forma como o *bit* RA é calculado pela ERB não faz parte do padrão.

Neste trabalho foi usado o conceito de elevação do nível de ruído (*noise rise*) como forma de medir o nível de congestionamento.

Para entender este conceito, suponhamos que R_{pot} seja a razão entre a potência interferente total no setor e a potência de ruído térmico, ou seja:

$$R_{pot} = \frac{N_i W}{i_0} ,$$

onde o carregamento no setor NtW pode ser expresso por: $NtW = NW + I_t$. Neste trabalho foi utilizado como limiar de sobrecarga o valor de $R_{pot} = 5$ [49]. Quando este valor é ultrapassado a ERB ativa o controle de congestionamento.

(r)	0	1	2	3	4	5
Taxa (Kbps)	0	9.6	19.2	38.4	76.8	153.6
p	1	3/16	1/16	1/32	1/32	0
q	0	0	1/16	1/16	1/8	1

Tabela 4.5: Probabilidade de transição para a taxa adjacente caso o controle de congestionamento não esteja (p) ou esteja (q) ativo

4.1.3 Camada de enlace

Algoritmo de escalonamento de pacotes

A principal funcionalidade implementada na camada de enlace é o algoritmo de escalonamento de pacotes conhecido como *Proportional Fair Scheduling* (PFS). O principal objetivo deste algoritmo é dar alguma prioridade a usuários que percebam condições mais favoráveis em seus canais ao mesmo tempo que procura impedir que usuários em condições inferiores tenham suas transmissões estagnadas.

Para que este objetivo seja alcançado, são usados conceitos *cross-layer*. Em outras palavras, mesmo fazendo parte da camada de enlace, o PFS lê valores tais como o SINR que, de maneira geral, estão disponíveis apenas à camada física.

Para facilitar a compreensão deste algoritmo, dividimos o PFS em dois passos. No primeiro passo o escalonador seleciona o usuário j , que é aquele que possua a maior relação $\frac{DRC_i(t)}{R_i(t)}$:

$$j = \underset{i}{\operatorname{arg\,max}} \left[\frac{DRC_i(t)}{R_i(t)} \right] \quad (4.17)$$

onde $DRC_i(t)$ é a mais recente estimativa de DRC feita pelo usuário i e $R_i(t)$ é a taxa média à qual este usuário vem recebendo *bits*.

No segundo passo, a taxa média do usuário i (para todo i) é atualizada como a seguir (média ponderada):

$$R_i(t+1) = (1 - \alpha) \cdot R_i(t) + \alpha \cdot CRT_i(t) \quad (4.18)$$

$$\begin{cases} CRT_i(t) = DRC_i(t) & \text{se } i = j \\ CRT_i(t) = 0 & \text{se } i \neq j \end{cases} \quad (4.19)$$

onde $CRT_i(t)$ é a taxa atual de *bits* do usuário i no instante t .

Cada terminal realiza estimativas do DRC ($DRC_i(t)$) de acordo com o SINR observado, por ele, no canal direto. Este valor é enviado à ERB e consiste na informação da camada física usada pelo PFS. A Tabela 4.4 apresenta o SINR mínimo necessário para cada uma das 12 taxas possíveis para o DRC.

Como pode ser visto na Equação 4.18, o parâmetro α ajusta o balanceamento existente entre a vazão agregada e a justiça. Por exemplo, quando $\alpha = 0$, R_i é atualizado com $1 \cdot R_i(t) + 0 \cdot CRT_i(t)$, ou seja, a taxa (CRT_i) nunca altera a média R_i . Com isso, a relação $\frac{DRC_i(t)}{R_i(t)}$ dependerá apenas da condição atual de cada usuário. Logo, o usuário com melhor condições terá prioridade máxima. Isso equivale à priorizar a vazão agregada.

Por outro lado, se $\alpha = 1$, R_i é atualizado com $0 \cdot R_i(t) + 1 \cdot CRT_i(t)$. Com isso, o último usuário a ser servido tem prioridade mínima, e outros usuários, mesmo que estejam em estado de baixa qualidade de canal, obterão maior relação $\frac{DRC_i(t)}{R_i(t)}$ sendo, assim, priorizados. Isso equivale a priorizar a justiça, em detrimento da vazão agregada.

Como será mostrado em nossos experimentos, este parâmetro é capaz de prover algum nível de justiça apenas em nível de *slots* alocados por usuários, o que

não necessariamente resulta em equidade em nível de usuário (e.g. vazão média individual).

4.1.4 Modelo dos Usuários

Para facilitar o entendimento do modelo de usuários, apresentamos a Figura 4.3 na qual é mostrado, em alto nível, a descrição de nosso modelo.

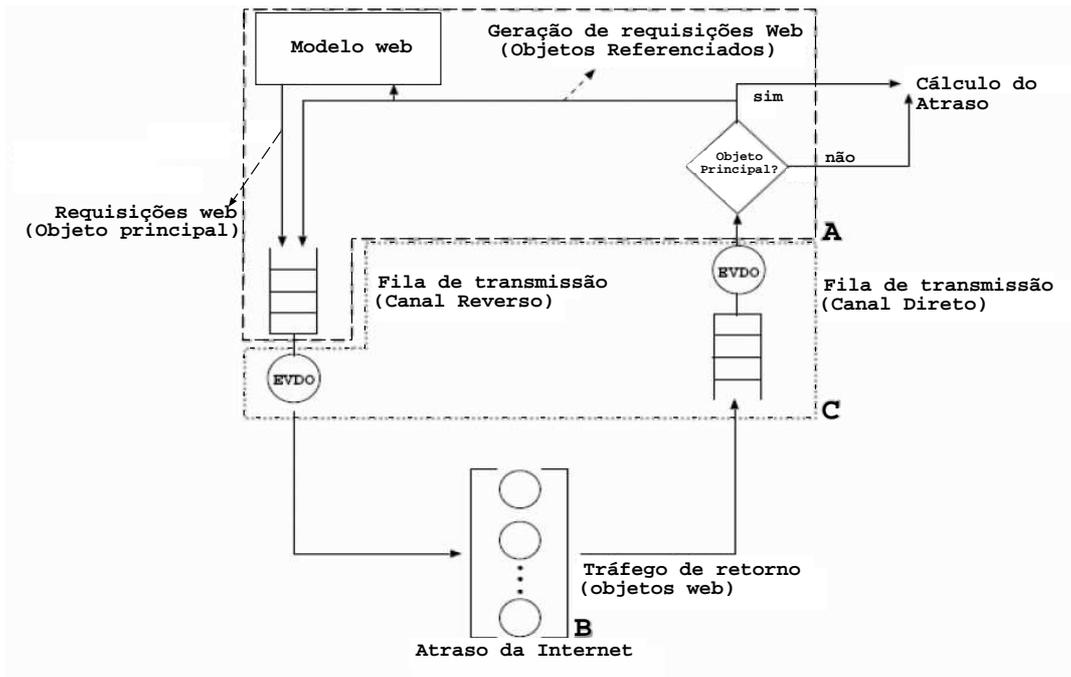


Figura 4.3: Visão Geral do modelo dos usuários

A figura é dividida em seções. A seção *A* representa o modelo de usuário *web*, a seção *B* representa o atraso sofrido pelos pacotes na Internet, e a seção *C* modela o protocolo EVDO, sendo mostrado no lado esquerdo da figura e o canal direto, no lado oposto.

Quando um usuário gera uma requisição *web*, o pacote referente à requisição espera na fila de transmissão do canal reverso até que haja oportunidade de transmissão. No canal reverso implementamos os seguintes subcanais: *pilot*, RRI, DRC, dados e subcanal de confirmação. A implementação destes canais é baseada em [5], a qual é amplamente usada pela literatura disponível [101, 102].

As requisições *web* eventualmente são encaminhadas para a Internet (Objeto Internet), onde sofrerá o atraso relativo ao RTT (*Round Trip Time*) o que representa

o tempo de viagem até o servidor *web* e a respectiva volta dos objetos *web* solicitados até a ERB, onde aguardarão na fila para serem transmitidos pelo canal direto (PFS).

Após descrever em linhas gerais o modelo, detalharemos o modelo de usuários *web* implementado.

Para o modelo de usuários proposto, assumimos que não há entrada e saída de usuários no sistema e que o usuário em nenhum momento transmite/recebe dados para/de mais de uma única antena. O modelo do usuário *web* foi construído com base em [103] e está ilustrado no lado esquerdo da Figura 4.3.

Quando um usuário *web* clica em uma ligação de hipertexto, várias requisições *web* podem se seguir. A primeira transmite a requisição do usuário ao servidor, solicitando a página principal. Quando os pacotes contendo a página principal chegam por completo ao cliente, são automaticamente geradas solicitações subsequentes de forma a obter, por exemplo, imagens, arquivos de sons ou outras páginas referenciadas. Modelamos o número de objetos referenciados por página *web* através de uma distribuição Pareto truncada para o valor máximo de 53, como mostrado na Tabela 4.6 [103, 104, 105].

Após ter emitido todas as requisições relativas a uma página *web*, e após ter recebido todo o conteúdo solicitado, e a página estiver completamente disponível na tela, o usuário passa um intervalo de tempo conferindo o resultado na tela. Usamos uma distribuição exponencial com $\lambda = 0.033190$, para representar esse intervalo de tempo de acordo com [106]. Após este intervalo, o usuário clica em uma nova ligação de hipertexto, iniciando novamente o processo de recuperação de uma página *web*.

Outras características importantes para a modelagem de tráfego *web* são: (1) o número de objetos referenciados por cada objeto *web* principal, (2) o tamanho, em *bytes*, dos objetos requisitados e (3) o tamanho das próprias requisições *web*. A Tabela 4.6 apresenta os valores considerados para estes parâmetros, de acordo com [103, 104, 105].

O atraso da Internet (RTT) é modelado da seguinte maneira: ao ser gerado, o pacote segue para uma fila do tipo “infinitos servidores” (do inglês *infinite server queue*) onde espera a sua vez para ser transmitido ao objeto Internet. Desta maneira simulamos o atraso gerado pela Internet (RTT) desde o momento que um pacote é transmitido até o momento em que a respectiva resposta é recebida pela ERB. Os

Parâmetro	Valor
Número de objetos referenciados por cada objeto principal	v.a Pareto truncada $k = 2, \alpha = 1.1, max = 53$
Tamanho do objeto principal	v.a. Lognormal $\mu = 10710, \sigma = 25032$ $min = 100, max = 2097152$
Tamanho dos objetos referenciados	v.a. Lognormal $\mu = 7758, \sigma = 126168$ $min = 50, max = 2097152$
Tamanho da requisição <i>web</i>	200 bytes
Tempo de leitura da página (<i>OFF</i>)	v.a. Exponencial, $\lambda = 0.033190$

Tabela 4.6: Parâmetros para o modelo de tráfego *web*

parâmetros considerados [107] para o RTT são $\mu = 0.32935$ e $\sigma = 0.0064513$ para uma variável aleatória Gaussiana.

4.1.5 Modelo de Mobilidade

O modelo usado neste trabalho para representar a mobilidade de usuários é o modelo *Random Waypoint* (RWP) [108]. A razão para esta escolha é que o RWP é comumente [109, 110, 111] usado como modelo sintético para mobilidade de usuários. A ideia geral do RWP é descrever de maneira simples o padrão de movimento de usuários independentes dentro de uma área definida.

Além disso, embora o RWP tenha sido proposto há vários anos e ser um dos modelos mais usados pelos engenheiros de redes, o trabalho de Le Boudec et. al. [14] recentemente trouxe novas contribuições. Resumidamente, os autores definem maneiras simples de se obter distribuição estacionária de nós e de velocidades na área considerada. Adicionalmente, os autores contribuem com um algoritmo para ser usado na fase inicial de experimentos de simulação, de modo que o experimento já inicie no estado estacionário, no que diz respeito ao posicionamento e às velocidades

dos usuários. O nome dado a esta capacidade é *Simulação Perfeita*.

O modelo RWP de [14] possui as seguintes características:

- cada nó se move através de uma linha reta de um *waypoint* (coordenada) a outro;
- os *waypoints* são uniformemente distribuídos sobre a área convexa, por exemplo, a célula.
- quando um nó chega ao *waypoint*, um novo ponto de destino é aleatoriamente sorteado;
- embora o modelo permita aleatoriedade na escolha da velocidade inicial, nosso modelo supõe a velocidade constante de 3km/h, típica de cenários pedestres de mobilidade.
- em nosso cenário, consideramos que o *thinking time*, termo usado para os instantes em que um nó permanece sem se movimentar antes de partir para o próximo *waypoint*, é nulo (i.e. igual a 0).

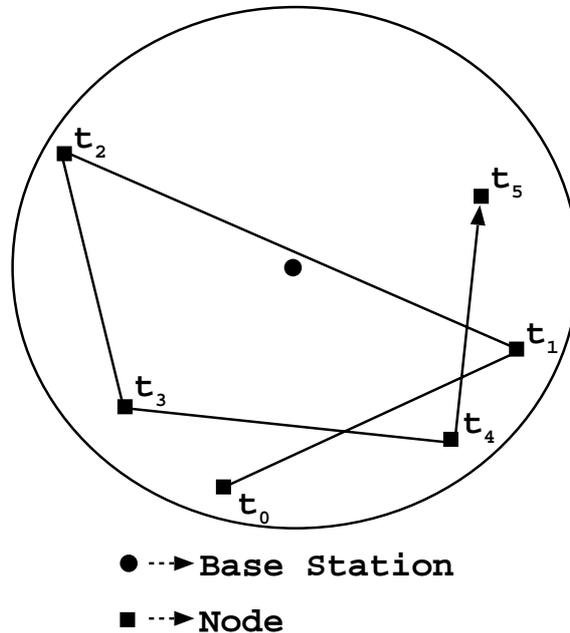


Figura 4.4: Visão geral do modelo *Random Waypoint*

Como exemplo, considere a Figura 4.4. Quando o nó alcança o *waypoint* t_1 , o *waypoint* t_2 é escolhido aleatoriamente. Imediatamente o nó passa a caminhar em direção a t_2 à velocidade escolhida (3km/h em nosso cenário).

Em nosso caso, optamos por atualizar a posição do nó uma vez a cada segundo. Desta maneira, a posição do nó é constantemente atualizada mesmo que ele não tenha alcançado o destino.

O RWP, como mostrado em [14], não alcança regime de estado estacionário em alguns cenários devido a fatores tais como a mudança ou o decaimento da velocidade à medida que a simulação progride.

Para resolver este impasse, os autores de [14] propuseram a ideia de *Simulação Perfeita*, onde o estado inicial é aleatoriamente amostrado do conjunto de estados de regime estacionário. Como nós usamos este procedimento em nossos experimentos, apresentamos o algoritmo a seguir.

Seja A o domínio no qual os nós estão contidos, δ o limite superior do diâmetro de A , X_t a posição do nó no instante t e V a velocidade

1. Sortear X_{t-1} e X_{t+1} , duas variáveis aleatórias uniforme $U(0, \delta)$ independentes,
2. gerar V , uma variável aleatória uniforme $U(V_{\min}, V_{\max})$,
3. repetir os passos 1 e 2 até que $V < \|X_{t+1} - X_{t-1}\|$,
4. gerar uma variável aleatória ($u: U(0, 1)$),
5. $X_t = (1 - u)X_{t-1} + uX_{t+1}$

Este algoritmo de amostragem garante a estacionaridade do RWP e tem outra vantagem: não há necessidade de se descartar o período inicial da simulação, já que se encontra em estado estacionário desde o início.

Em nossos experimentos, o início dos experimentos de simulação são descartados devido ao período transiente do modelo EVDO unido ao modelo de tráfego. A adição do modelo de mobilidade não trouxe necessidade de aumento no período inicial de simulação que é descartado.

4.2 Experimentos e resultados

Nesta seção apresentamos vários experimentos realizados com o modelo proposto.

Os resultados dos experimentos de simulação foram obtidos utilizando-se a ferramenta de simulação do ambiente Tangram-II [112, 113] disponível para computadores pessoais que rodam o sistema operacional Linux. Esta é uma ferramenta de código aberto, onde a descrição do modelo é facilitada por uma linguagem de alto nível baseada na linguagem C.

Como essa ferramenta foi desenvolvida e é mantida em nosso laboratório, há maior facilidade (e também interesse) em implementarmos novas funcionalidades necessárias para o desenvolvimento dos modelos propostos nesta tese.

A seguir, listamos o conjunto de parâmetros usados para configurar nosso modelo. Em seguida, apresentamos e discutimos cada experimento para tentar responder as questões levantadas no Capítulo 1: As adaptações da camada física e da camada de enlace empregadas em redes sem fio de longa distância são capazes de manter a qualidade de serviço e a justiça a um nível satisfatório?

4.2.1 Cenários de experimentos e resultados obtidos para o modelo EVDO

O primeiro grupo de experimentos tem o objetivo de determinar o número de usuários que uma célula pode servir mantendo uma qualidade mínima, *do ponto de vista do usuário*, em função de um conjunto de parâmetros. Por fim, apresentamos os resultados obtidos com as respectivas discussões e respostas para as questões levantadas.

Dividimos os experimentos em **6** estágios, de acordo com o desenvolvimento do trabalho:

- primeiramente propomos um modelo que representa uma célula e que detalha características das camadas físicas e de enlace.
- em seguida agregamos, a este sistema, modelos de usuários estáticos com o intuito de gerar carga suficiente para estressar o sistema.
- a terceira parte destina-se a resolver o modelo para determinar o número de usuários que uma célula pode servir mantendo uma qualidade mínima, *do ponto de vista do usuário*, que será determinada em função de um conjunto de

parâmetros, bem como estudar raio máximo de alcance da antena em diversas condições de interferência.

- a quarta parte consiste em investigar a equidade, ou seja, a diferença de qualidade de serviço entre usuários em função de sua localização em relação ao centro da célula.
- na quinta parte, adicionamos mobilidade aos usuários do modelo proposto no intuito de avaliar o seu efeito sobre a equidade;
- em seguida, na sexta parte, reunimos as limitações encontradas, do ponto de vista da qualidade de serviço, e propomos uma solução simples para contornar tais problemas;

Descrição	Valor
População de usuários	10-80 usuários
Raio da célula	1 km
Potência máxima de transmissão do TA	23 dBm
Potência máxima de transmissão da ERB	55.8 dBm
Perda de penetração para usuários estáticos	10 dB
Ruido Térmico	-165 dB
Parâmetro α do PFS	0.001
<i>Cable Loss</i>	3 dB
Ganho na antena da ERB	17 dB
Sensibilidade do receptor (Cliente)	-119 dBm
Velocidade dos usuários móveis	3 Kmh

Tabela 4.7: Valores configurados para os experimentos realizados

Os seguintes parâmetros foram usados nos experimentos realizados. A interferência inter-celular foi configurada para 40% da interferência intra-celular; os usuários são aleatoriamente posicionados na célula; o algoritmo de controle de congestionamento foi configurado com p e q de acordo com [49] (ver [114, 71] para definição de p e q); foi usado grau de confiança de 95% para o cálculo do intervalo

de confiança e, para todos os resultados obtidos, o intervalo de confiança foi de 19%¹ (ou menos) dos valores médios obtidos. Note que optamos por não mostrar os intervalos de confiança em nossas figuras no intuito de facilitar sua visualização; os outros parâmetros do modelo são apresentados na Tabela 4.7.

4.2.2 Experimentos realizados e resultados obtidos

Cenário sem mobilidade

No primeiro conjunto de experimentos calculamos o *goodput* e o atraso médio. Nós definimos o *goodput* como sendo a razão L/d , onde L representa o tamanho médio de uma página *web* completa e d é o atraso médio computado desde o instante em que o usuário aponta em uma hiper-ligação até o momento em que a página é completamente recebida. Dois cenários foram considerados para o primeiro conjunto de experimentos: (1) não há mobilidade de usuários e (2) 30% dos usuários estão se movimentando pela célula.

Definimos as seguintes métricas apresentadas nas Figuras 4.5 e 4.6. Seja g_i o *goodput* médio do usuário obtido em uma rodada de simulação. Então:

$$\text{Média} = \frac{\sum_{i=1}^r g_i}{r},$$

onde r é o número de rodadas da simulação. A Figura 4.5 mostra os valores mínimo e o máximo de g_i , ou seja,

$$\text{Mínimo} = \min [g_i], \forall i, e$$

$$\text{Máximo} = \max [g_i], \forall i.$$

Como esperado, o *goodput* tende a diminuir à medida que a população cresce. Isso ocorre devido ao compartilhamento de recursos. Note que, para as populações consideradas, a diferença entre o menor e o maior valor de *goodput* observados varia de 7 a 25% em torno da média.

¹Este intervalo de confiança é alto, quando comparado aos outros resultados obtidos nesta tese, pois o modelo de usuário *web* envolve distribuições de cauda longa. Ressaltamos que este intervalo de confiança não altera nossas conclusões.

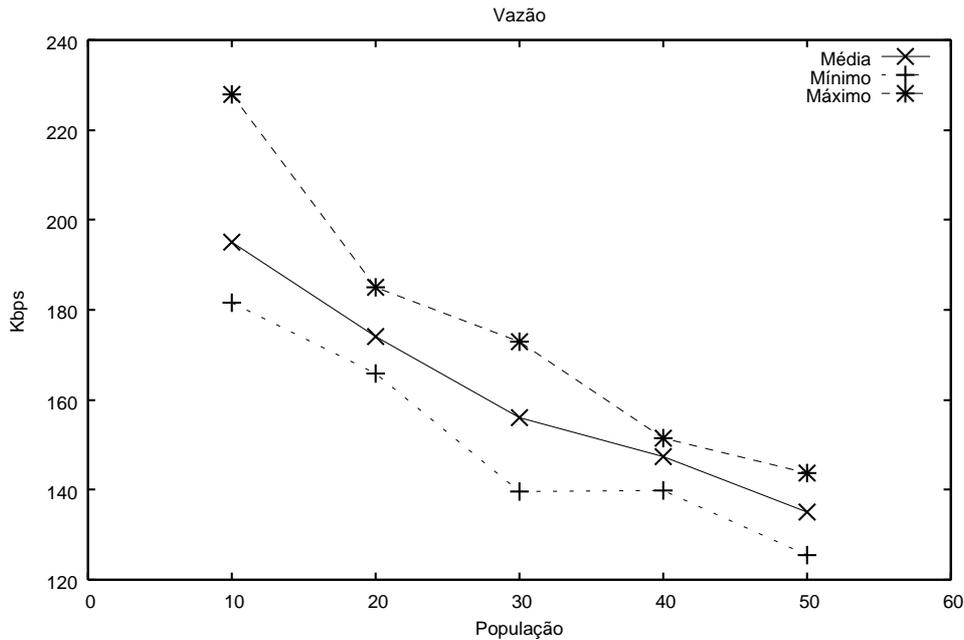


Figura 4.5: *Goodput* médio em função da população

O atraso médio em função da população de usuários é mostrado na Figura 4.6, e o *goodput* é mostrado na Figura 4.5. Como esperado, o *goodput* diminui e o atraso cresce com o aumento da população. Note que o erro relativo do *goodput* é bem menor que o do atraso. Este também é um comportamento esperado, e ocorre inclusive em modelos M/M/1. No caso do modelo proposto, é importante enfatizar que a variância do atraso é ainda maior por causa do uso de distribuições de cauda longa. A diferença entre o menor e o maior valor é, na maioria dos casos, da mesma ordem de grandeza que o atraso médio para os casos de maior população, apesar da diferença relativa ser bem maior quando comparada à do *goodput*. Isto ocorre principalmente devido ao uso de distribuições de cauda longa na modelagem do tamanho dos objetos *web*. Para uma discussão mais aprofundada sobre a auto-similaridade do tráfego *web*, referenciamos o trabalho de Crovella [105].

A seguir apresentamos o segundo conjunto de experimentos onde temos o objetivo de avaliar a justiça em relação ao atraso e *goodput*. Para estes experimentos, o raio da célula foi dividido em dez anéis igualmente espaçados, conforme mostrado na Figura 4.13, sendo o anel 1 o mais próximo à ERB e o anel 10 o mais distante. Note que ainda não consideramos mobilidade de usuários.

Nas Figuras 4.7 e 4.8 podemos observar a variação do *goodput* e atraso médios em função da distância em relação à estação base e da população. É possível observar

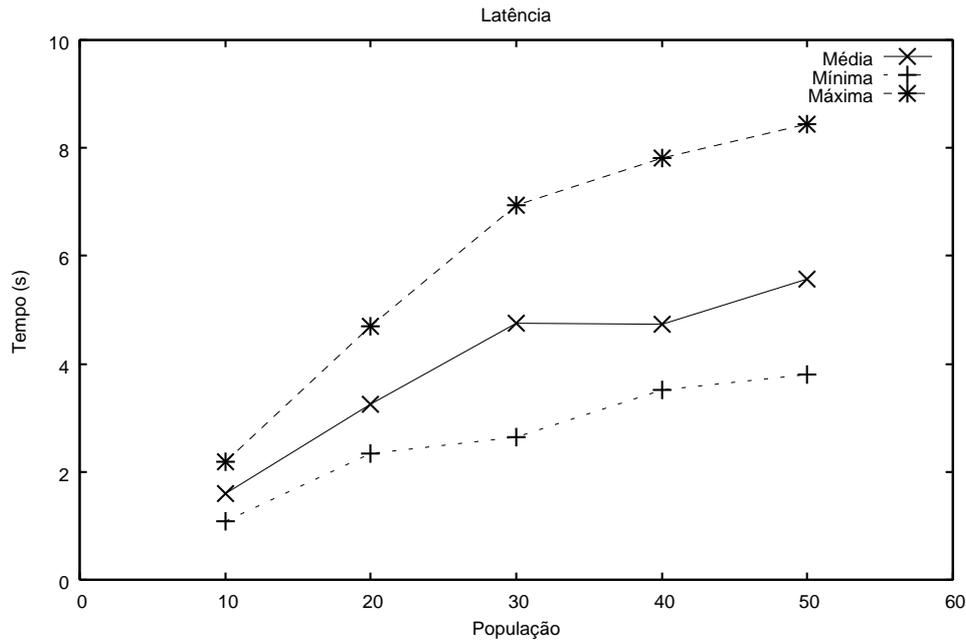


Figura 4.6: Atraso médio em função da população

que o *goodput* no anel 10 é 20% do valor encontrado no anel 1, quando a população foi configurada para 80 usuários. O atraso apresenta resultado análogo. Há um aumento de 3s (anel 1) para 50s (anel 10) indicando um significativo decréscimo da QoS à medida que a distância em relação à ERB aumenta.

Cenário misto de usuários sem e com mobilidade

No próximo grupo de experimentos consideramos que um subconjunto da população se move. Neste novo cenário, chamado *Cenário 2*, 30% dos usuários são móveis (Note que, no *Cenário 1* não havia mobilidade de usuários). As Figuras 4.9 e 4.10 comparam estes dois grupos de experimentos (cenários 1 e 2).

Durante estes experimentos, monitoramos a carga do sistema. Embora tenhamos optado por não incluir uma figura apresentando a carga do sistema em função da população considerada, é importante compreendermos as situações de carga para discutirmos a diferença observada na qualidade de serviço.

Monitorando-se a carga do sistema observou-se que, para os casos de 10 a 40 usuários, o sistema está com carga baixa. Com 60 usuários a carga do sistema é de aproximadamente 84%, ou seja, o sistema está fortemente carregado. Quando a população é configurada para 80 usuários a carga se aproxima de 95% (o sistema está sobrecarregado). Nós definimos a *carga* do sistema como sendo a fração de *slots*

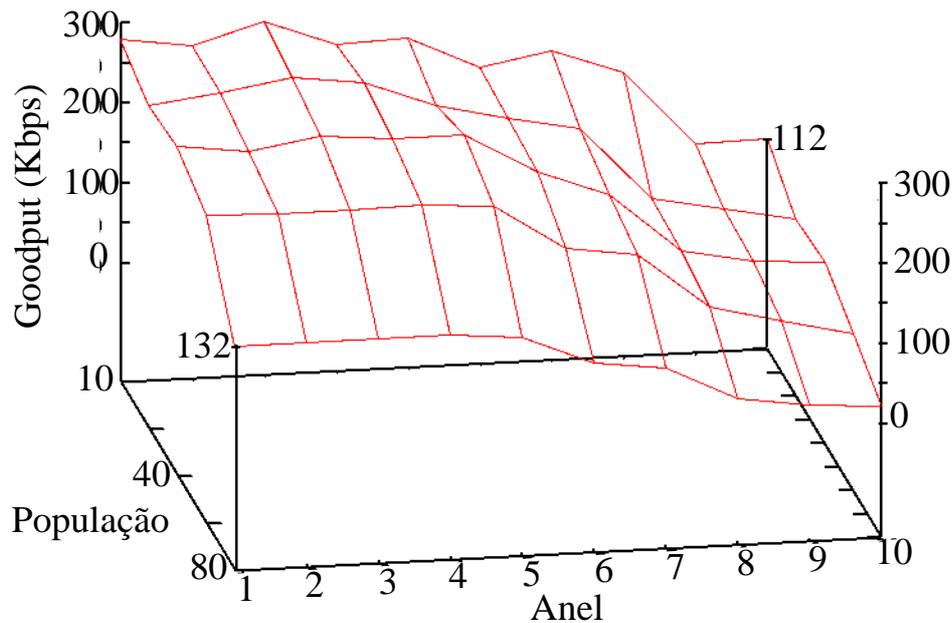


Figura 4.7: *Goodput* em função da população por anel de distância

do canal direto ocupados com transmissão de dados.

Para melhor compreensão das Figuras 4.9 e 4.10 considere os três casos acima descritos: baixa carga, forte carga e sobrecarga do sistema.

Quando o sistema está com baixa carga, o número de *slots* ociosos é relativamente alto. Portanto, quando um usuário móvel encontra boas condições no canal, é alta a probabilidade deste usuário encontrar um *slot* livre para transmissão. Por isso, a presença de usuários móveis no sistema não afeta significativamente o atraso dos usuários estáticos.

Quando o sistema está fortemente carregado, há um número relativamente baixo de *slots* ociosos. Assim, quando o cenário é trocado e passa-se a considerar usuários móveis, a contenção significativamente aumentará. Como uma fração dos usuários móveis encontrar-se-á próximos à ERB, os usuários deste subconjunto perceberão, com maior probabilidade, melhores condições do canal e serão priorizados pelo algoritmo de escalonamento de pacotes. Como consequência, os usuários mais distantes são severamente prejudicados. Isso pode ser observado no caso onde a população foi configurada para 60 usuários, nas Figuras 4.9 e 4.10.

No caso onde o sistema se encontra sobrecarregado, há maior demanda do que capacidade e surge o problema de insuficiência de *slots* para servir aos usuários. Por isso, tanto o cenário que considera apenas usuários estáticos quanto o cenário que

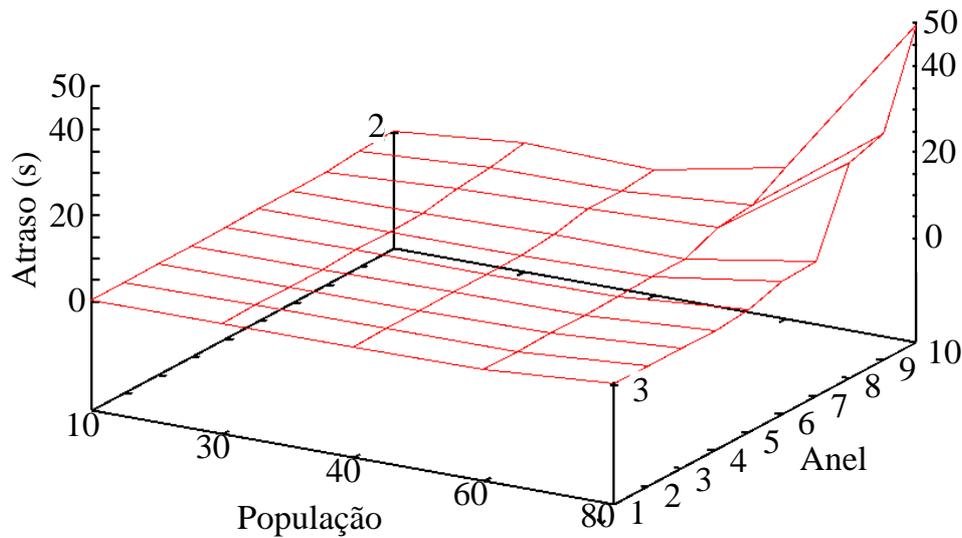


Figura 4.8: Atraso em função da população e do anel de distância em relação à ERB

considera móvel uma fração da população sofre de séria deterioração do atraso e do *goodput*.

Os casos de baixa carga e de sobrecarga são desinteressantes, pois estão fora do nível de operação tipicamente projetado para sistemas de computação. Por isso focaremos no caso onde o sistema está fortemente carregado.

As Figuras 4.11 e 4.12 apresentam resultados mais detalhados para o caso de 60 usuários. Nessas figuras, o atraso e o *goodput* são apresentados para os cenários 1 e 2. Se compararmos o cenário 1 (sem mobilidade) com o cenário 2 (30% da população é considerada móvel), será possível notar que o atraso médio cresceu de 2 para aproximadamente 8 segundos.

Um comportamento análogo é mostrado na Figura 4.12, onde o *goodput* médio foi de 175Kbps no cenário 1 e 125Kbps no cenário 2. Conseqüentemente, é fácil ver que a introdução de usuários móveis diminui significativamente o *goodput* e aumenta o atraso.

Em seguida, continuando o foco sobre um sistema fortemente carregado (60 usuários) consideramos um novo grupo de experimentos. O objetivo é mostrar que a degradação da qualidade de serviço, observada quando comparamos os cenários 1

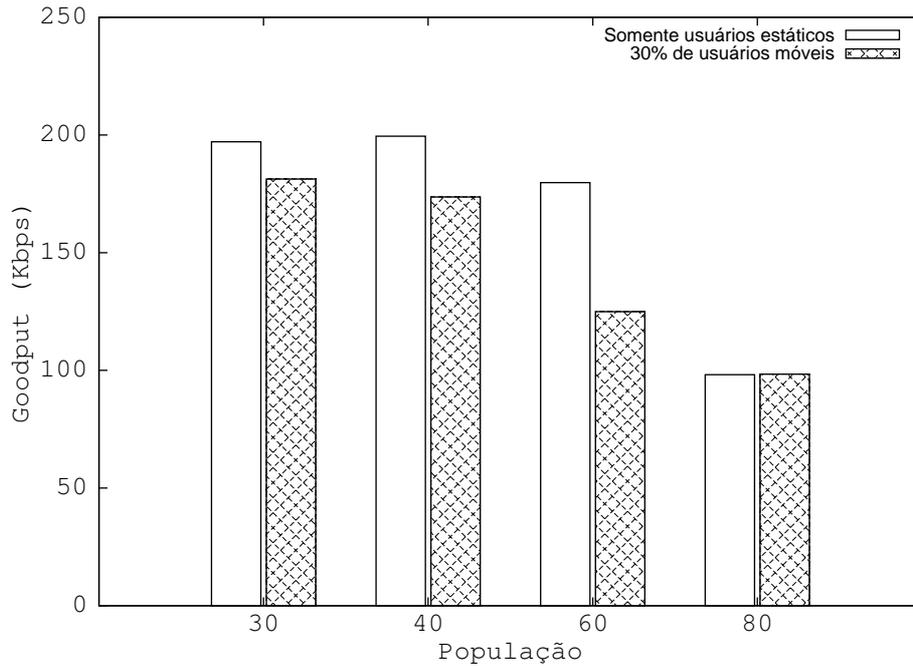


Figura 4.9: *Goodput* médio percebido pelo usuário em função da população considerada (usuários móveis e estáticos vs. estáticos)

e 2, é acompanhada de séria degradação da justiça em nível de usuário, tanto em relação ao atraso quando ao *goodput*. Note que, para os cenários que consideram mobilidade, há um anel adicional que chamamos de *mob.* o qual agrega os resultados para o grupo de usuários com mobilidade.

Além disso, desejamos analisar o efeito da variação da população e da inserção de usuários móveis sobre a justiça em nível de usuário. Para isso consideramos, para fins de comparação, um cenário onde apenas um usuário estático se encontra presente. O objetivo deste cenário é estabelecer um valor de referência para as métricas que buscamos: *goodput* e atraso, para cada anel de distância, quando não há contenção. Mais adiante compararemos estes valores de referência com outros experimentos no intuito de avaliar alterações na QoS e na justiça em nível de usuário.

Os resultados deste experimento mostraram que quando apenas um usuário estático se encontra presente no primeiro anel de distância, seu atraso médio é de 0.68s. Caso este mesmo usuário seja posicionado nos anéis 8, 9 e 10, seu atraso médio cresce para 2.1s, 2.4s e 3.6s, respectivamente. Note que o atraso médio de um único usuário no último anel de distância é aproximadamente cinco vezes maior do que um único usuário posicionado no primeiro anel de distância.

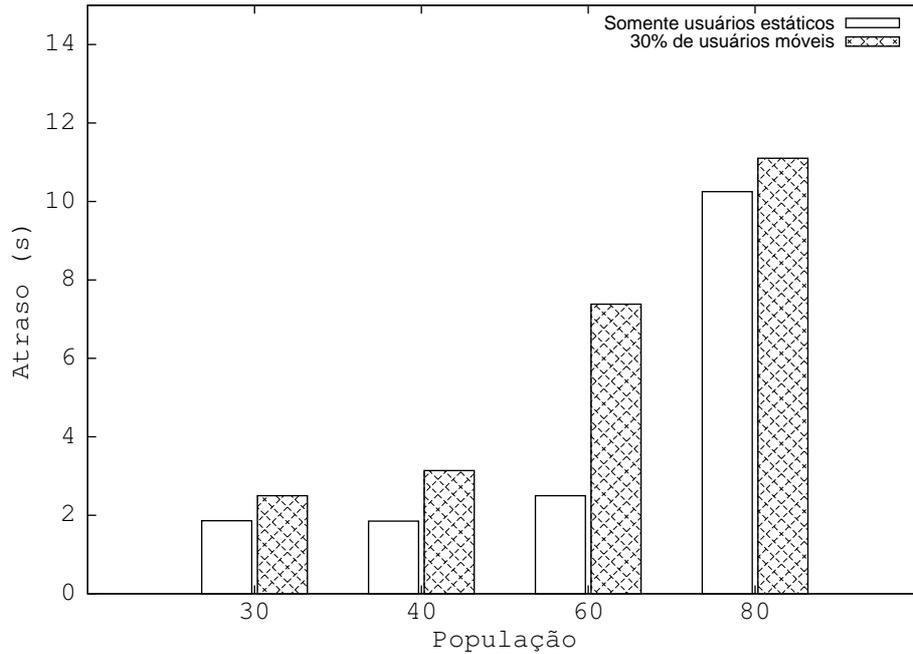


Figura 4.10: Atraso médio observado em função da população considerada (usuários móveis e estáticos vs. estáticos)

Na Figura 4.13, observa-se que para uma população de 60 usuários estáticos, o atraso de um usuário no último anel de distância cresce dez vezes quando comparado ao atraso médio do primeiro anel. Para os experimentos realizados com 30% da população sendo considerada de usuários móveis, este crescimento aumenta em 15 vezes. Em resumo, observamos que tanto o aumento da população quando a presença de usuários móveis causam forte crescimento da injustiça em nível de usuário.

4.2.3 Discussão: Há como melhorar a qualidade de serviço em relação à justiça (*goodput* e atraso)?

De acordo com os resultados obtidos até agora, fomos capazes de observar que há uma injustiça, em nível de usuário, em redes sem fio. É importante notar que esta injustiça possui duas causas principais:

- causa (a): como discutido no capítulo 2.1.3, o teorema de Shannon e Hartley afirma que a maior taxa de bits alcançável de modo a que a taxa de erros não cresça fora do controle é inversamente proporcional à distância que separa dois pontos;

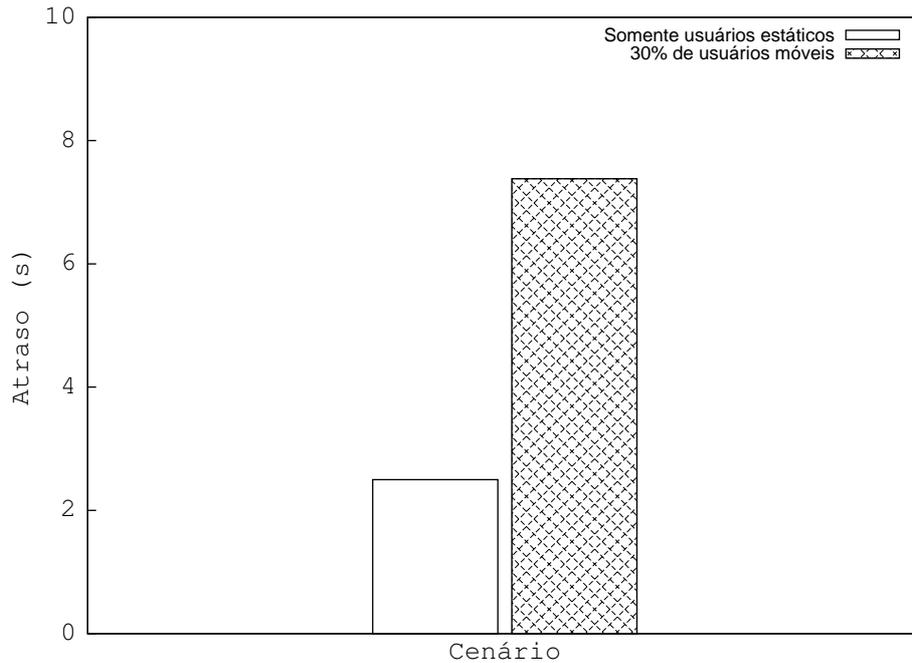


Figura 4.11: Atraso médio (60 usuários) para os cenários 1 e 2, respectivamente

- causa (b): os algoritmos de escalonamento de pacotes usados em redes sem fio de longa distância (ex: EVDO, WIMAX, HSDPA, etc...) utilizam a qualidade do canal (SINR) como métrica para escolha do nó a ser servido. Além disso, os algoritmos de escalonamento mais comumente usados procuram aumentar a vazão da rede através de alguma priorização de usuários que possuam melhor qualidade instantânea do canal.

No Capítulo 3 foram revisados diversos trabalhos que atacaram o caso (b). Estes trabalhos trouxeram inúmeras propostas de novos algoritmos que permitem, no melhor dos casos, a realização de um ajuste fino entre o que se deseja alcançar: maior vazão agregada da rede, ou maior justiça em nível de usuários. O fato de aumentar justiça, invariavelmente deteriora a vazão alcançada pela rede, e vice-versa.

Além disso, os trabalhos discutidos no Capítulo 3, avaliam a justiça e a qualidade de serviço usando métricas que tipicamente refletem o ponto de vista da rede. Discutimos, também, que o ajuste de parâmetros de algoritmos de escalonamento dificilmente resulta em um aumento significativo na justiça em nível de usuário. Este ajuste é feito através do uso de parâmetros que dão alta prioridade a usuários que enfrentam piores condições de meio físico. Neste caso, a justiça entre os usuários aumenta, porém a vazão agregada fica seriamente degradada.

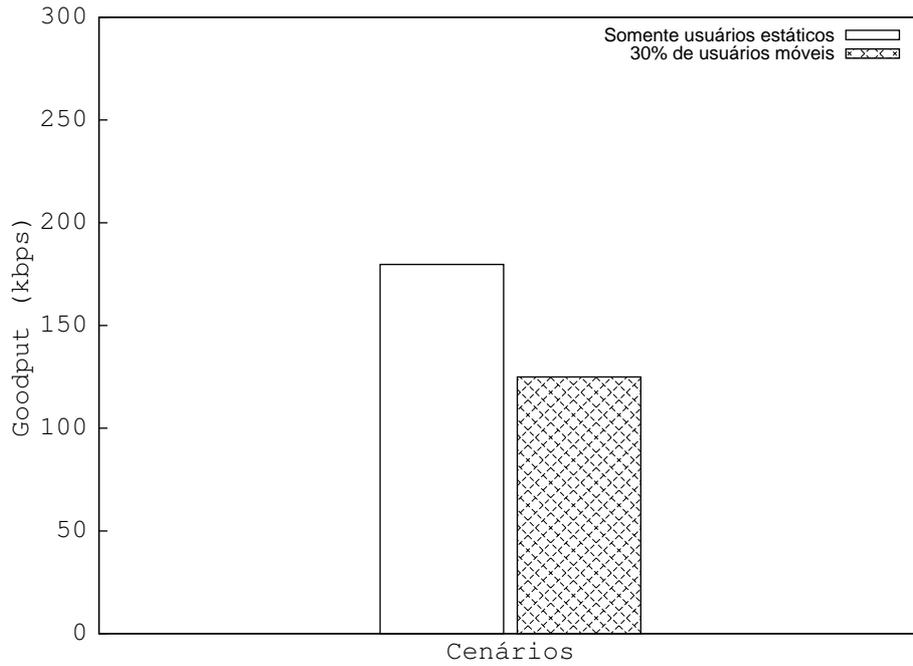


Figura 4.12: *Goodput* médio (60 usuários) para os cenários 1 e 2, respectivamente

Para reforçar nossa observação, realizamos um conjunto de experimentos onde variamos o parâmetro α do algoritmo PFS. Como mencionado anteriormente, o parâmetro α na Equação 4.18 ajusta o grau de justiça do sistema, no que diz respeito ao número de oportunidades de transmissão, e um valor típico para este parâmetro é $\alpha = 0.001$. Assim, buscamos analisar sua influência sobre a justiça em nível de usuário.

Neste experimento específico optamos por analisar o caso de 30 usuários. Com isso queremos evidenciar que, mesmo com o sistema em baixas cargas, a tentativa de melhorar a justiça através de ajustes no parâmetro do algoritmo de escalonamento degrada a vazão agregada do sistema e não resulta em melhora significativa na justiça.

A Figura 4.15 mostra o *goodput* para a população de 30 usuários. Neste experimento o parâmetro α foi configurado para 0.001 e 1. Note que quando $\alpha = 1$ o escalonamento ocorre de acordo com o *Round Robin*. A partir desta figura, pode-se observar que o *goodput* dos usuários dos últimos anéis permanece baixo enquanto o *goodput* dos usuários dos primeiros anéis foi significativamente prejudicado. Portanto não parece adequado modificar o parâmetro α na tentativa de melhorar a justiça em nível de usuário. Isto acontece pois a fração de *slots* alocados a usuários

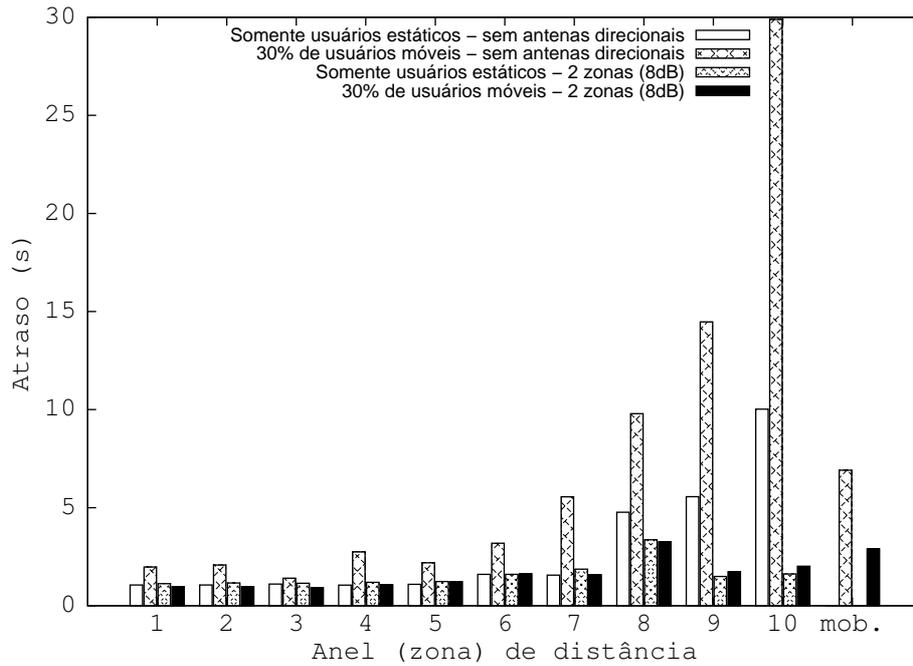


Figura 4.13: Atraso médio em função da distância para cada cenário considerado

com piores condições de SINR aumenta com α , mas esses usuários somente são capazes de receber dados às menores taxas disponíveis. Por esta razão, a taxa de bits agregada do canal direto também é fortemente prejudicada.

Os resultados até aqui apresentados (ex: Figuras 4.7, 4.8 e 4.15) evidenciam que a justiça (em nível de rede, ex: número de *slots*) tradicionalmente discutida em artigos relacionados ao nosso trabalho não reflete em justiça em nível de usuário (ex: vazão e atraso).

4.2.4 Experimentos: Melhorando a justiça em nível de usuário

Neste trabalho propomos uma maneira de reduzir a injustiça intrínseca da camada física e dos algoritmos de escalonamento de redes sem fio. Para isso tomamos uma abordagem diferente dos trabalhos relacionados e buscamos melhorar as condições do nível físico (causa (a)). Neste sentido, adicionamos antenas direcionais aos usuários localizados nos anéis de distância mais prejudicados. Estas antenas podem ser bem baratas, como uma Yagi por exemplo. Além disso, nossos experimentos mostraram que basta uma antena de ganho modesto como, por exemplo, 8dB.

Assim, dois novos cenários serão considerados no próximo conjunto de experimen-

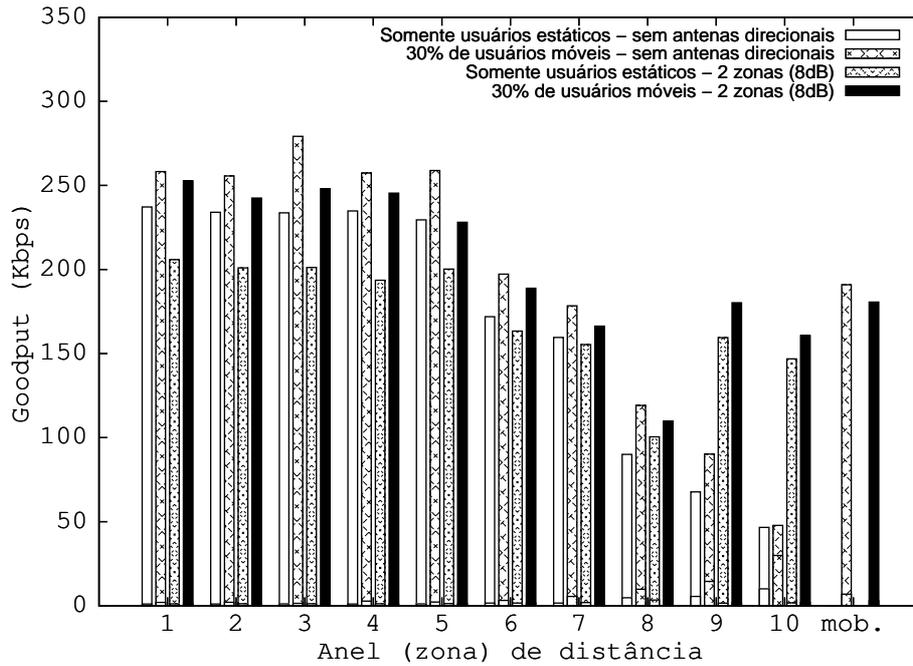


Figura 4.14: *Goodput* médio em função da distância para cada cenário considerados. Nestes novos cenários os usuários localizados nos dois últimos anéis de distância usam antenas direcionais com ganho de $8dB$. No cenário 3 apenas usuários estáticos são considerados e no cenário 4 consideramos que 30% da população é composta de usuários móveis.

A eficácia da solução proposta é mostrada nas Figuras 4.13 e 4.14. Na Figura 4.13 apresentamos o atraso médio para todos os 4 cenários considerados neste trabalho. Note que as diferenças entre os atrasos médios nos últimos anéis de distância diminuí significativamente se compararmos o cenário 1 com o cenário 3 e o cenário 2 com o 4.

Por exemplo, o atraso de um usuário no décimo anel diminuí de 29 segundos (cenário 2) para 2 segundos (cenário 4). Isso ocorre devido a adição de antenas direcionais aos dois últimos anéis de usuários estáticos. Note que o atraso dos usuários móveis, mostrado nas duas últimas barras (mob.) também diminuí quando comparamos o cenário 2 com o cenário 4.

Na Figura 4.14, os resultados para os cenários 3 e 4 mostram que o *goodput* nos anéis 9 e 10, que usam antenas direcionais, é significativamente melhorado se comparados com os cenários 1 e 2, respectivamente. O valor do *goodput* nos anéis 9 e 10 é aproximadamente 75% do valor encontrado para o primeiro anel de distância

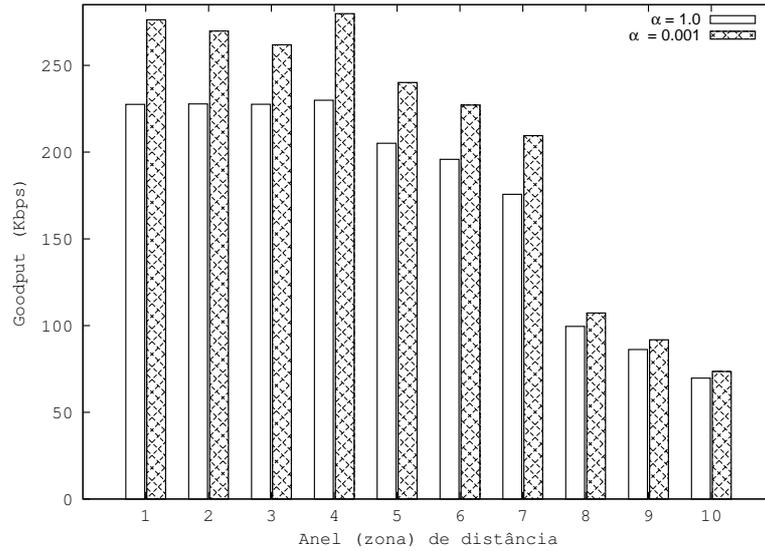


Figura 4.15: *Goodput* em função de α

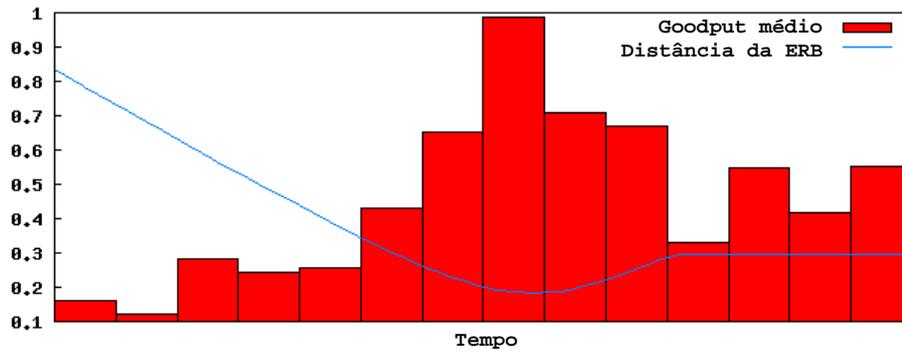


Figura 4.16: Caminho amostral de uma simulação: *goodput* de um usuário em função de sua distância à ERB

nos cenários 3 e 4. Nos cenários 1 e 2, este valor era de 20%.

Para melhor quantificar o ganho na justiça em nível de usuário definimos as seguintes métricas:

- seja σ_g^2 a variância do *goodput* entre os anéis de distância,
- seja σ_d^2 a variância do atraso entre os anéis de distância,
- seja d_g a diferença entre o menor e o maior *goodput* para cada anel de distância e
- seja d_d a diferença entre o menor e o maior atraso para cada anel de distância.

Cenário	σ_g^2	σ_d^2	d_g	d_d
1 - Sem mobilidade: não há antenas direcionais	5844.49	9.05	190.57	8.97
2 - Com mobilidade: não há antenas direcionais	6142.96	72.55	231.45	28.48
3 - Sem mobilidade: antenas direcionais nos últimos 2 anéis	1149.62	0.45	105.39	0.45
4 - Com mobilidade: antenas direcionais nos últimos 2 anéis	2144.12	0.62	142.86	2.34

Tabela 4.8: σ_g^2 , σ_d^2 , d_g e d_d quantificando a justiça em nível de usuário

É importante observar que, para todas as métricas definidas acima, são considerados todos os anéis de distância, incluindo o anel *mob.*.

Através da comparação do cenário 1 com o cenário 3 e do cenário 2 com o cenário 4, na Tabela 4.8, pode-se claramente notar que a justiça em nível de usuário é significativamente melhorada mesmo quando há 30% de usuários móveis no sistema.

Capítulo 5

Uma proposta para priorização de tráfego de vídeo escalável

Este capítulo trata do segundo grupo de contribuições trazido por esta tese, onde nosso principal objetivo é mostrar que é possível alcançar importante nível de priorização dos pacotes mais importantes em fluxos de vídeo escalável, sem o emprego de esquemas elaborados para escalonamento e descarte, assim como motivado na Subseção 1.2.2.

Primeiramente, na Seção 5.1, apresentamos dois conceitos fundamentais para este capítulo: (i) nossa proposta para enquadramento dos pacotes e (ii) nossa definição para nível de rajada (do inglês *burstiness*) e o esquema proposto para transmissão de pacotes.

Em seguida, na Seção 5.2, propomos um modelo analítico capaz de capturar o processo de perda aleatório que é chave para nossa proposta. Resolvemos o modelo para cenários típicos de redes sem fio, e realizamos a combinação dos resultados obtidos com nossa proposta de enquadramento para discutir a viabilidade e as vantagens alcançadas quando nossa ideia é aplicada.

Na Seção 5.3 adicionamos mais funcionalidades ao modelo analítico, com o objetivo de contornar uma limitação: a suposição de quadros de vídeo com tamanho fixo. O modelo obtido é resolvido por simulação e, além de resultados análogos aos discutidos na seção anterior, calculamos, através de um arcabouço do qual o modelo de simulação faz parte, o PSNR do vídeo recebido. Além de analisar a viabilidade e as vantagens alcançadas com o uso de nossa ideia quando cenários de vídeos reais

são considerados, o objetivo final desta seção é avaliar o impacto da priorização de perda alcançada sobre a qualidade objetiva do vídeo.

Na Seção 5.4 introduzimos novas funcionalidades ao modelo da Seção 5.3 a considerar um cenário mais realista do que os anteriores com relação ao último roteador no caminho fonte-destino.

5.1 Enquadramento de pacotes e definição de nível de rajada

Para que seja possível entender os modelos a serem apresentados, é importante apresentarmos o esquema de enquadramento considerado neste capítulo.

Considere a Figura 5.1. Cada fluxo de vídeo é, naturalmente, dividido em quadros. Um quadro de vídeo escalável é ilustrado na Figura 5.1(a). Em nossa proposta, cada quadro é dividido em pacotes tal que os *bits* relativos à camada base (BL) são inseridos no(s) primeiro(s) pacote(s) a ser(em) transmitido(s). Por exemplo, o pacote 0 criado na parte inferior da Figura 5.1(a).

Os pacotes subsequentes contêm os *bits* da camada EL ordenados de acordo com sua relevância para a qualidade final da decodificação. Por exemplo, os pacotes de número de sequência 1 a $n - 1$ criados na parte inferior da Figura 5.1(a). Como os

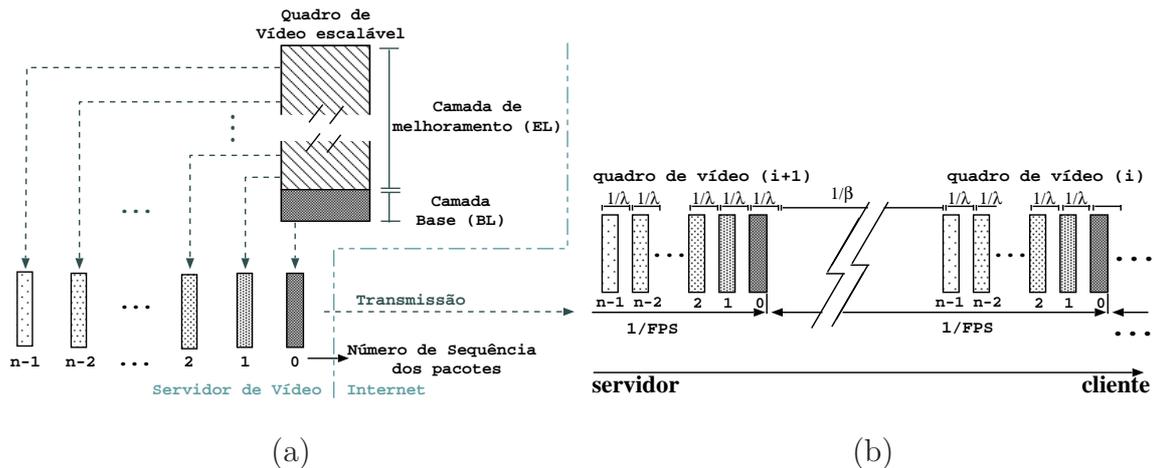


Figura 5.1: (a) Esquema de enquadramento proposto; (b) esquema de transmissão de pacotes proposto

bits de um quadro de vídeo escalável são gerados em ordem decrescente de relevância

para a qualidade da decodificação, o enquadramento da EL é trivial: os primeiros x *bits* gerados pelo codificador são incluídos no campo de dados do segundo pacote a ser transmitido, ou seja, pacote de número de sequência 1. Os *bits* $x + 1$ a $2x$ são inseridos no pacote 2, e assim por diante.

Os pacotes são transmitidos em ordem decrescente de relevância, ou seja, o primeiro pacote a ser transmitido é o pacote 0, seguido pelo pacote 1, e assim sucessivamente, como ilustrado na Figura 5.1 (b).

Claramente, o enquadramento é feito de maneira natural a quem realiza esta tarefa. O ponto chave está na maneira como os pacotes são transmitidos. Todos os pacotes relativos a um mesmo quadro de vídeo são transmitidos em uma mesma rajada e a variável a ser controlada é o intervalo entre dois pacotes consecutivos, como ilustrado nas Figuras 5.1 (b) e 5.2.

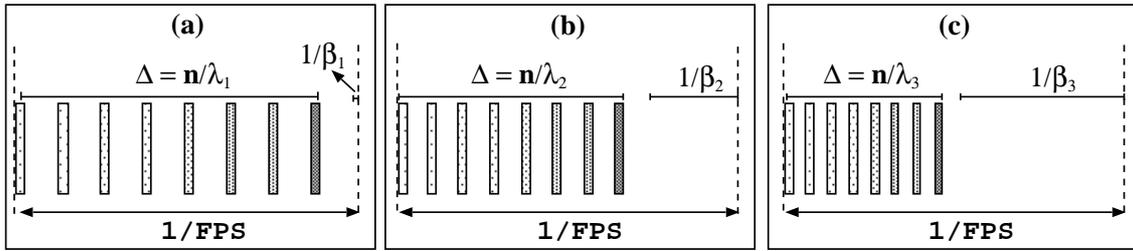


Figura 5.2: Três exemplos ilustrando nossa definição de nível de rajada e a relação entre λ e β .

Seja $\frac{1}{\lambda}$, o intervalo entre a emissão de dois pacotes consecutivos e $\frac{1}{FPS}$ o intervalo entre dois quadros consecutivos de vídeo, onde FPS é a taxa de quadros por segundo (*Frames Per Second*) do vídeo. A duração Δ de uma rajada de pacotes é dada por $\Delta = \frac{n}{\lambda}$, onde n é o número de pacotes por quadro. Equivalentemente,

$$\Delta = \frac{1}{FPS} - \frac{1}{\beta}, \quad (5.1)$$

onde o intervalo entre o final de uma rajada de pacotes e o início da próxima é definido como $1/\beta$. Claramente, nenhum pacote é emitido durante este período.

Note-se que em nossa definição de Δ , consideramos tamanhos de rajadas fixas em n . Essa é uma das suposições de nosso modelo analítico apresentado na Seção 5.2. Na realidade, contudo, diferentes quadros de vídeo (ex: I,B,P) possuem tamanhos potencialmente diferentes. O caso onde n é variável será coberto pelos modelos de simulação apresentados nas seções 5.3 e 5.4.

Na Seção 5.2 será explicado como a combinação do enquadramento proposto e da ordenação da emissão de pacotes são usados para alcançar o objetivo de obter melhor qualidade na decodificação de fluxos de vídeo. Para isso, controlaremos apenas um parâmetro: o espaçamento entre pacotes ou, equivalentemente, o *nível de rajada*.

Um exemplo de como variamos o espaçamento de pacotes é apresentado na Figura 5.2. Considerando $\lambda_1 < \lambda_2 < \lambda_3$ e $\beta_1 > \beta_2 > \beta_3$, os exemplos (a), (b) e (c) ilustram 3 casos em ordem crescente de nível de rajada. Note que a alteração do nível de rajada não modifica na carga oferecida.

5.2 Proposta de um modelo analítico para avaliação de perda intra-rajada em fluxos de vídeo

Esta seção apresenta o modelo analítico proposto, cujo objetivo é calcular a distribuição da perda *intra-rajada*. Para isso, é considerado o cenário geral descrito na Figura 1.2 do Capítulo 1, que é rerepresentada, por conveniência, na Figura 5.3.

Usamos o termo perda intra-rajada para destacar o fato de que estudamos separadamente o nível de perda alcançado para cada pacote i (veja Figura 5.1) das rajadas. Esta é, até onde sabemos, uma nova maneira de estudar o processo de perda em redes de comutação de pacotes. Este novo ponto de vista é fundamental para o desenvolvimento de nossa proposta.

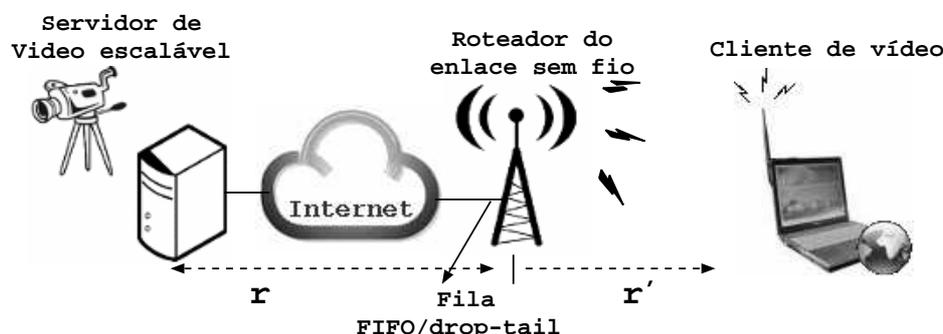


Figura 5.3: Cenário geral considerado neste capítulo

É importante ressaltar que o modelo analítico proposto é propositalmente simples e, por isso, não inclui detalhes inerentes a redes sem fio nem do tráfego de vídeo

escalável. Estas funcionalidades são consideradas em dois modelos de simulação incrementais a serem apresentados posteriormente, nas Seções 5.3 e 5.4.

O modelo analítico, apesar de simplificado, captura o processo básico de perda observado também nos modelos detalhados de simulação, e nos permite construir a base de nossa proposta. Os modelos detalhados de simulação são úteis para avaliar a aplicabilidade do método proposto a cenário realistas como, por exemplo, redes de acesso ADSL (*Asymmetric Digital Subscriber Line*) [115] ou qualquer rede de acesso à Internet onde o último enlace é o gargalo do caminho servidor-cliente. Na Figura 5.3 este cenário seria representado por $r > r'$, onde r é a banda entre o servidor e o roteador do enlace sem fio e r' é a banda entre o enlace sem fio e o cliente.

5.2.1 Descrição do modelo

Desejamos mostrar que é possível melhorar a qualidade do vídeo no receptor através do controle do nível de rajada de um tráfego de vídeo. No modelo analítico, esse tráfego chega a uma fila de *buffer* finito e compete pelo servidor com tráfego de fundo também modelado.

As seguintes suposições são consideradas para o modelo analítico proposto:

- tamanho fixo de pacotes;
- rajadas com o mesmo número de pacotes;
- no servidor (que representa o roteador do enlace sem fio), o tempo entre chegadas de pacotes possui distribuição exponencial, com média igual a $\frac{1}{\lambda}$. O tempo entre chegadas representa as variações do retardo no caminho entre o servidor e o roteador do enlace sem fio. Cabe ressaltar que os pacotes são gerados pelo servidor a intervalo de tempos fixos;
- tempo de serviço exponencial, com média $\frac{1}{\mu}$, no último enlace;
- a reordenação de pacotes, gerada pela Internet, é considerada desprezível. De fato, um estudo de Towsley et al. [116] mostrou que 4% dos pacotes chegam ao destinatário fora de sequência, com uma pequena porcentagem ($\approx 1\%$) sendo devido à reordenação. A maioria das ocorrências é devido à recuperação

de perda (ex: retransmissões do TCP). Como tratamos de transmissão de vídeo, que ocorre através do protocolo UDP, é razoável desconsiderarmos a reordenação;

- tráfego de fundo, agregado, é representado por um processo de Poisson.

A Figura 5.4 apresenta os módulos principais do modelo proposto. O módulo *Last_Queue* é uma abstração de uma fila com serviço exponencial de taxa μ , disciplina de serviço FIFO e política de descartes *Drop-Tail*. Este módulo representa a fila do último salto (ex: enlace sem fio) no caminho servidor-cliente. Assumimos que *Last_Queue* é compartilhado por outros tráfegos (tráfego de fundo) modelados como sendo uma fonte Poisson [52].

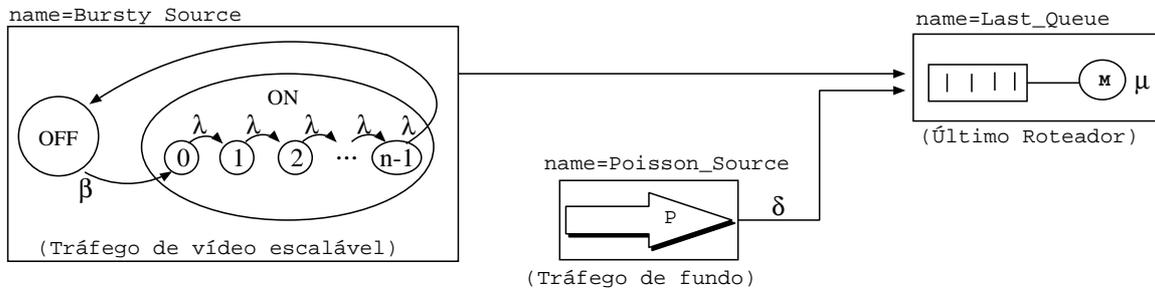


Figura 5.4: Modelo analítico proposto: visão geral

O módulo de nome *Bursty_Source* representa o processo de chegada de pacotes de vídeo na fila do roteador sem fio. Enquanto *Bursty_Source* estiver no estado *OFF*, não ocorrerá geração de pacotes de vídeo para *Last_Queue*. Assumimos que o período *OFF* possui distribuição exponencial com média $\frac{1}{\beta}$.

Quando, eventualmente, o estado de *Bursty_Source* muda para *ON*, um número fixo de pacotes é enviado ao módulo *Last_Queue*. Cada pacote contém parte de um quadro de vídeo escalável gerado no servidor. O tempo entre chegadas de pacotes é modelado como uma distribuição exponencial de média $\frac{1}{\lambda}$.

Relembramos que, embora os pacotes da fonte de vídeo possam ser gerados a intervalos constantes, o tempo entre chegadas não será determinístico devido ao *jitter* introduzido pela Internet. Portanto, é razoável supor que o intervalo entre a chegada de pacotes de vídeo é exponencialmente distribuído. Nos modelos de simulação veremos que essas suposições não influenciam o comportamento a ser mostrado.

Seja f a taxa de quadros por segundo (FPS) do vídeo. Pelo modelo proposto, a duração média do tempo de um ciclo *ON-OFF* é o intervalo entre dois quadros consecutivos, ou seja, igual a $\frac{1}{f}$. Claramente, $\frac{1}{\beta} + \frac{n}{\lambda} = \frac{1}{f}$, e β pode ser escrito tanto em função de λ quanto de f :

$$\beta = \frac{f\lambda}{(\lambda - (nf))} \quad \text{para } \lambda > nf. \quad (5.2)$$

No modelo analítico proposto, o nível de rajada, b , é definido como sendo a razão entre a taxa de chegada de pacotes enquanto em estado *ON* (λ) e a taxa média de chegada de pacotes nf :

$$b = \frac{\lambda}{nf} \quad (b > 1). \quad (5.3)$$

Portanto,

$$\beta = \frac{bf}{(b-1)}. \quad (5.4)$$

Através do modelo descrito acima é fácil controlar b . Se fixarmos a taxa de geração de quadros de vídeo em f , b pode ser ajustado através de λ , de acordo com a Equação 5.3. Além disso, usando a Equação 5.4, β pode ser obtido em função de b .

O módulo *Poisson_Source* gera tráfego Poisson que representa o tráfego de fundo agregado [117, 118], que compartilha recursos de *Last_Queue* com o tráfego de vídeo. Nós definimos δ como a taxa média de chegada de tráfego de fundo.

Seja ρ a carga do sistema representada pela razão entre a taxa total de chegada de pacotes vídeo mais tráfego de fundo ($\delta + nf$) e a taxa de serviço μ :

$$\rho = \frac{\delta + nf}{\mu}. \quad (5.5)$$

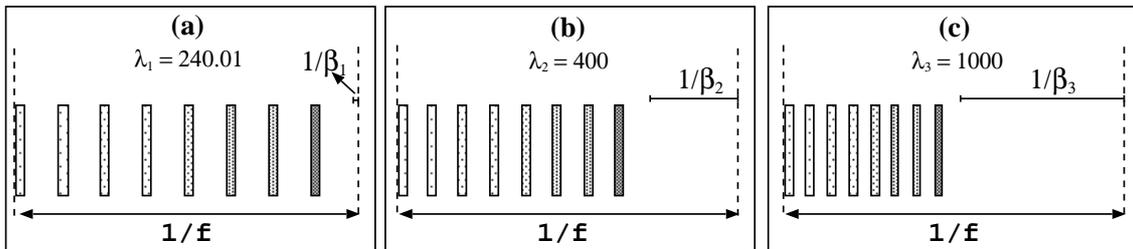


Figura 5.5: Três exemplos ilustrando nossa definição de nível de rajada e a relação entre λ e β .

A Figura 5.5 ilustra três exemplos de valores para λ e β . Note que $\frac{1}{\beta}$ cresce à medida que λ aumentada (veja equações 5.3 e 5.4), de forma que a carga média e o tempo médio entre chegada de quadros (rajadas) permaneça constante.

Isso ocorre pois o tempo em OFF ($\frac{1}{\beta}$) é definido em função de λ fixando-se o tempo médio entre dois quadros consecutivos ($\frac{1}{f}$). A carga ρ (Equação 5.5) depende, portanto, apenas da taxa média de pacotes (nf) e independe tanto do intervalo médio entre pacotes durante a rajada ($\frac{1}{\lambda}$) quanto do nível de rajada (b).

Apresentados os módulos do modelo analítico proposto, podemos passar para o próximo passo que consiste em calcular a fração de perda de cada i -ésimo pacote das rajadas, ou seja, a fração de perda intra-rajada. Para isso, precisamos definir o estado do modelo, e gerar sua respectiva cadeia de Markov correspondente (diagrama de transição entre estados). Destacaremos cada transição de estado que incorre em perda de um pacote de vídeo.

Seja S um estado, do modelo analítico proposto, definido pela concatenação de três variáveis de estado (s_1, s_2 e s_3), onde s_1 indica o estado (ON (1) ou OFF (0)), s_2 representa o número de sequência do pacote da rajada presente ($(0, 1, \dots, n-1)$) e s_3 é o número de pacotes aguardando na fila de *Last_Queue*, com *buffer* de capacidade igual a q pacotes.

A cadeia de Markov para este modelo analítico é ilustrada na Figura 5.6. Na última linha desta cadeia, as transições tracejadas em vermelho representam especificamente a chegada de cada pacote de vídeo que resulta em perda. Note que o conjunto de estados com $s_3 = q$ contém todos os estados representando fila cheia. Então, uma transição

$$(1, s_2, q) \rightarrow (1, ((s_2 + 1) \bmod n), q) \quad (5.6)$$

indica a perda de um pacote de vídeo de número de sequência s_2 .

5.2.2 Definição das medidas de interesse

Defina γ como o número médio de pacotes de vídeo perdidos por unidade de tempo, ou seja, a taxa média de perda de pacotes de vídeo. Como o modelo proposto é markoviano, a fração de tempo que o sistema permanece no estado (s_1, s_2, s_3) é

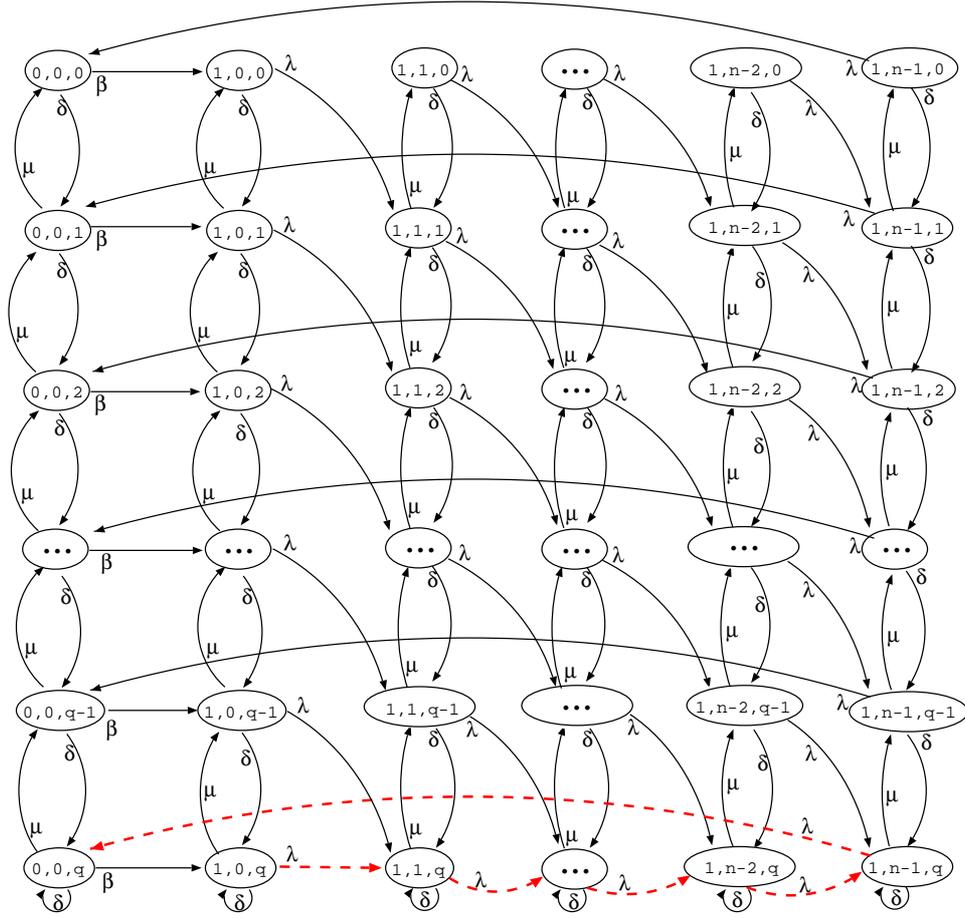


Figura 5.6: Representação da cadeia de Markov do modelo analítico

dada por $\pi_{(s_1, s_2, s_3)}$. O cálculo de γ pode, portanto, ser feito como a seguir:

$$\gamma = \sum_{i=0}^{n-1} \lambda \pi_{1,i,q} \quad .$$

Adicionalmente, seja γ^i a taxa média de perda de pacotes com número de sequência igual a (i) . Então,

$$\gamma^i = \lambda \pi_{1,i,q}, \quad \text{para } i = 0, \dots, n-1.$$

Definimos F^i como a razão entre a taxa média de perda de pacotes de número de sequência (i) e a taxa média de geração de pacotes de número de sequência (i) . Adicionalmente, definimos F como a razão entre a taxa média de perda de pacotes de vídeo e a taxa média de chegadas de todos os pacotes de vídeo. Logo:

$$F^i = \frac{\gamma^i}{f} \quad \text{e} \quad F = \frac{\gamma}{nf}. \quad (5.7)$$

5.2.3 Obtendo solução para o modelo

Para explicar como obtemos a solução para o modelo analítico proposto, apresentaremos um exemplo para avaliar a estrutura de sua matriz geradora infinitesimal (Q). Em seguida, mostraremos que esta matriz possui uma estrutura conhecida, para a qual existem algoritmos eficientes para solução em estado estacionário. Finalmente, discutiremos que qualquer parametrização do modelo resulta em matriz com a mesma estrutura.

Considere o diagrama de transição de estados apresentado na Figura 5.7.

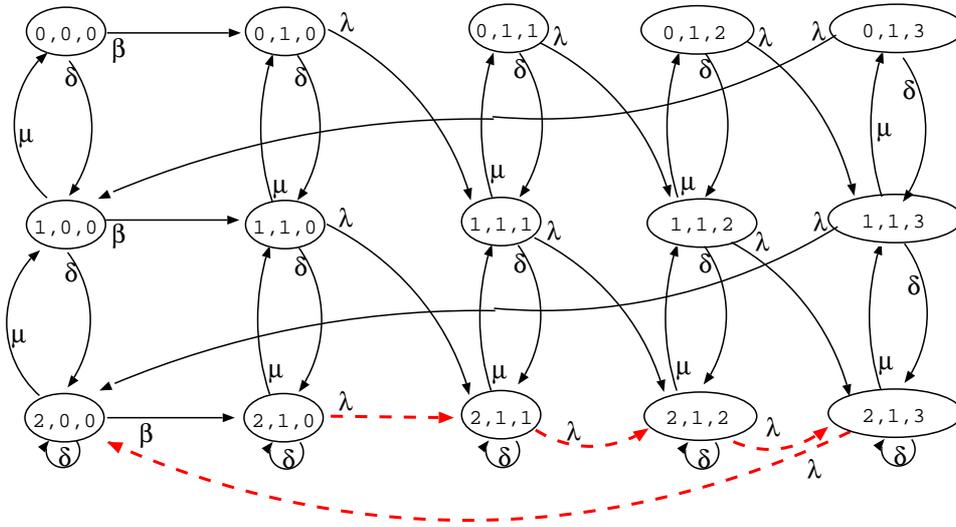


Figura 5.7: Exemplo de modelo para $n=4$ e $q=2$

Para se obter uma solução eficiente do modelo, buscamos por alguma estrutura especial. Como a ordenação das variáveis de estado pode resultar em estruturas diferentes, observamos que, para o modelo proposto, a ordenação de estados que resulta em uma estrutura conhecida e computacionalmente eficiente é dada por (s_2, s_1, s_3) , de acordo com nossa definição de espaço de estados.

Considere um exemplo do modelo proposto onde $n = 4$ e $q = 2$. Construimos, então, a matriz \mathbf{Q} com taxas de mudança entre os estados do processo, como apresentado abaixo. Para facilitar a visualização, optamos por adicionar à primeira linha e à primeira coluna valores $(s_2 s_1 s_3)$ representando cada estado. Além disso, transições com taxa igual a zero foram ocultadas:

$$Q = \begin{bmatrix} & 000 & 010 & 011 & 012 & 013 & 100 & 110 & 111 & 112 & 113 & 200 & 210 & 211 & 212 & 213 \\ 000 & x & \beta & & & & \delta & & & & & & & & & \\ 010 & & y & & & & & \delta & \lambda & & & & & & & \\ 011 & & & y & & & & & \delta & \lambda & & & & & & \\ 012 & & & & y & & & & & \delta & \lambda & & & & & \\ 013 & & & & & y & \lambda & & & & \delta & & & & & \\ 100 & \mu & & & & & z & \beta & & & & \delta & & & & \\ 110 & & \mu & & & & & z & & & & & \delta & \lambda & & \\ 111 & & & \mu & & & & & z & & & & & \delta & \lambda & \\ 112 & & & & \mu & & & & & z & & & & & \delta & \lambda \\ 113 & & & & & \mu & & & & & z & \lambda & & & & \delta \\ 200 & & & & & & \mu & & & & & w & \beta & & & \\ 210 & & & & & & & \mu & & & & & k & \lambda & & \\ 211 & & & & & & & & \mu & & & & & k & \lambda & \\ 212 & & & & & & & & & \mu & & & & & k & \lambda \\ 213 & & & & & & & & & & \mu & \lambda & & & & k \end{bmatrix}$$

onde $x = -(\beta + \delta)$, $y = -(\delta + \lambda)$, $z = -(\beta + \delta + \mu)$, $w = -(\beta + \mu)$, $k = -(\lambda + \mu)$ e $j = -(\lambda + \mu + \delta)$.

Se dividirmos a matriz Q em submatrizes de dimensão $(n - 1) \times (n - 1)$ (neste exemplo, 3×3), obteremos a matriz Q' :

$$Q' = \begin{bmatrix} \begin{array}{cc|cc} x & \beta & \delta & \\ & y & \delta & \lambda \\ & & \delta & \lambda \\ & & & \delta & \lambda \\ & & & & \delta & \lambda \\ & & & & & \delta \end{array} \\ \hline \begin{array}{cc|cc} \mu & & z & \beta \\ & \mu & z & \\ & & z & \\ & & & z \\ & & & & z \\ & & & & & z \end{array} \\ \hline \begin{array}{cc|cc} & & \mu & \\ & & \mu & \\ & & & \mu \\ & & & & \mu \\ & & & & & \mu \end{array} \\ \hline \begin{array}{cc|cc} & & w & \beta \\ & & k & \lambda \\ & & & k & \lambda \\ & & & & k & \lambda \\ & & & & & k \end{array} \end{bmatrix}.$$

Claramente, a matriz Q' pode ser dividida em submatrizes, como a seguir:

$$Q_0 = \begin{bmatrix} x & \beta & & & \\ & y & & & \\ & & y & & \\ & & & y & \\ & & & & y \end{bmatrix}, \quad A_0 = \begin{bmatrix} \delta & & & & \\ & \delta & \lambda & & \\ & & \delta & \lambda & \\ & & & \delta & \lambda \\ \lambda & & & & \delta \end{bmatrix},$$

$$A_1 = \begin{bmatrix} z & \beta & & & \\ & z & & & \\ & & z & & \\ & & & z & \\ & & & & z \end{bmatrix}, \quad A_2 = \begin{bmatrix} \mu & & & & \\ & \mu & & & \\ & & \mu & & \\ & & & \mu & \\ & & & & \mu \end{bmatrix} \text{ e}$$

$$Q_M = \begin{bmatrix} w & \beta & & & \\ & k & \lambda & & \\ & & k & \lambda & \\ & & & k & \lambda \\ \lambda & & & & k \end{bmatrix}.$$

Se escrevermos a matriz Q' em função destas submatrizes, obtemos uma matriz tridiagonal, Q_{BD} , análoga à matriz Q :

$$Q_{BD} = \begin{bmatrix} Q_0 & A_0 & & \\ A_2 & A_1 & A_0 & \\ & A_2 & Q_M & \end{bmatrix}$$

Com este exemplo, queremos mostrar que independentemente dos parâmetros q e n considerados no modelo, a matriz geradora infinitesimal (Q) sempre possui estrutura em blocos (*block Quasi-Birth-Death* - QBD) [119, 51].

Claramente, alterações nos parâmetros q e n não farão com que Q deixe de possuir estrutura QBD. Isso acontece pois o parâmetro n está relacionado com a dimensão das matrizes Q_0 , A_0 , A_1 , A_2 e Q_M . De maneira análoga, o parâmetro q se relaciona com a dimensão da matriz Q_{BD} (ex: ocorrências das submatrizes A_0 , A_1 , A_2).

Portanto, a matriz geradora infinitesimal para o modelo analítico proposto sem-

pre terá estrutura de um processo QBD, independentemente da parametrização do modelo.

Embora seja difícil obter solução fechada para tal classe de modelos markovianos, existem algoritmos eficientes [51, 120] para sua solução numérica, ou seja, obtenção da solução em estado estacionário (vetor π). Com isso, o cálculo de medidas de interesse (Equação 5.7) é computacionalmente eficiente.

Para implementar e resolver o modelo proposto, usamos o ambiente de modelagem e análise de desempenho denominado Tangram-II [120, 121]. Esta ferramenta agrega diversos algoritmos para solução de modelos markovianos e não-markovianos, incluindo a classe de modelos que nos interessa.

5.2.4 Parâmetros considerados

Consideramos um conjunto de parâmetros, que estão de acordo com o que é normalmente encontrado em aplicações de fluxo de vídeo em cenários comuns de redes sem fio. Para todas as análises realizadas nesta seção, a taxa (f) de quadros por segundo dos vídeos é igual a $30FPS$ e a taxa do vídeo é equivalente a 2.80Mbps. Este valor para taxa de *bits* do vídeo equivale a um fluxo de vídeo de alta qualidade. As taxas de quadro e de *bits* escolhidas implicam que a geração de pacotes de vídeo seja de 240 pacotes por segundo. Adicionalmente, considerando pacotes com comprimento total de 1500 *bytes*, o número de pacotes gerados por quadro é igual a 8, ou seja, o tamanho das rajadas é fixa em $n = 8$ pacotes.

A capacidade de serviço μ é de 5.80Mbps, o que se aproxima da taxa máxima de serviço de redes 3G de acesso sem fio à Internet, como algumas versões do EVDO e HSDPA. Portanto, a taxa de serviço do módulo *Last_Queue* da Figura 5.4 é de $\mu = 484$ pacotes por segundo.

Escolhemos três valores para a capacidade do *buffer*: $q = 25$, $q = 50$ e $q = 150$. Os primeiros valores ($q = 25$ e $q = 50$) são comumente usados na literatura e representam tamanhos comuns de *buffers* em equipamentos de rede sem fio. O último valor ($q = 150$) representa um *buffer* superdimensionado, maior do que qualquer valor razoável. Isso foi feito para mostrar que os resultados são válidos para uma ampla faixa de tamanhos de *buffers*.

O nível de rajada, b , foi variado de $b + \epsilon$ ¹ a 83.33. De maneira análoga, como $b = \frac{\lambda}{nf}$, λ varia de 240.01 a 2000.

A carga de tráfego de fundo, gerado pelo módulo *Cross-Traffic* da Figura 5.4, varia de 99 a 294 pacotes por segundo. Esta parametrização equivale, de acordo com a Equação (5.5), a valores de ρ que representam desde cenários de carga média ($\rho = 0.7$) até cenários de sobrecarga ($\rho = 1.1$).

Por conveniência, a Tabela 5.1 resume todos os parâmetros considerados no estudo apresentado mais adiante.

Parâmetro	Valor/Faixa	Explicação	Correspondência
f	30	quadros por segundo (FPS)	vídeo de 30 FPS
n	8	tamanho de rajada fixo em n (estado <i>ON</i>)	8 pacotes por rajada
nf	$8 \times 30 = 240$	taxa média de pacotes de vídeo	2.88Mbps
μ	484	taxa média de serviço	5.88Mbps
q	25, 50, 150	capacidade do buffer em pacotes	fila com atraso máximo de 0.052s a 0.31s
δ	99 a 295	tráfego de fundo de 1.18 a 3.54Mbps	carga (ρ) de 70% a 110%
λ	240.01 a 20000	b de 1 (suave) a 83.33 (em fortes rajadas)	2.88Mbps a 240Mbps

Tabela 5.1: Parâmetros considerados para solução do modelo analítico

5.2.5 Resultados e discussão

A Figura 5.8(a) mostra a fração de pacotes de vídeo (F) perdidos em função do nível de rajada b . Como esperado, para um dado valor de ρ , a fração de perda aumenta consideravelmente à medida que b cresce. O aumento da fração de perda em função do nível de rajada de um tráfego é um fenômeno conhecido e esta observação apenas corrobora o que seria esperado. Adiante, comentaremos mais sobre esse assunto.

¹ Como b deve ser maior que 1 para que β seja positivo (veja Equação (5.4)), ϵ é definido como um valor bem pequeno (0.41×10^{-5}) em relação à faixa de valores considerados para b .

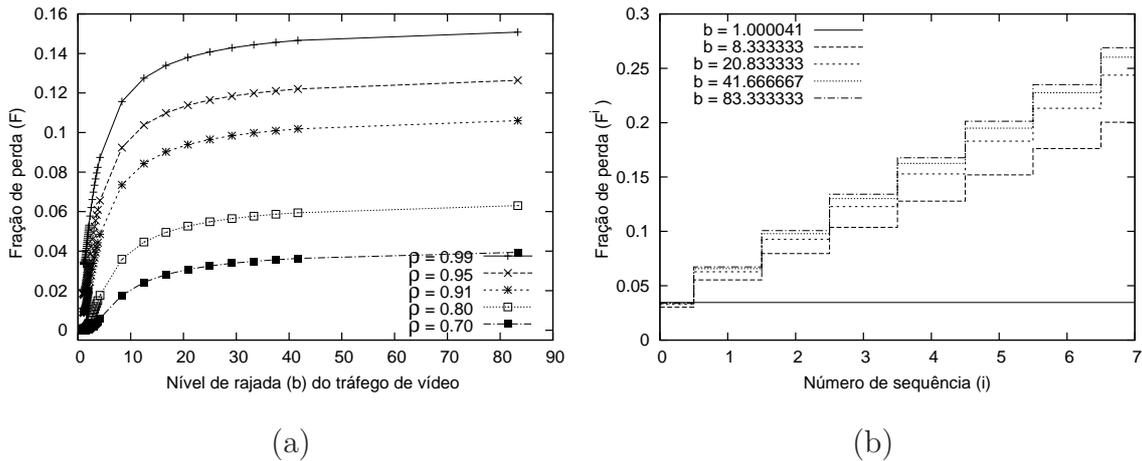


Figura 5.8: (a) Fração de perda do tráfego de vídeo, $q=25$; (b) Distribuição da perda intra-rajada, para $\rho = 0.99$ e $q=25$

A Figura 5.8(b) apresenta a fração de perda de pacotes para cada i -ésimo pacote gerado em cada rajada (quadro de vídeo). O eixo x representa o número de sequência i do pacote, sendo 0 o primeiro pacote de cada rajada.

É importante lembrar que os pacotes com $i = 0$ contêm, segundo o esquema de enquadramento proposto, $bits$ relativos à BL, enquanto os pacotes seguintes contêm $bits$ relativos à camada EL.

A Figura 5.8(b) mostra que, exceto quando o valor de b é muito próximo a 1, a fração de perda cresce com a ordem do pacote transmitido na rajada. Em outras palavras, quando $b \approx 1$ (pacotes de vídeo são emitidos de modo a que o intervalo entre pacotes de uma mesma rajada seja igual ao intervalo entre o final de uma rajada e o início da seguinte, o que chamamos de tráfego suave - *smooth traffic*), a fração de perda é aproximadamente a mesma para todos os pacotes. Entretanto, à medida que b aumenta, cada pacote subsequente enfrenta nível de perda mais elevado, em relação ao pacote anterior, sendo que o primeiro pacote ($i = 0$) sofre a menor perda, ou seja, $F^0 < F^1 < \dots, F^{n-1}$.

Alguns trabalhos disponíveis na literatura [122, 123, 124] mostram que a perda de pacotes é proporcional ao nível de rajada, assim como mostrado na Figura 5.8(a). Este resultado sugere que um tráfego mais suave ($b \approx 1$) é uma boa opção para transmissão de fluxos de vídeo sobre redes de comutação de pacotes. De fato, este é o método de transmissão mais comumente usado.

Entretanto, os resultados apresentados na Figura 5.8(b) mostram que é possível

ajustar o valor de b com o objetivo de obter um processo de perda de mais favorável para a qualidade de uma transmissão de vídeo do que quando $b \approx 1$. Este é o ponto chave para mostrar que é possível melhorar a qualidade do fluxo de vídeo recebido através do ajuste de b , em relação ao que é tradicionalmente feito ($b \approx 1$).

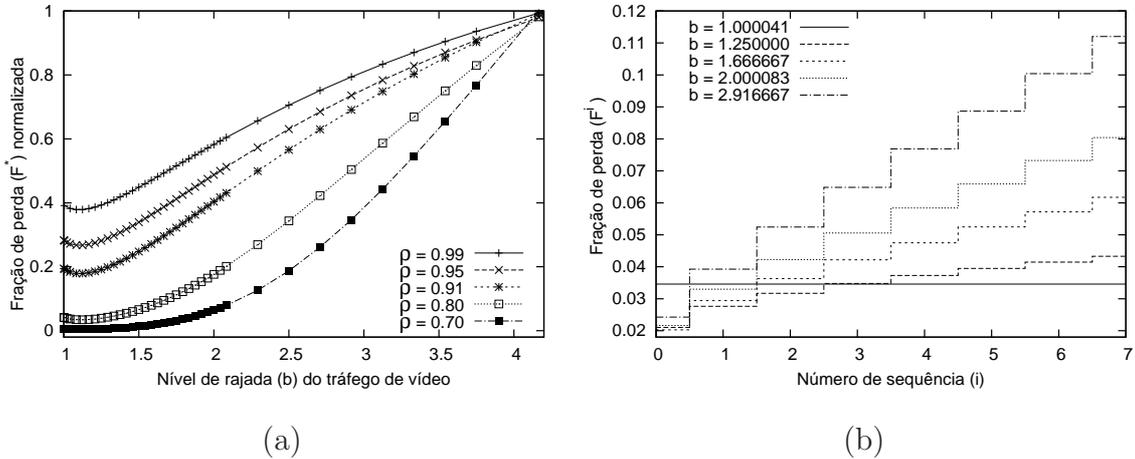


Figura 5.9: (a) Fração de perda total (F) (normalizado), para valores menores de b e $q=25$; (b) Distribuição da perda intra-rajada ($\rho = 0.99$), para valores menores de b e $q=25$

Ilustramos, através das Figuras 5.9(a) e 5.9(b) que, se o ajuste do parâmetro b for corretamente realizado, é possível obter valores menores para F^i ($i = 0, 1, 2$) do que quando $b \approx 1$ (tráfego suave), sem aumento significativo de F , como será discutido mais adiante.

Nossa proposta de enquadramento de pacotes explora este fato. Como apresentado na Seção 5.1, quando um quadro de vídeo escalável é gerado, os pacotes são transmitidos em ordem decrescente de significância para a qualidade da imagem de vídeo a ser decodificada no receptor. Como consequência, os pacotes mais importantes das rajadas são os primeiros e estes deveriam ser preservados com maior prioridade em comparação aos pacotes da cauda das rajadas.

Note que a Figura 5.9(a) é um detalhamento (*zoom*) do lado esquerdo da Figura 5.8(a). Nós focamos nos valores de *nível de rajada* de $1 < b \leq 4.2$, ou seja, o tráfego de vídeo é gradualmente alterado de suave para moderados níveis de rajada. Para facilitar a comparação entre curvas representando diferentes cargas, optou-se por normalizar cada função $F(\rho)$ para um dado ρ através da divisão de $F(\rho)$ pelo valor máximo obtido (F_{max}) para aquele ρ . Os valores normalizados são indicados

por F^* . Isso explica porque todas as funções $F^*(\rho)$ alcançam 1 para $b = 4.2$. Como mostrado na Figura 5.9(a), não há aumento significativo em F^* até que $b > 1.2$ (para $\rho > 0.99$) ou $b > 1.7$ (para $0.7 < \rho < 0.95$).

A Figura 5.9(b) mostra a fração de perda F^i para o i -ésimo pacote de cada rajada, para os valores de b comentados acima. O ponto chave a ser observado é que, nesta faixa de b , F^0 é menor do que quando $b \approx 1$. Isso significa que, através do ajuste correto do nível de rajada (b) do tráfego de vídeo, é possível obter menor nível de perda dos primeiros pacotes de cada quadro, quando comparado ao caso onde $b \approx 1$ (tráfego suave), sendo que $b \approx 1$ é o valor tradicionalmente usado. Desta maneira, favorece-se os pacotes mais importantes para a qualidade do vídeo decodificado.

Esse ganho é alcançado, de acordo com a Figura 5.9(b), ao custo do aumento da fração de perda dos pacotes que carregam *bits* menos importantes para a decodificação de vídeo. De fato, há um balanceamento decorrente da seleção de b na faixa $1.2 < b < 3$: parte do aumento da perda dos pacotes mais ao final da rajada é compensado pela diminuição da perda dos pacotes do início da rajada. Logo, intuitivamente, a qualidade do fluxo de vídeos escaláveis pode ser melhorada, já que é melhor perder quadros que contenham *bits* menos importantes, sobretudo os últimos *bits* da EL (últimos pacotes das rajadas) a perder *bits* da camada base.

Na Figura 5.10(a), apresentamos a fração de perda intra-rajada, ou seja, F^i para cada pacote individual de uma rajada, como função de b para $q = 25$. Note que para o 0-ésimo pacote (ex: primeiro pacote de cada rajada, que contém *bits* da BL), o valor mais baixo para a fração de perda (F^i) é alcançada quando $1.2 < b < 1.5$. Por outro lado, F^i , para $i \geq 4$, aumenta de $\approx 25\%$ a 30% . Se b é ajustado para valores maiores que 2, F^i volta a crescer até atingir o comportamento ilustrado na Figura 5.8(b).

Nota-se claramente que há uma faixa de valores do nível de rajada que resultam em sensível diminuição da fração de perda de pacotes BL sem aumentar significativamente a fração de perda total (F). Por exemplo, a escolha de $b \approx 1.4$ resulta na redução de 40% da fração de perda de pacotes BL (F^0), quando comparado com o caso onde $b \approx 1$, com apenas um pequeno impacto na fração de perda total ($F^0 = F = 3.5\%$ quando $b \approx 1$; e $F^0 = 2.0\%$, $F = 3.7\%$ quando $b = 1.4$). Estes

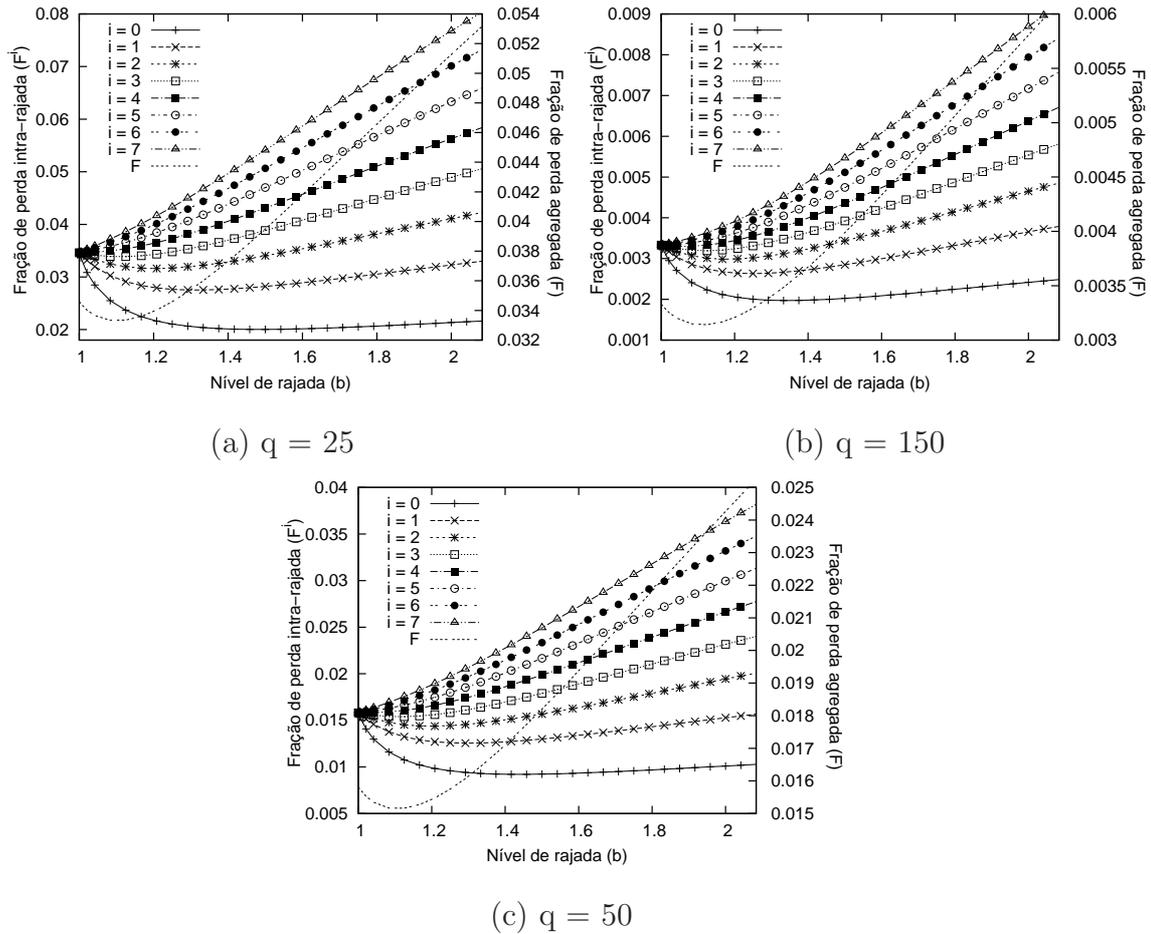


Figura 5.10: ($\rho = 0.99$): Fração de perda intra-rajada (F^i) em função do nível de rajada b

são resultados interessantes já que os pacotes BL são, de longe, os mais importantes para a decodificação de cada quadro de vídeo e, por isso, devem ser preservados.

As Figuras 5.10(b) e 5.10(c) mostram que é possível tirar conclusões análogas às apresentadas acima para outros valores de tamanho de *buffer* ($q = 150$ e $q = 50$, respectivamente). É interessante mostrar que embora os valores absolutos de F^i e F sejam menores quando $q = 150$ e $q = 50$, a faixa de b que favorece pacotes BL não muda significativamente.

O próximo grupo de resultados mostra que o processo de perda observado até agora continua válido para diferentes valores de carga e de tamanhos de *buffer*. Na Figura 5.11 consideramos $\rho = 1.10$ e na Figura 5.12, $\rho = 0.89$. Em ambos os casos, três valores para q são considerados: 25, 50 e 150.

Quando um cenário de sobrecarga é considerado ($\rho = 1.10$), o processo de perda é muito similar ao mostrado na Figura 5.10 (onde $\rho = 0.99$), com a diferença de que

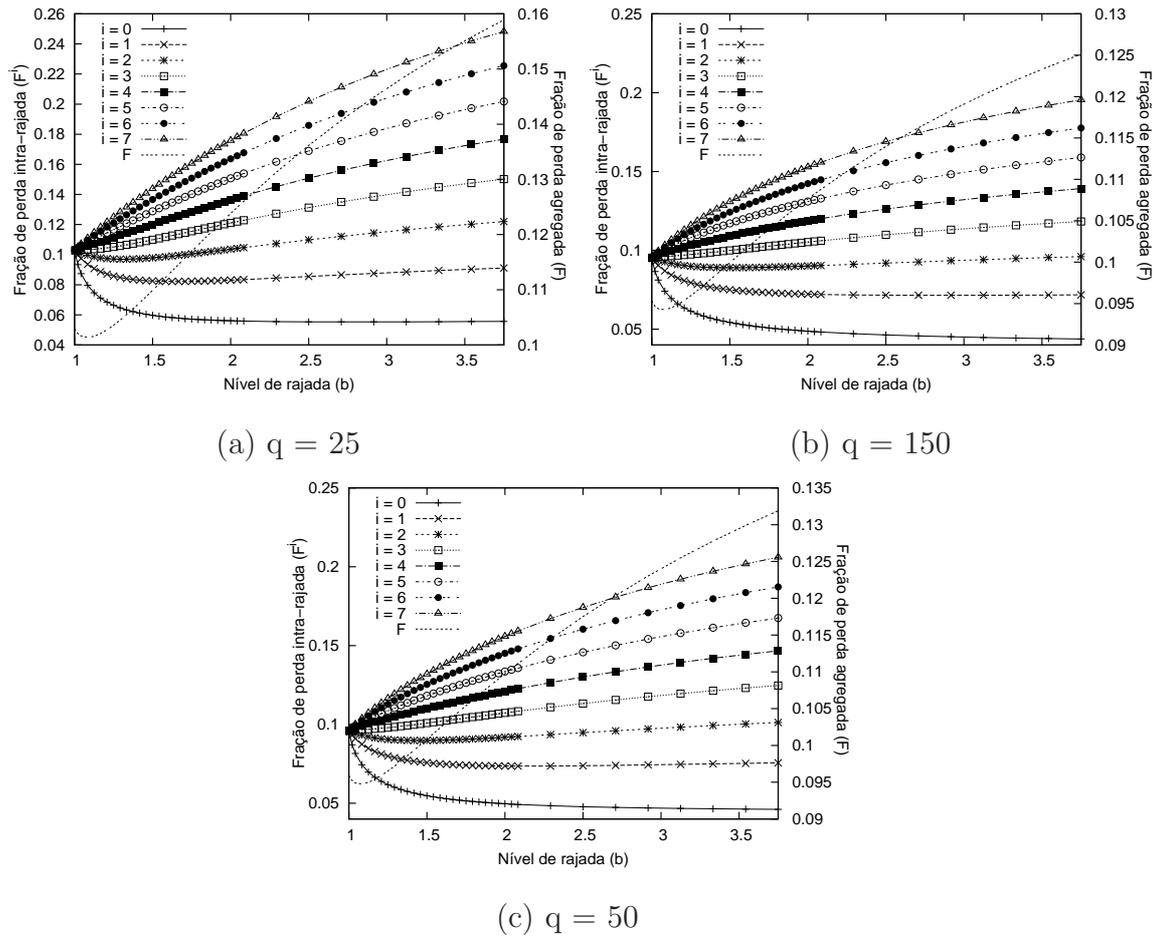


Figura 5.11: ($\rho = 1.105$): Fração de perda intra-rajada (F^i) em função do nível de rajada b

o valor absoluto de F^i e F são maiores.

Conclusões similares podem ser tiradas através da observação da Figura 5.12. É importante notar que, para a ampla faixa de cargas e tamanhos de *buffers* considerados, a faixa de valores de b que produzem menor perda de pacotes BL, com pouco aumento da perda total, está sempre no intervalo $[1.2, 1.8]$.

Este resultado indica que o valor de b pode ser escolhido no mesmo intervalo independente o tamanho de *buffer* (q) e da carga (ρ) do roteador sem fio.

Como mencionado na Seção 5.2, foram desenvolvidos modelos de simulação realistas, cujo objetivo é estender o estudo realizado até agora, relaxando as principais simplificações levadas em conta pelo modelo analítico proposto. Como veremos, os resultados obtidos através dos modelos de simulação mostram o mesmo comportamento geral da distribuição de perda intra-rajada discutidos até aqui.

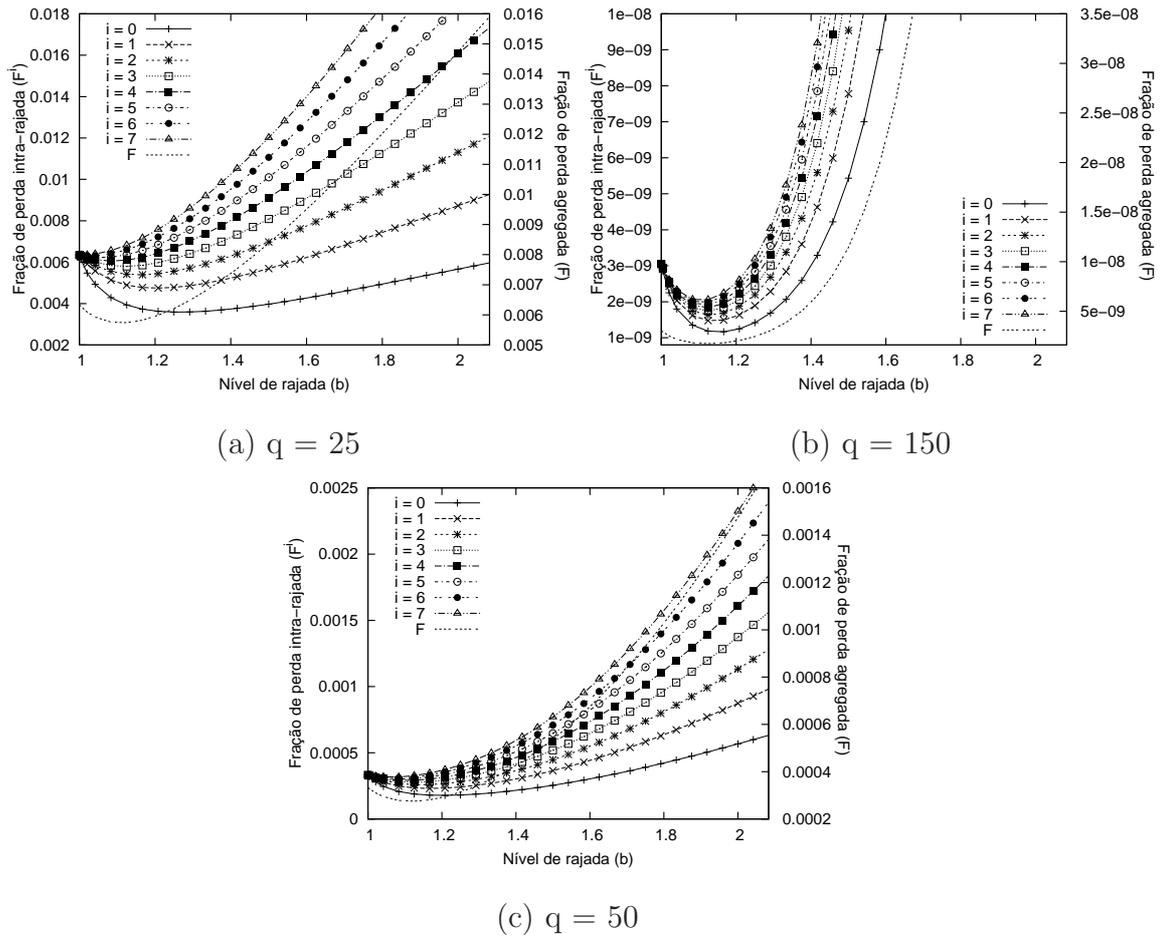


Figura 5.12: ($\rho = 0.89$): Fração de perda intra-rajada (F^i) em função do nível de rajada b

5.3 Modelo de simulação A: vídeos reais

Nesta seção propomos um modelo de simulação que, assim como o modelo analítico apresentado na Seção 5.2, considera o cenário geral da Figura 5.3. Entretanto, o modelo de simulação inclui maiores detalhes do que o analítico e ainda nos permite calcular a qualidade (PSNR [30]) do vídeo recebido pelo cliente, como mostraremos adiante.

5.3.1 Descrição do modelo

Uma visão geral do modelo de simulação A é apresentada na Figura 5.13. Os módulos *Cross-Traffic* e *Last-Router* são basicamente os mesmos.

A principal diferença entre este modelo e o analítico está na substituição do módulo *Bursty-Source* por um novo módulo: *Video-Source*. Assim como no modelo

analítico, o tempo entre chegadas de pacotes é modelado por uma distribuição exponencial. Entretanto, o número de pacotes por quadro de vídeo e o enquadramento são mais próximos a uma transmissão real de vídeo.

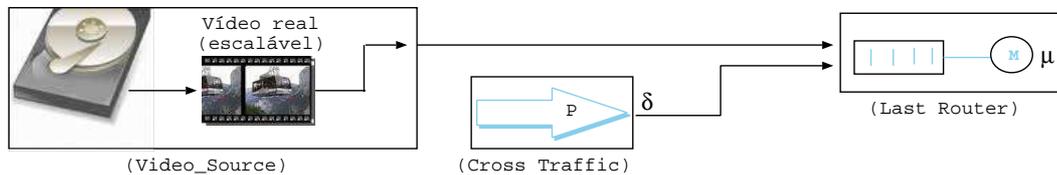


Figura 5.13: Visão geral do modelo de simulação A

Durante a implementação do módulo *Video_Source*, estudou-se duas abordagens para representar padrões de tráfego gerados em fluxos de vídeo escalável: a implementação de modelos sintéticos ou a implementação de um emulador que representa fluxos reais.

De fato, diversos trabalhos disponíveis na literatura [125, 126, 127, 128, 129, 130, 131] propõem modelos matemáticos que caracterizam tráfegos de fontes de fluxos de vídeo escalável. Entretanto, o uso desta abordagem em nosso modelo de simulação seria um fator limitante pois não seria possível estudar o impacto de diferentes padrões de perda sobre a qualidade (ex: PSNR) do vídeo recebido pelo usuário final.

Portanto, ao invés de implementar, no módulo *Video_Source*, um modelo matemático sintético disponível na literatura, optamos por usar *traces* reais de cenas de vídeo bem conhecidas [132] da área de codificação de vídeo. Esta escolha foi tomada com o objetivo de construir um modelo de simulação com o qual seja possível não só estudar o processo de perda de fluxos de vídeos em filas FIFO, mas também calcular a qualidade resultante do vídeo decodificado no receptor. Um resumo das diferenças entre o modelo de simulação e o analítico é apresentado na Tabela 5.2.

Primeiramente, ao contrário do modelo analítico, onde o número de pacotes por rajada (n) é fixo, o modelo de simulação define n_j como sendo o número de pacotes gerados por cada quadro (j) de um vídeo real.

Como consequência, o enquadramento de pacotes também reflete um cenário mais realista, pois depende do tamanho de cada camada i em cada quadro j , como ilustrado na Figura 5.14

Outra diferença está na relação entre a perda de um pacote i e seu efeito sobre

Funcionalidade	Modelo analítico	Modelo de simulação
Tamanho da rajada (em pacotes)	fixo (n)	variável (n_j) (<i>traces</i> de vídeos reais)
Enquadramento de pacotes	uma camada por pacote	variável, de acordo com n_j
Fração de perda	i -ésimo pacote = i -ésima camada	variável, de acordo com o enquadramento
Escalabilidade	FGS (<i>Fine Granular Scalability</i>)	MGS (<i>Medium Grained Scalability</i>)
PSNR do vídeo decodificado	não disponível	disponível

Tabela 5.2: Analítico vs. simulação A: resumo das diferenças

a EL. No modelo de simulação, a perda de um pacote pode afetar mais de uma camada, de acordo com o enquadramento. Por exemplo, na Figura 5.14(a), a perda do pacote de número de sequência igual a 1 significaria a perda de parte da BL e, também, de parte da primeira camada EL; e a perda do pacote de número de sequência igual a 3 implicaria na perda das de parte das camadas EL 1 e 2. Por isso, apresentaremos, mais adiante, novas medidas de interesse para o modelo de simulação. Essas medidas buscam medir separadamente (i) a perda intra-rajada e (ii) a fração de perda de cada camada de vídeo.

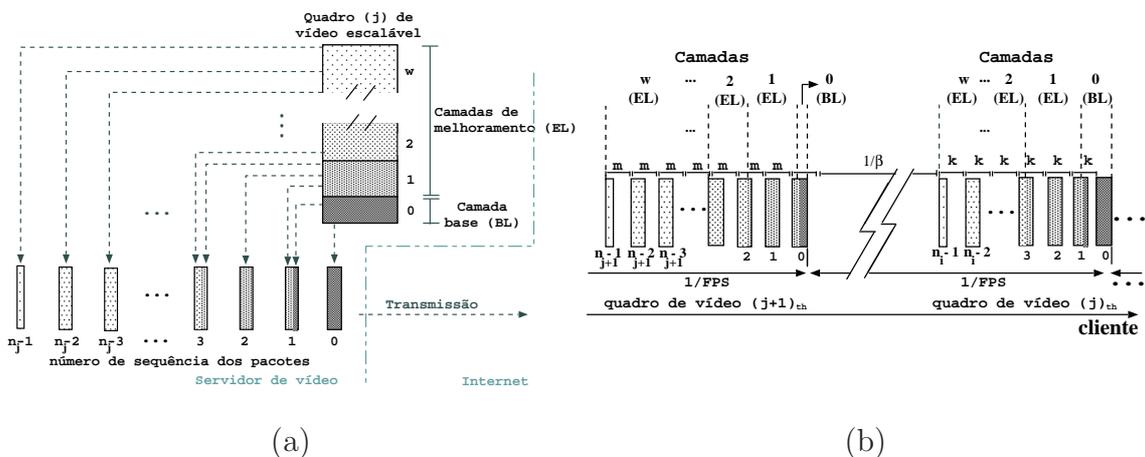


Figura 5.14: (a) Simulação: enquadramento; (b) transmissão de pacotes

A terceira diferença entre os modelos analítico e de simulação reside no esquema de codificação escalável considerado. O modelo analítico supõe escalabilidade FGS (*Fine Granular Scalability*), enquanto o modelo de simulação foi implementado considerando-se escalabilidade MGS [132] (*Medium Grained Scalability*).

A principal razão para considerarmos escalabilidade MGS para o modelo de simulação foi o fato de que o software de referência (JSVM/SVC [133, 134, 91] - *Joint Scalable Video Model/Scalable Video Codec*) usado para codificar/decodificar vídeos escaláveis, em sua versão corrente, não suporta o uso de FGS. Segundo a documentação da ferramenta JSVM/SVC v9.19.3, o MGS passou a ser usado no lugar do FGS, pois este último ainda não contava com uma implementação computacionalmente viável. Como tínhamos o objetivo de construir um modelo de simulação capaz de calcular o PSNR do vídeo recebido no receptor, precisávamos codificar fluxos de vídeo para alimentar o módulo *Video_Source*, e depois decodificarmos o fluxo recebido no receptor. Portanto, o MGS foi escolhido.

A capacidade de obter o PSNR resultante de um fluxo de vídeo, em função do nível de rajada do vídeo emitido, constitui a última diferença entre o modelo analítico e o de simulação, como resumido na Tabela 5.2.

Note que a definição de nível de rajada é diferente para o modelo analítico, onde as rajadas possuem tamanho fixo, e para o modelo de simulação, onde cada quadro pode gerar quantidade diferente de pacotes (rajada de tamanho variável). Embora as duas definições sejam análogas, é importante entender em detalhes o significado de nível de rajada para o modelo de simulação. Para tal, introduzimos a definição a seguir.

Suponha uma sequência de vídeo codificado usando MGS de w camadas. A camada 0 corresponde à BL e as camadas 1 a $w - 1$ correspondem à ELs. Cada quadro de vídeo deste fluxo possuirá exatamente w camadas. Suponha, também, n_j como o número de pacotes necessários para transmitir o quadro j .

A Figura 5.14 (a) ilustra o exemplo para 2 quadros consecutivos de vídeo. Neste exemplo, a camada BL do quadro (j) ocupa todo o pacote 0 e parte do pacote 1. Os pacotes 1, 2 e o início do pacote 3 contém a camada EL 1. Os pacotes $n_j - 2$ e $n_j - 1$ contém a camada EL w . Estes são, portanto, os pacotes de menor importância para a qualidade do vídeo decodificado.

No exemplo da Figura 5.14 (b), k representa o espaçamento entre pacotes do quadro j , e m representa o espaçamento entre pacotes do quadro $j + 1$. k e m podem ser diferentes, já que n_j e n_{j+1} assumem valores arbitrários, de acordo com o *trace* de vídeo. A definição do nível de rajada precisa levar em conta essas diferenças e, por isso, redefinimos b a seguir.

Para o modelo de simulação, o nível de rajada b é definido como

$$b = \frac{\lambda_j}{n_j f} \quad (5.8)$$

onde λ_j é a taxa de envio de pacotes para o quadro j . Como n_j varia para cada quadro j , λ_j deve ser calculado quadro-a-quadro, para que b se mantenha fixo ao longo do tempo: $\lambda_j = bn_j f$.

Por que n_j assume diferentes valores para cada quadro j ? Isso ocorre pois fluxos de vídeo produzem tráfego com comportamento conhecido como VBR (*Variable Bit Rate*), devido à maneira como os codificadores de vídeo realizam a compressão. Primeiramente apenas 1 em cada grupo de quadros (GOP - *Group of Pictures*) pode ser decodificado independentemente de outros. Estes quadros são chamados de quadros I . A diferença na imagem (ex: movimento), desde o último quadro I , gera dados que são comprimidos em um quadro chamado P . Por fim, as diferenças entre imagens de um quadro I e um quadro P são computadas e comprimidas em quadros denominados B . Em geral, portanto, quadros P são menores que quadros I , e quadros B são menores do que quadros P .

Outra fonte de variabilidade no tamanho de quadros é a característica da cena digitalizada. Cenas com muitos detalhes na imagem ou muito movimento entre quadros geram mais dados do que cenas mais simples ou de baixa movimentação.

O módulo *Video_Source* (Figura 5.13) realiza a leitura do quadro de vídeo (j) a ser enquadrado em n_j pacotes. O enquadramento é feito seguindo a mesma ideia apresentada na Figura 5.14(a), ou seja, a camada BL é colocada no(s) primeiro(s) pacote(s) de cada rajada e as camadas EL são enquadradas em ordem decrescente de importância para a qualidade final do vídeo.

Cada pacote é, então, enviado ao módulo *Last_Router*. Caso haja perda (ex: fila cheia), o módulo *Last_Router* computa qual pacote foi perdido e quais camadas serão afetadas, de acordo com o conteúdo do pacote. Estas informações são reunidas para, ao final da simulação, gerarem um registro detalhado do que foi perdido e do

que foi recebido pelo cliente.

Uma vez explicado cada módulo do modelo de simulação A, indicamos como calcular a qualidade objetiva (PSNR) do vídeo recebido pelo cliente.

O modelo de simulação proposto é apenas uma parte de um arcabouço criado com o objetivo de calcular o PSNR resultante de um experimento de fluxo de vídeo, como ilustrado na Figura 5.15, onde cada experimento de fluxo de vídeo escalável é composto de *VII* passos.

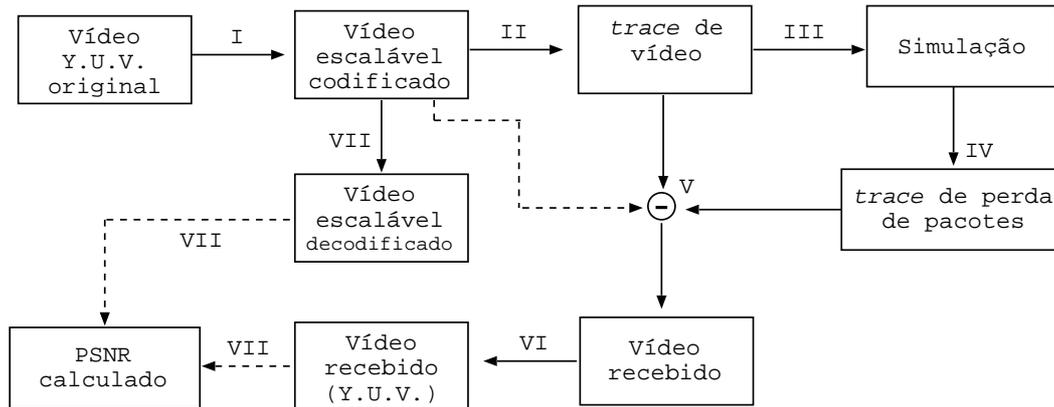


Figura 5.15: Estrutura dos experimentos de simulação: desde (I) a codificação até (VII) o cálculo do PSNR ao final da transmissão

No passo *I*, uma cena de vídeo é codificada usando a ferramenta JSVM/SVC para produzir um vídeo escalável comprimido (codificado). No passo *II*, gera-se um *trace* do vídeo codificado. Este *trace* contém o tamanho em *bytes* de cada camada de cada quadro de vídeo.

A simulação é alimentada, no passo *III*, com o *trace* do vídeo (o módulo *Video-Source* do modelo de simulação implementa, internamente, o esquema de enquadramento proposto). Então, o modelo de simulação é executado, no passo *IV*, gerando um registro de cada pacote perdido e recebido pelo cliente, incluindo seu conteúdo.

O passo *V* consiste em usar o registro de perda de pacotes para extrair, do *trace* de pacotes de vídeo, o que foi perdido durante a transmissão. Com isso é possível usar o software JSMV/SVC para gerar um arquivo de vídeo codificado que representa o que foi recebido pelo cliente.

Este fluxo é decodificado no passo *VI*, e no *VII* é comparado ao vídeo antes de ser transmitido, através do cálculo do PSNR.

5.3.2 Definição das medidas de interesse

A primeira medida de interesse do modelo de simulação A é análoga à obtida no modelo analítico através da Equação 5.7.

Usando o registro de perda de pacotes gerado durante a simulação, é possível calcular a fração de perda ϕ_a como sendo a razão entre o total de pacotes com número de sequência (a) perdidos e o total de pacotes (a) gerados, para $0 \leq a \leq \max(n_j) - 1$, onde $\max(n_j)$ é o maior número de pacotes necessários para transmitir um quadro considerando todo o fluxo de vídeo. A fração total de pacotes perdidos, F , também é calculada.

Uma característica importante do modelo de simulação é o uso de vídeos reais e, com isso, a perda de um pacote i não equivale precisamente à perda da camada de vídeo i , como era considerado no modelo analítico (veja Tabela 5.2). Portanto, é necessário calcular separadamente, para o modelo de simulação, a fração de perda (L_c) de cada camada c de vídeo escalável, como a razão entre o número total t_c de camadas c indecodificáveis no receptor e θ : $L_c = \frac{t_c}{\theta}$, onde θ representa o número de quadros do fluxo transmitido. Note que cada quadro contém exatamente w camadas, a BL e $w - 1$ ELs. Este vídeo contém, portanto, θ unidades de cada uma das i -ésimas camadas ($0 \leq i \leq w - 1$).

Para a avaliação da qualidade do vídeo recebido pelo cliente, usamos o PSNR, que é uma métrica comumente usada para avaliar a qualidade de aplicações de vídeo. O mais comum na literatura, como visto no Capítulo 3, é a apresentação do PSNR quadro-a-quadro ou o cálculo do PSNR médio. Contudo, essas métricas não permitem realizar a avaliação que pretendemos da qualidade de vídeo. Portanto, usaremos outra métrica para avaliar o PSNR, como explicado a seguir.

Suponha uma situação onde 6% dos dados transmitidos são perdidos. No trabalho desenvolvido por esta tese, estamos interessados em mostrar que com o ajuste correto do nível de rajada, os períodos de perda são significativamente menos perceptíveis ao usuário final.

Obviamente, se a transmissão enfrenta 6% de perda, a maioria dos pacotes chegarão ao receptor. Com isso, caso o PSNR médio seja usado como métrica, as diferenças de PSNR medidas durante os períodos de perda seriam diluídas pelos períodos mais longos onde não houve perda. Portanto, como esperado para o va-

lor médio de uma variável aleatória, o cálculo do PSNR médio não permite avaliar períodos curtos onde a qualidade do vídeo pode ficar muito prejudicada devido a perdas.

Por outro lado, o uso do PSNR quadro-a-quadro para vários valores de *nível de rajada* (b) se torna proibitivo, pois a quantidade de informação gerada é muito grande, o que dificulta a análise dos dados.

Para contornar esses problemas, usaremos o histograma do PSNR quadro-a-quadro para avaliar a qualidade do vídeo recebido. Esta métrica permite avaliar a eficiência da nossa proposta através da observação da fração de quadros obtidos para cada intervalo de PSNR.

5.3.3 Parâmetros considerados

Cada experimento de simulação é construído como explicado a seguir.

Primeiramente, define-se o perfil de vídeo a ser simulado. Para cada perfil de vídeo são executadas 12 rodadas de simulação, cada uma com 4000 segundos de duração. Se chega-se ao final de um *trace* de vídeo durante o experimento, inicia-se novamente sua leitura do início.

Foi usado grau de confiança de 95% para o cálculo do intervalo de confiança e, para todos os resultados obtidos, o intervalo de confiança foi de 6% a 11% dos valores médios obtidos. Com o objetivo de obtermos uma apresentação mais clara dos resultados, optamos por não usar as barras de intervalos de confiança nos gráficos em que os resultados são apresentados.

Todos os vídeos usados nos experimentos de simulação (modelo A) foram codificados com os parâmetros apresentados na Tabela 5.3. Todos os vídeos foram gerados com 6 camadas de escalabilidade MGS, com $GOP = IbbbPbbbPbbbPbbb$ ($GOP=16$). A qualidade de cada camada é definida através dos parâmetros QP (*Quantization Parameter*) usados em cada camada. Os mesmos parâmetros de codificação foram usados para todos os perfis (cenas) de vídeos nos estudos feitos através do modelo de simulação A.

Um total de 5 vídeos bem conhecidos foram usados: Foreman, Mobile, Football, Crew e Star Wars IV. Cada vídeo apresenta características distintas de movimentação e complexidade, o que resulta em fluxos de vídeo com padrões de tráfego

	Camada 0 (BL)	Camada 1	Camada 2	Camada 3	Camada 4	Camada 5
QP	40	36	32	29	26	22

Tabela 5.3: Parâmetros de Quantização (QP) usados em todos os perfis de vídeo sensivelmente diferentes, como ilustrado nas Figuras 5.16 e 5.17, e resumido na Tabela 5.4.

É importante notar que o uso de vídeos que geram padrões de tráfego diferentes é importante para analisarmos a viabilidade da ideia proposta para os principais perfis de vídeo.

O vídeo Foreman representa aplicações como tele-jornais ou programas de entrevista, onde há baixa movimentação. Já aplicações de transmissão de esporte, com alta movimentação, são representadas pelo vídeo Football. Nas Figuras 5.16 (a) e (c) vê-se que o tráfego resultante do perfil Foreman possui baixa variabilidade temporal, enquanto o perfil Football resulta em tráfego de variabilidade sensivelmente maior. Esta figura apresenta o número de pacotes por quadro para todo o vídeo. A linha tracejada representa a média de pacotes por quadro.

Durante a descrição do objeto *Video_Source*, comentamos sobre o que causa variabilidade nos tamanhos dos quadros. No perfil Foreman (Figura 5.16 (a)), temos um exemplo de variabilidade gerada por movimentação na cena. Pode-se ver que o número de pacotes por quadro muda muito no final da sequência (a partir do quadro de número 220). Isso ocorre pois ao final da cena, a câmera é movimentada.

Ainda observando a Figura 5.16 (a), é possível, inclusive, identificar cada tipo de quadro: quadros *I*, que geram aproximadamente 20 pacotes cada um; *P* que formam uma linha intermediária de pacotes/quadro e os quadros *B* que formam a linha inferior na figura por serem os menores.

O mesmo padrão também é observado no vídeo Mobile (Figura 5.16 (b)), onde os quadros *I* geram, em média, 50 pacotes; os quadros *P* geram ≈ 35 pacotes e os quadros *B* geram ≈ 9 pacotes. Este vídeo representa um perfil de movimentação média, mas com alta complexidade de imagens. Por essa razão, o algoritmo de codificação não atinge a mesma compressão que no caso anterior (Foreman), produzindo rajadas bem maiores.

Já o vídeo Crew representa uma cena mista, contendo trechos de baixa e de

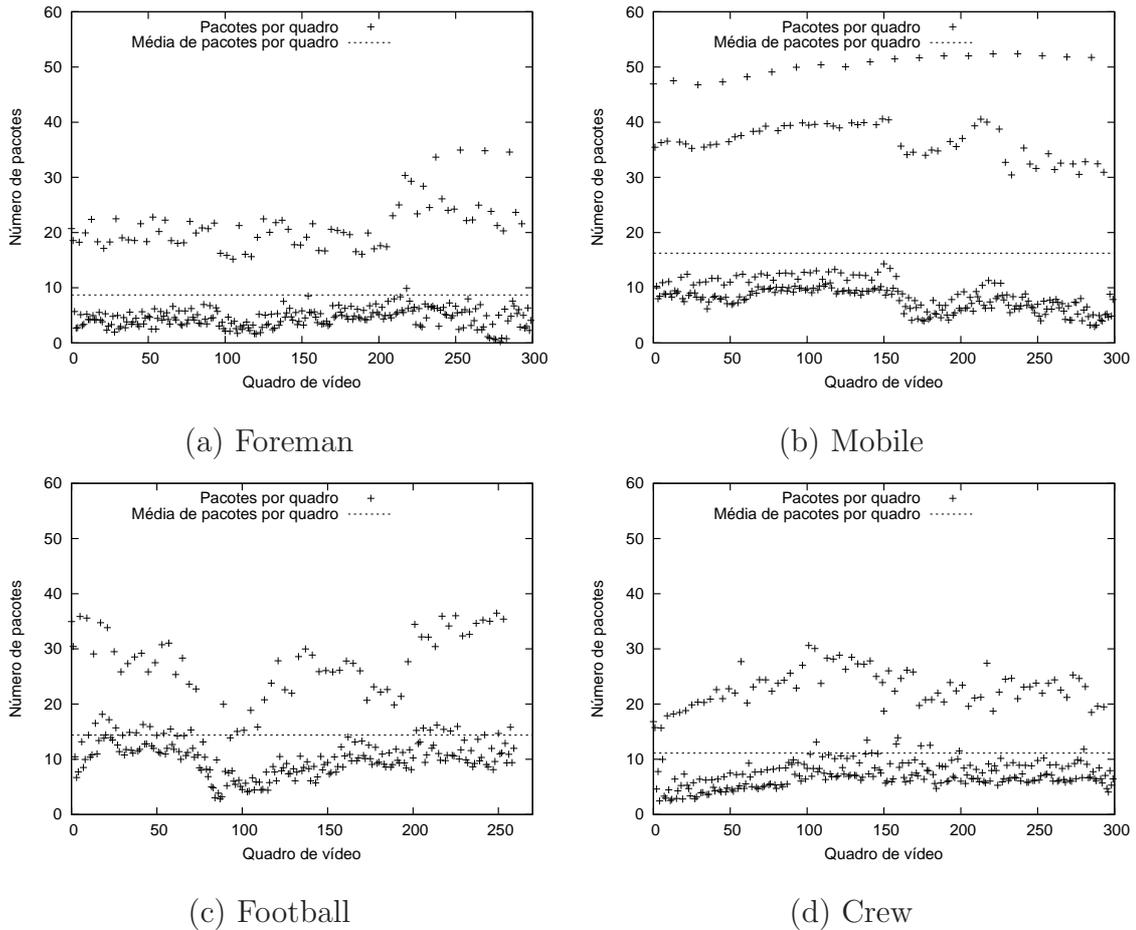


Figura 5.16: Pacotes por quadro para 4 perfis de vídeo escalável

alta movimentação. Este vídeo possui, como peculiaridade, a presença de *flashes* de fotografia que causam mudança repentina em quadros específicos (ex: o quadro onde as fotos são tiradas). A Figura 5.16 (d) mostra como estas fotografias geram quadros isolados de tamanho maior que a média geral.

Por fim, o vídeo *Star Wars IV*, apresentado na Figura 5.17, representa um cenário misto de longa duração, totalizando 1 hora de vídeo obtido do início do filme *Star Wars IV* (note que os perfis de vídeo apresentados, anteriormente, possuem ≈ 10 segundos de duração). Podemos ver na Figura 5.17 que o tráfego gerado por este vídeo alterna períodos de baixa, média e alta movimentação.

Como vimos, diferentes perfis de vídeo geram diferentes padrões de tráfego. Claramente, espera-se que as cargas oferecidas por diferentes perfis também sejam diferentes. Entretanto, desejamos realizar experimentos de simulação, onde a carga oferecida à fila (*Last_Queue*) seja basicamente a mesma, mesmo quando diferentes

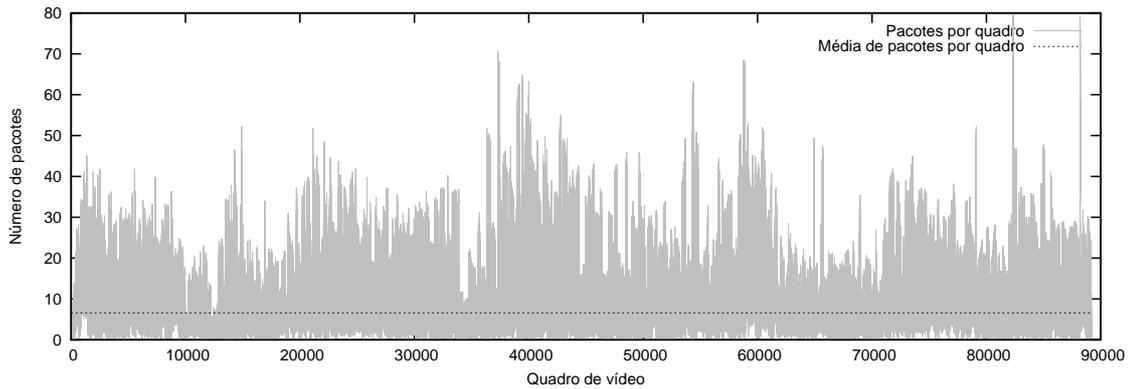


Figura 5.17: Pacotes por quadro para o perfil de vídeo escalável Star Wars IV

perfis são considerados, de forma a facilitar a comparação dos resultados.

A Tabela 5.4 apresenta a taxa média resultante de cada vídeo codificado. De fato, os valores são diferentes. Para garantir que a carga oferecida à fila (*Last-Queue*) seja aproximadamente a mesma, parametrizamos o modelo de simulação, como explicado a seguir.

Video	Duração	FPS	Taxa bits (Mbps)	Resolução	taxa média (pacotes/seg)	Perfil
Foreman	10s	30	3.02	352x288	261.55	Talking head
Mobile	10s	30	5.61	352x288	488.19	Slow panning
Football	8.67s	30	4.97	352x288	431.33	High movement
Crew	10s	30	3.84	352x288	333.49	High scene transition
Star Wars IV	3720s	24	1.8	672x288 (wide)	158.02	Mixed movement

Tabela 5.4: Resumo dos perfis de vídeo

Todos os experimentos de simulação, realizados através do modelo A , foram executados considerando a taxa de serviço do objeto *Last-Queue* de 1500 pacotes por segundo, e a carga foi ajustada para $\rho = 0.99$.

Para alcançar a mesma carga para diferentes perfis de vídeo, ajustamos a carga do tráfego de fundo, no objeto *Cross_Traffic*, para que a soma total do tráfego que chega à fila (*Last_Queue*) seja aproximadamente a mesma (0.99×1500).

Por exemplo, a taxa de pacotes para o vídeo Foreman é de 261.55 pacotes/s (veja Tabela 5.4). Assim, ajustamos a taxa de geração de pacotes de tráfego de fundo (*Cross_Traffic*) para 1224 pacotes/s. Com isso, a carga resultante é de $\rho = \frac{261.55+1224}{1500} = 0.9904$. Analogamente, o vídeo Mobile possui carga de 488.19 pacotes/s. Portanto, ajustamos a taxa do tráfego de fundo para 997 pacotes/s. Com isso, a carga para este vídeo é equivalente a $\rho = \frac{997+488.19}{1500} = 0.9901$. Os outros perfis seguem a mesma ideia.

5.3.4 Resultados e discussão

Considerando o perfil de vídeo Football, as Figuras 5.18 (a) e (b) apresentam, respectivamente, L_c , (ϕ_a) e F . Todas essas métricas são apresentadas em função do nível de rajada.

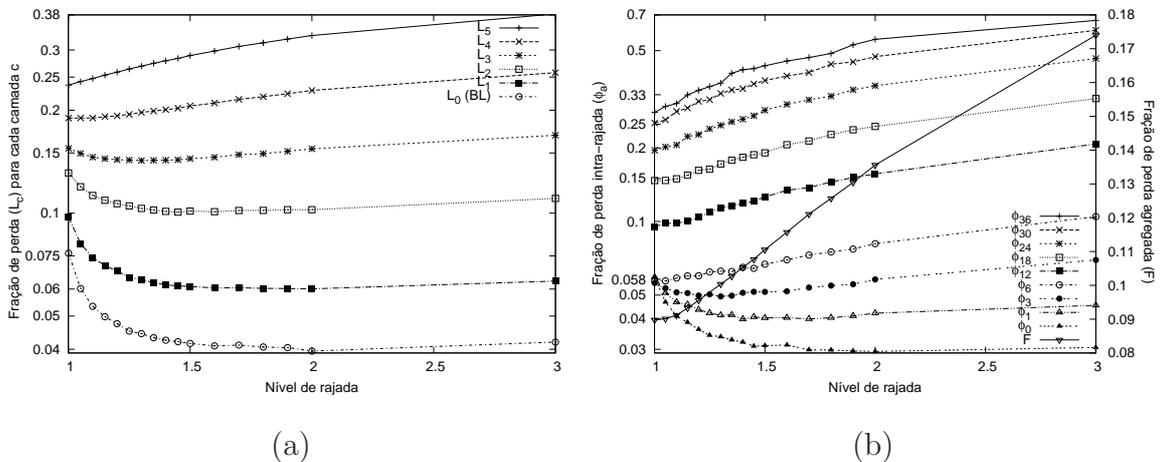


Figura 5.18: Perfil Football: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$

Note que, como apresentado na Figura 5.16, os vídeos reais usados nos experimentos de simulação geram rajadas contendo dezenas de pacotes. Por isso seria potencialmente confuso apresentarmos, para cada experimento, ϕ_a para todos os pacotes a das rajadas, já que teríamos um número muito grande de curvas por gráfico.

Então, optou-se por apresentar a fração de perda (ϕ_a) de apenas alguns dos pacotes. Com isso, os gráficos alcançam maior nível de clareza.

A Figura 5.18 (b), apresenta ϕ_a para $a = \{0, 1, 3, 6, 12, 18, 24, 30, 36\}$. Observa-se que quando $b = 1.5$, ϕ_0 é aproximadamente 50% menor do que quando $b \approx 1$. Além disso, todos os pacotes com $a \leq 3$ também apresentaram melhora no nível de perda para $1 < b \leq 2$.

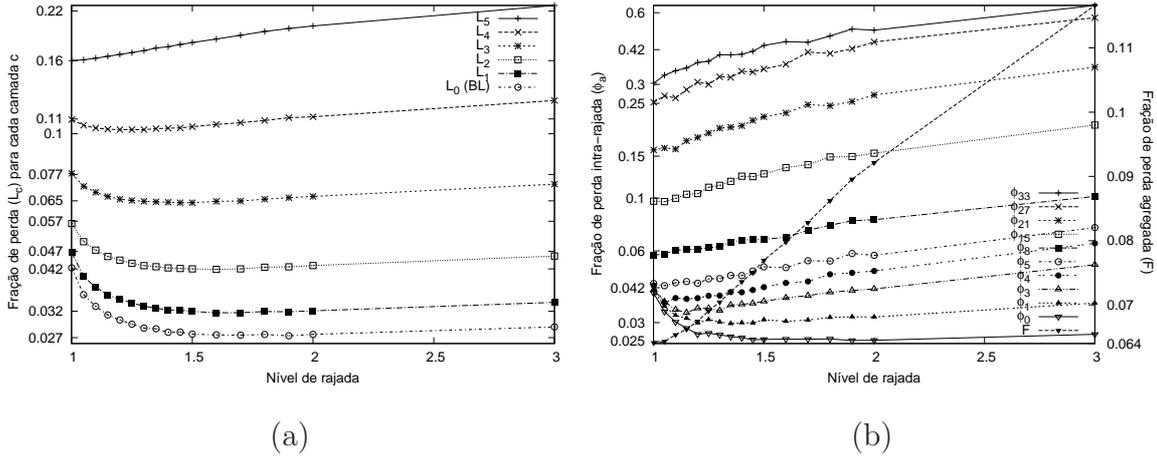


Figura 5.19: Perfil Foreman: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$

No lado direito da Figura 5.18 (b) observa-se que F cresce monotonicamente à medida que b é ajustado para valores maiores que 1. Isso ocorre pois o ajuste de b entre $[1.2, 2]$ gera redução de ϕ_a apenas para $a \leq 3$, ou seja, apenas os primeiros 4 pacotes das rajadas são beneficiados (ex: redução da perda, quando comparado ao caso onde $b \approx 1$) pelo ajuste do nível de rajada. Todos os outros pacotes ($3 < a \leq 36$) sofrem nível mais elevado de perda para qualquer valor de $b > 1$.

Ainda considerando o vídeo Football, os resultados (Figura 5.18 (a)) indicam que L_0 (fração de perda da BL) foi reduzida de 4.2% para $\approx 2.7\%$ quando b foi ajustado de 1 para 1.7, ou seja, redução de 35.7%. Além disso, todas as outras camadas também apresentaram redução na perda, com exceção da camada 5, já que L_5 aumenta de $\approx 16\%$ para $\approx 20\%$ quando b é ajustado de 1 para 1.7.

Resultados análogos, obtidos através de experimentos considerando os perfis de vídeo Foreman, Mobile, Crew e Star Wars IV, são apresentados nas Figuras 5.19, 5.20, 5.21 e 5.22, respectivamente.

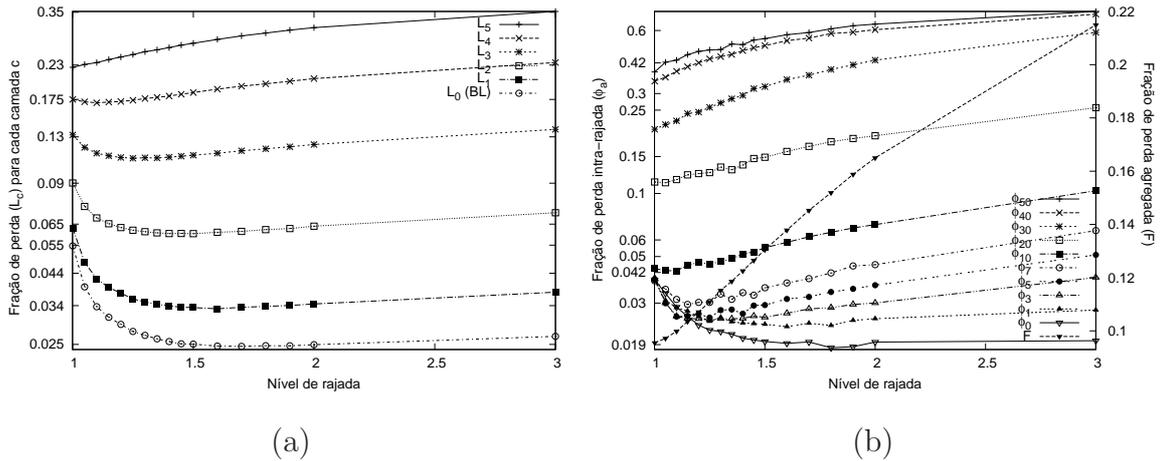


Figura 5.20: Perfil Mobile: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$

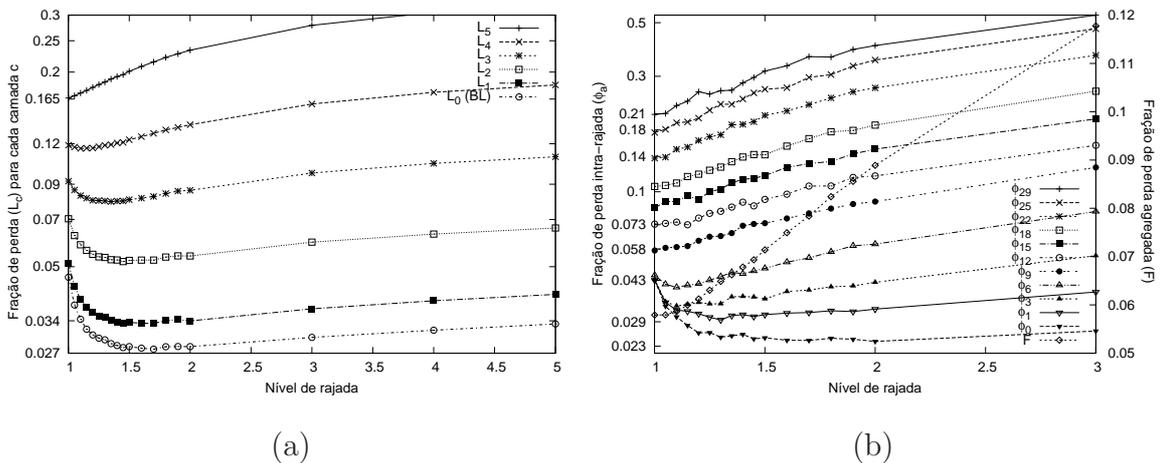


Figura 5.21: Perfil Crew: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$

Dois observações importantes são destacadas para este conjunto de resultados:

(i) O modelo analítico apresentado anteriormente, indica um processo de perda onde, comparado ao caso em que $b \approx 1$, o uso de um determinado b ($1 < b \leq 2$) resulta em melhora no nível de perda dos primeiros pacotes de uma rajada ao custo do aumento do nível de perda dos pacotes subsequentes. O modelo de simulação A indica que este mesmo comportamento também pode ser observado para o cenário de rajadas de comprimento (pacotes) variável, que corresponde a fluxos reais de

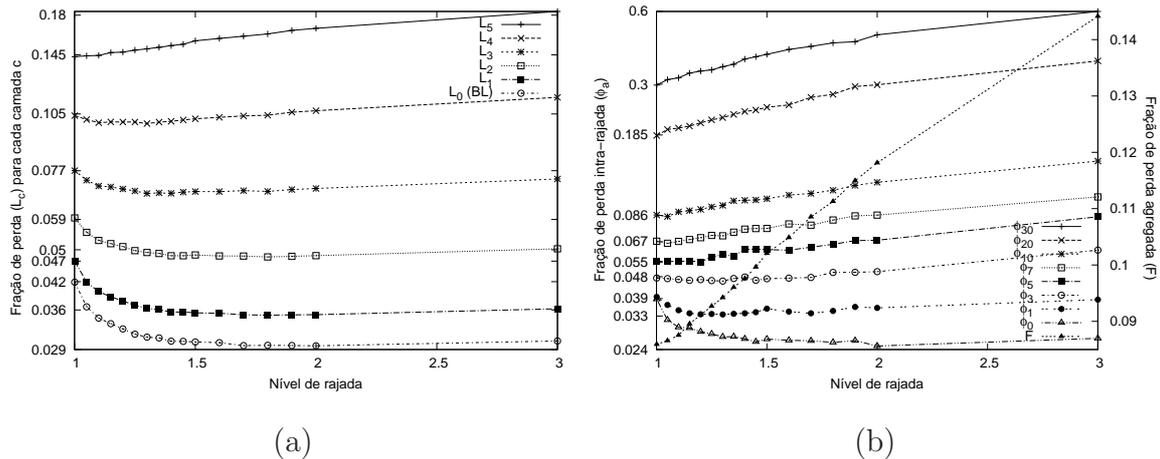


Figura 5.22: Perfil Star Wars IV: (a) Fração de perda (L_c) para cada camada c de vídeo escalável, $q=25$; (b) Fração de perda intra-rajada (ϕ_a) para cada pacote de número de sequência a , $\rho = 0.99$ e $q=25$

vídeo escalável.

(ii) É comum, na literatura, a avaliação do efeito do nível de perda agregada (F) [135, 136, 137, 81] sobre a qualidade de fluxos de vídeo, ao invés de se considerar o processo de perda intra-rajada. O conhecimento geral é que níveis mais elevados de perda (F) sempre resultam em vídeo de qualidade inferior. Contudo, para o caso de vídeos escaláveis, onde cada rajada possui pacotes com diferentes níveis de importância, observamos que o uso de $b > 1$ resulta em sensíveis melhoras em $L_c \forall c \in \{0, 1, 2, 3, 4\}$, apesar do valor de F aumentar com b . Na Seção 5.3.4, discutiremos como este novo resultado se relaciona com as melhorias na qualidade objetiva do vídeo (PSNR).

A princípio, a observação (ii) pode parecer contra-intuitiva. Contudo, seu entendimento é simples: basta recordar a relação de dependência entre camadas (intra-quadro) intrínseca à codificação escalável (Figura 5.23(a), em vermelho), e entre quadros de vídeo (inter-quadro) (Figura 5.23(b), em azul) .

Por exemplo, a decodificação da camada 1 depende da correta decodificação da camada 0 (BL). De maneira análoga, a camada 2 depende da 1 e assim por diante. Portanto, a redução da perda de pacotes BL beneficia, em efeito cascata.

Além disso, a hierarquia entre quadros (I, P, B) também colabora para o comportamento destacado na Observação (ii). Por exemplo, a incapacidade de decodificar a BL de um quadro I inviabiliza a decodificação de todo o GOP. Portanto, a redução

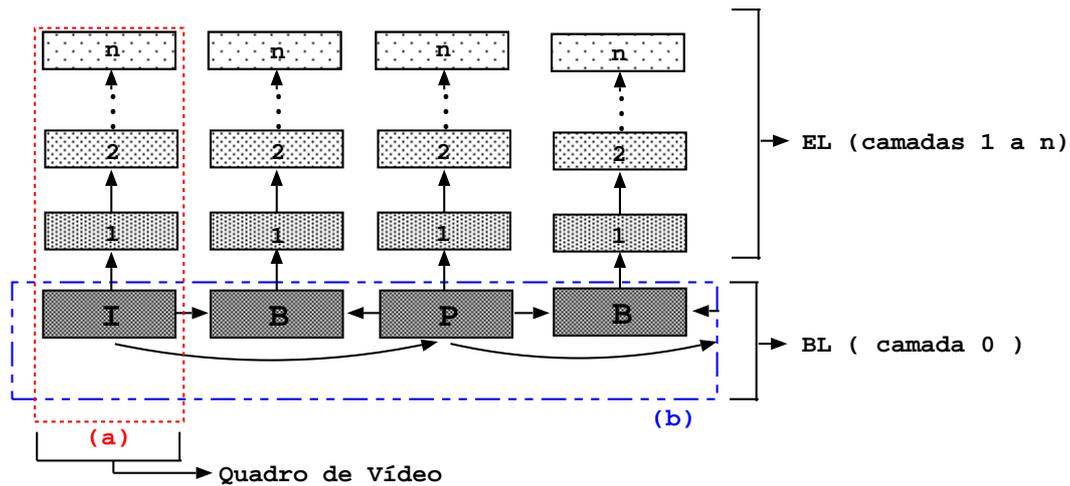


Figura 5.23: Dependência intra-quadro (a) e inter-quadro (b)

da perda de pacotes contendo a camada BL também gera benefícios que alcançam quadros hierarquicamente inferiores.

Estudo do ganho de qualidade de vídeo alcançado

A seguir, apresentaremos histogramas do PSNR quadro-a-quadro para os perfis de vídeo considerados pelo modelo de simulação *A*. Estes resultados são importantes pois, enquanto as métricas apresentadas na seção anterior não permitem avaliar objetivamente a qualidade do vídeo percebida pelos usuários, os resultados desta seção permitem esta análise através do uso do PSNR. Avaliamos como a priorização de pacotes mais importantes, alcançada pela nossa proposta, resulta em ganhos de qualidade do vídeo decodificado no cliente.

A Figura 5.24 apresenta os resultados para o perfil Football. Repare que esta figura apresenta a fração de quadros com PSNR dentro da faixa $[x - 1, x + 1[$ dB, para valores de x (PSNR), em dB, que correspondem a três importantes faixas de qualidade [138, 81] percebida pelo usuário: (i) ruim, pobre e regular, quando $x < 30dB$; (ii) bom (aceitável), quando $30dB < x < 38dB$; e (ii) excelente para $x > 38dB$.

As barras vermelhas representam o ponto referência, onde o tráfego tem o nível de rajada $b \approx 1$, ou seja, tráfego suave. Também são apresentados outros casos de b , que são representados por pontos de diferentes cores e formatos.

Em específico, a linha azul tracejada representa uma boa escolha para b . Para

a Figura 5.24, a curva selecionada foi para $b = 1.6$. Comparando as curvas onde $b = 1.6$ com $b = 1.001$, percebe-se o efeito positivo na qualidade com o uso de um nível de rajada bem ajustado.

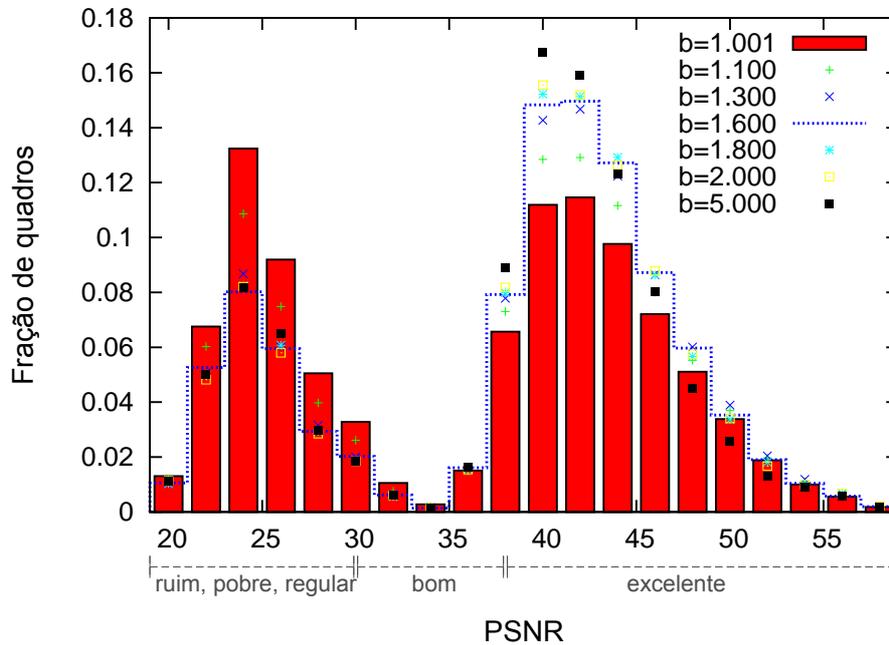


Figura 5.24: Perfil Football: histograma do PSNR quadro-a-quadro

As Figuras 5.18 (a) e (b), apresentadas na seção anterior, mostraram que o uso de $b = 1.6$ resultou em um importante nível de priorização de pacotes mais importantes para a qualidade do vídeo (menor perda nos primeiros pacotes das rajadas). Os resultados, mostrados na Figura 5.24, são importantes pois mostram que a priorização alcançada pela nossa proposta aponta uma significativa melhora do PSNR dos quadros de vídeo.

Por exemplo, a fração de quadros com $\text{PSNR} \approx 24\text{dB}$ diminui de 13.7% quando $b = 1.001$ para 8% quando $b = 1.6$, ou seja, uma redução de 41% dos quadros de vídeo com $\text{PSNR} \approx 24\text{dB}$. Também houve diminuição significativa na fração de quadros com PSNR dentro de outras faixas pertencentes ao grupo de baixa qualidade. Isso ocorre porque o correto ajuste de b (ex: 1.6) resulta em um padrão de perda onde elas ocorrem com maior probabilidade em pacotes de menor importância para a qualidade do vídeo. Como consequência, a qualidade do vídeo é menos afetada pelas perdas. Este é um típico exemplo do resultado que buscamos com nossa proposta.

Além disso, uma vez que o cliente recebe maior quantidade de quadros de maior

prioridade, é possível decodificar mais ELs, o que causa aumento na fração de quadros com PSNR classificados como muito bom ou excelente. Na Figura 5.24, por exemplo, quando $b = 1.6$, o aumento da fração de quadros com PSNR entre $37dB$ e $49dB$ é explicada pela migração de parte dos quadros com PSNR entre $21dB$ e $32dB$ para a área de alta qualidade ($37dB$ a $49dB$). A região com $PSNR \geq 50dB$ permaneceu praticamente inalterada para $b = 1.6$.

Note que não é apropriado ajustar b para valores maiores do que 1.6. Por exemplo, quando $b = 5$ observa-se comportamento menos vantajoso, do que quando $b = 1.6$, ou seja, a fração de quadros com PSNR pobre ($29 - 30dB$) volta a aumentar, enquanto ocorre diminuição de até 20% da fração de quadros com PSNR de mais alta qualidade ($PSNR > 45dB$).

Como a deterioração na qualidade é ainda maior para $b > 5$, não apresentamos resultados para estes casos. Abaixo, recordamos o fenômeno que está por trás da inviabilidade do uso de *níveis de rajada* ainda mais altos.

A partir de um ponto (ex: $b = 2$), o aumento do nível de rajada não diminui a perda dos pacotes do início das rajadas e aumenta a perda dos pacotes subsequentes. Como comentado na Seção 5.2.5, este fenômeno é bem conhecido pela literatura [122, 123, 124].

A seguir, observe a Figura 5.25 onde apresentamos resultados de PSNR para o perfil de vídeo Crew.

De modo geral, os resultados são similares aos anteriores, com a diferença principal de que o aumento do número de quadros com excelente qualidade se deu não só pela diminuição dos quadros com qualidade ruim, pobre e regular, mas também pela diminuição de quadros com qualidade boa ($[30db - 37B]$). Outra diferença é que o ponto escolhido como melhor configuração de nível de rajada foi de $b = 1.5$, contra $b = 1.6$ para o perfil Football.

Além disso, assim como no perfil anterior, observamos que o uso de $b = 5$ gera resultados piores do que o melhor caso. Observe, na Figura 5.25, que a escolha de $b = 5$ resulta na redução da fração de quadros com qualidade superior a $47dB$, quando comparado com $b = 1.5$ e $b = 1.001$. Portanto, não há incentivo para se aumentar b além do ponto escolhido ($b = 1.5$).

Neste caso, é importante entender que o aumento do número de quadros com

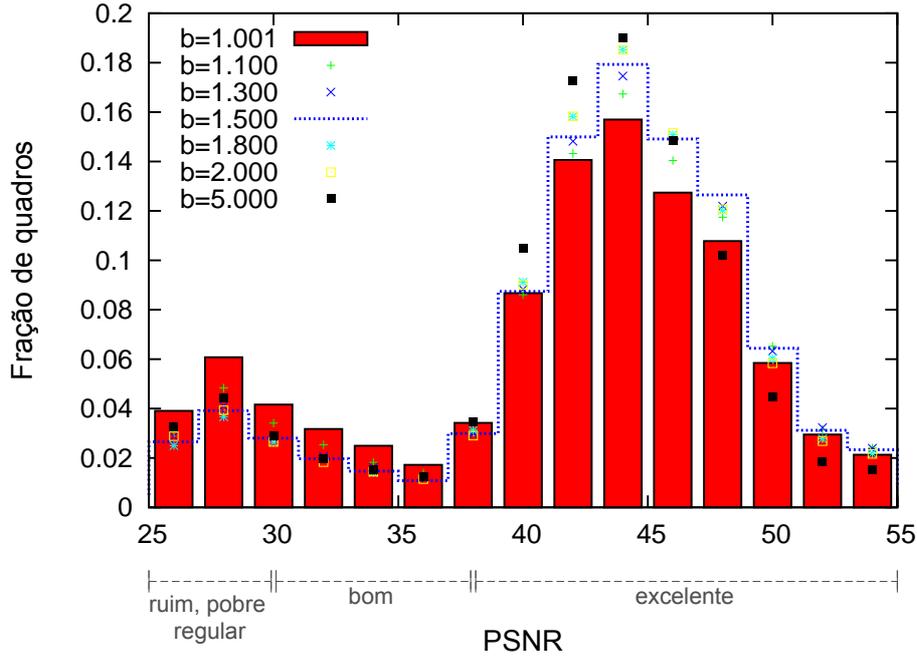


Figura 5.25: Perfil Crew: histograma do PSNR quadro-a-quadro

PSNR entre $39dB$ e $46dB$ se deu, em boa parte, devido à migração dos quadros com qualidade acima de $46dB$.

O próximo perfil de vídeo a ser analisado é o de Foreman, apresentado na Figura 5.26. A primeira diferença, em comparação aos perfis anteriores, está no formato da curva do histograma apresentado para este perfil. É importante destacar que essa diferença é resultado da combinação entre a parametrização do codificador, mostrado na Tabela 5.3, e as características do perfil de vídeo em questão. No caso do Foreman, onde o perfil é de baixa movimentação, o algoritmo de compressão consegue melhores resultados do que para perfis de maior movimentação. Por isso, observamos que a maior parte dos quadros possui qualidade boa ou excelente (ex: $PSNR > 36$).

O fato de que o vídeo recebido possui apenas uma pequena fração dos quadros com qualidade menor do que $30dB$, pode, a princípio, induzir à impressão de que não há a necessidade de ajustar o nível de rajada, já que a qualidade já seria satisfatória. Contudo, os resultados mostram que a escolha de $b = 1.3$ resultou em uma qualidade muito melhor. A fração de quadros de qualidade inferior a $36dB$ diminuiu de 25% a 30%, levando ao aumento do número de quadros com qualidade excelente ($[42dB - 50dB]$).

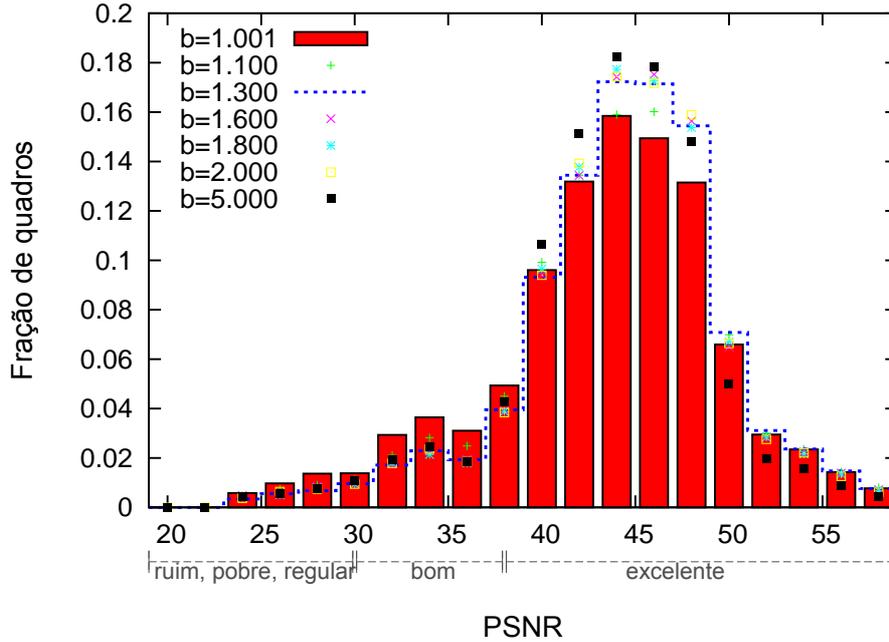


Figura 5.26: Foreman: histograma do PSNR quadro-a-quadro

O perfil Mobile, que representa mobilidade moderada e alta complexidade de imagens, é apresentado na Figura 5.27. Nota-se que a qualidade, devido às características deste perfil, é muito mais sensível ao nível de rajada do que os perfis anteriores, ou seja, o ajuste de b para 1.6 já resulta na deterioração do número na fração de quadros com $PSNR > 42dB$.

A princípio, escolhemos $b = 1.3$ como um bom valor para o nível de rajada. Comparado com o uso de $b \approx 1$, o uso de $b = 1.3$ gera redução de 25% a 50% nas frações de quadros com qualidade inferior a $31dB$, e o respectivo aumento nas frações de quadros com boa e excelente qualidade ($[35dB - 43dB]$). Todavia, observe que houve redução na fração de quadros com $PSNR > 44$.

Uma escolha mais conservadora seria de $b = 1.1$, onde não há diminuição do número de quadros com $PSNR > 44dB$. Mas, em contrapartida, a fração de quadros de baixa qualidade não é tão reduzida quanto na configuração $b = 1.3$.

Quando $b = 2$ ou $b = 5$, há forte redução na fração de quadros com qualidade superior a $39dB$. Os quadros desta faixa de qualidade migraram para a faixa de $[33dB - 38dB]$, indicando que parte importante dos quadros sofreram piora de qualidade. Além disso, estes valores de b resultaram no aumento do número de quadros com $PSNR < 31dB$, representando aumento no número de quadros com qualidade

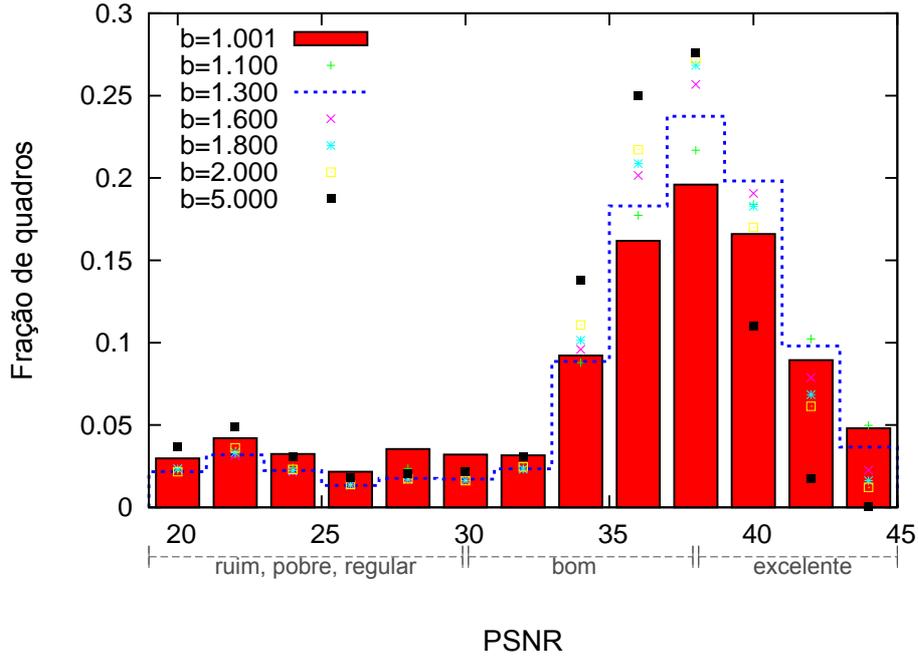


Figura 5.27: Perfil Mobile: histograma do PSNR quadro-a-quadro

ruim, pobre e regular.

O perfil de vídeo Star Wars IV, usado para representar cenários de fluxo de vídeo de longa duração com cenas mistas de baixa e alta movimentação, é apresentado na Figura 5.28, onde observa-se resultados análogos aos comentados para os outros perfis de vídeo discutidos anteriormente nesta seção.

Os resultados apresentados nesta seção nos permitiram concluir que o padrão de perda obtido através do ajuste correto do nível de rajada de fato resulta em significativa melhora na qualidade do vídeo recebido pelo cliente.

Para os parâmetros experimentados, b deve ser escolhido no intervalo [1.3 a 1.6] para obtenção de um vídeo de melhor qualidade. A escolha do melhor valor para b depende do estado da rede, da parametrização da codificação escolhida e do perfil do vídeo em questão.

O estudo de um algoritmo para selecionar um bom valor para b ou de um algoritmo que ajuste b dinamicamente, de acordo com o cenário em questão e suas respectivas flutuações, deixado como trabalhos futuros.

Há outro ponto que consideramos importante avaliar ainda nesta tese. O modelo de simulação e o modelo analítico, apresentados até agora, supõem uma fila FIFO, com serviço exponencial, no roteador do último salto da Internet. Na próxima seção,

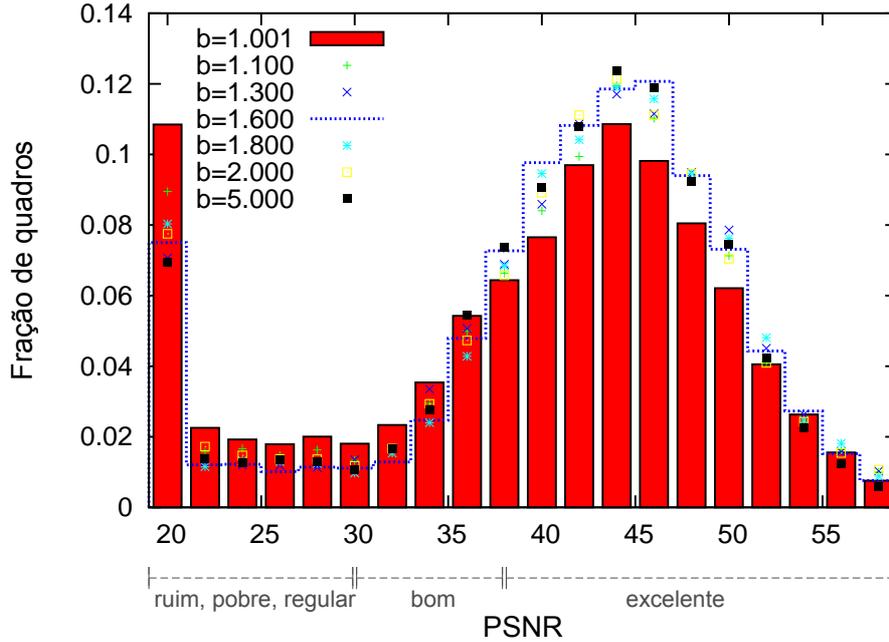


Figura 5.28: StarWars IV: histograma do PSNR quadro-a-quadro

introduziremos uma variação ao modelo de simulação A, onde o módulo (*Last_Queue*) é substituído por um serviço mais realista.

5.4 Modelo de simulação B: vídeos reais, traces de uma rede EVDO

Nesta seção, estudamos o processo de perda intra-rajada em função do nível de rajada (b), considerando que o cliente está conectado à Internet através de uma rede sem fio EVDO.

O principal objetivo deste estudo é avaliar se o padrão de perda observado, através dos modelos analítico e de simulação A, são válidos quando o modelo para o roteador do último salto do caminho fonte-destino reflete um cenário mais realista. Para isso, apresentamos a visão geral ilustrada na Figura 5.29.

5.4.1 Descrição do modelo

O modelo de simulação B é uma variação do modelo apresentado na Seção 5.3. Nesta variação, o módulo *Last_Queue* foi substituído pelo módulo *EVDO_Router*.

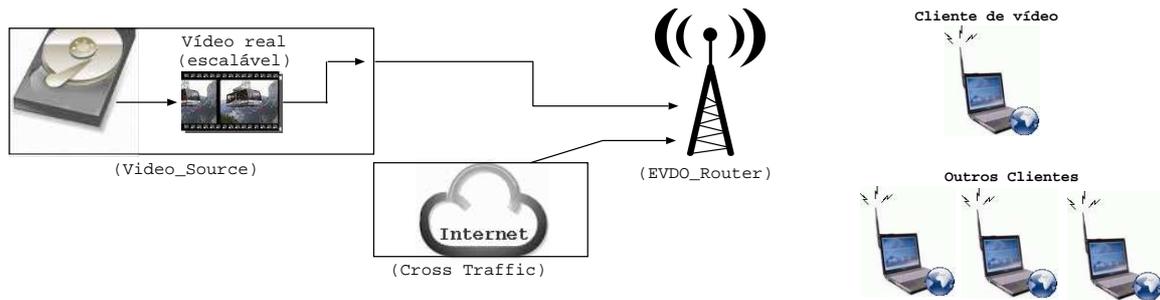


Figura 5.29: Visão geral do modelo de simulação B

O módulo *EVDO_Router* é alimentado com *traces* reais, obtidos através de medições realizadas em uma rede EVDO rev.0, pertencente à empresa Qualcomm, localizada na cidade de San Diego, Califórnia, EUA. O perfil de mobilidade é veicular suburbano.

Os arquivos contendo as medições da qualidade (SINR) do canal, a cada *slot* de tempo, foram gentilmente cedidos pelos autores do artigo de Attar *et al.* [139]. Foram cedidos 27 *traces*, cada um contendo medições com duração aproximada de 4.5 horas.

Para cada experimento executado, escolhe-se aleatoriamente, para cada usuário, um *trace* para representar a flutuação da qualidade de seu canal. O escalonamento de serviço para os usuários é dado de acordo com o algoritmo PFS (*Proportional Fair Scheduler*) apresentado na Seção 4.1.3.

O modelo de simulação B é capaz de prover as mesmas medidas de interesse que o modelo A, inclusive resultados de PSNR para os vídeos recebidos pelo receptor. Entretanto, para deixar esta seção mais sucinta, optamos por apresentar apenas a métrica L_c em função de b . Recorde que L_c representa a fração de camadas c indecodificáveis no receptor, como definido na Seção 5.3.2.

5.4.2 Parâmetros considerados

Cada experimento de simulação é construído como explicado a seguir.

Primeiramente, define-se o perfil de vídeo a ser simulado. Para cada perfil de vídeo são executadas 12 rodadas de simulação, cada uma com 4000 segundos de duração. Se um *trace* de vídeo ou da rede chega ao final durante o experimento, sua leitura inicia-se novamente. Um *trace* EVDO é aleatoriamente escolhido para

alimentar cada rodada de simulação (sem repetição).

O intervalo de confiança (95%) varia de 5% a 10% em torno dos valores médios obtidos para as métricas. Com o objetivo de obtermos uma apresentação mais clara dos resultados, optamos por não apresentar as barras de intervalos de confiança nos gráficos em que os resultados são apresentados.

A codificação dos vídeos precisou ser refeita para os experimentos do modelo B, como explicado abaixo.

No Capítulo 2, comentamos que a capacidade máxima de uma rede EVDO rev.0 é de 2.4Mbps. Entretanto, devido às variações na qualidade do canal, a taxa média de serviço é notadamente mais baixa que a taxa nominal. Para o conjunto de *traces* EVDO que usamos em nossos experimentos, a taxa média de serviço foi de ≈ 1.38 Mbps. Entretanto, os vídeos usados nos experimentos do modelo A possuem taxas bem acima deste valor. Por esta razão, os vídeos precisaram ser recodificados com o objetivo de obter taxas que ofereçam cargas compatíveis com a capacidade de serviço.

Em específico, buscamos escolher parâmetros de quantização (QP) que resultassem em cargas de 0.98 a 0.99. A Tabela 5.5 apresenta os valores de QP usados para cada perfil de vídeo e a respectiva taxa média de cada fluxo gerado.

	(BL)	EL_1	EL_2	EL_3	EL_4	EL_5	(Taxa média)
QP (Football)	45	42	40	38	35	33.5	1373.30Kbps
QP (Foreman)	38	36	33	30	29	27	1354.00Kbps
QP (Crew)	40	37	35	33	31	29.50	1363.50Kbps
QP (Mobile)	44	42	40	38	36	34.60	1354.50Kbps
QP (Star Wars IV)	44	41	39	37	35	33	1371.80Mbps

Tabela 5.5: Parâmetros de Quantização (QP) usados e a taxa média resultante para cada perfil de vídeo

Por exemplo, o perfil Football gera um fluxo de taxa média de 1373.3Kbps e, então, a carga oferecida por este vídeo é de $\frac{1373.3}{1380} = 0.995$.

A capacidade do *buffer* do módulo *EVDO_Router* foi configurada para $q = 25$ pacotes, assim como nos experimentos realizados com o modelo A. É importante ressaltar que redes 3G, como a rede EVDO, considerada nesta seção, oferecem, em

geral, filas diferentes para cada usuário conectado à rede sem fio. Em outras palavras, não há compartilhamento de fila. Como veremos nos resultados do modelo B, esta característica aponta resultados ainda melhores do que os encontrados para os casos onde a fila é compartilhada (modelos analítico e de simulação A).

Embora a fila do roteador EVDO não seja compartilhada entre os clientes, é importante estudar cenários onde a rede é compartilhada por outros usuários para analisar a interferência do compartilhamento do canal entre usuários. Consideramos, portanto, dois cenários para cada resultado apresentado na seção seguinte: com e sem disputa. No cenário sem disputa, supomos que apenas o cliente do fluxo de vídeo usa a célula EVDO ao qual está conectado. No caso com disputa, consideramos 3 outros usuários para os quais sempre há dados a serem servidos.

Para o cliente de vídeo, a presença de outros clientes concorrendo por recursos de sua célula é percebida apenas como variações na taxa de serviço, já que as filas do módulo *EVDO_Router* não são compartilhadas.

Veremos que os resultados com e sem disputa apresentam padrão de L_c semelhantes, embora o valor absoluto das métricas obtidas seja bem diferente.

5.4.3 Resultados e discussão

Nesta seção apresentamos os resultados para dois cenários considerados para o modelo de simulação B: (a) sem disputa e (b) com disputa.

As Figuras 5.30 (a) e (b) apresentam L_c em função de b para o perfil de vídeo Football nos cenários sem disputa e com disputa, respectivamente.

Para o caso sem disputa, o fluxo que usou o menor nível de rajada ($b \approx 1$) resultou em $L_0 = 7.8\%$, $L_1 = 8.2\%$, $L_2 = 9\%$, $L_3 = 9.1\%$, $L_4 = 9.5\%$, $L_5 = 9.8\%$. Por outro lado, quando $b = 2$, $L_0 = 3\%$, $L_1 = 5\%$, $L_2 = 5\%$, $L_3 = 7\%$, $L_4 = 9\%$ e $L_5 = 12.1\%$. Com isso, vemos que o ajuste $b = 2$ resultou em 61% de redução em L_0 . Reduções menores, mas também significativas, ocorreram em L_1 , L_2 , L_3 e L_4 . Já L_5 sofreu acréscimo de 12%.

De fato, esses resultados são análogos aos apresentados na Figura 5.18 (a), o que nos aponta para o mesmo processo de perda observado nos estudos realizados nas seções anteriores. O interessante é que a redução da perda das primeiras camadas (ex: BL) foi ainda maior para o modelo B (ex: 61%), quando comparado ao modelo A

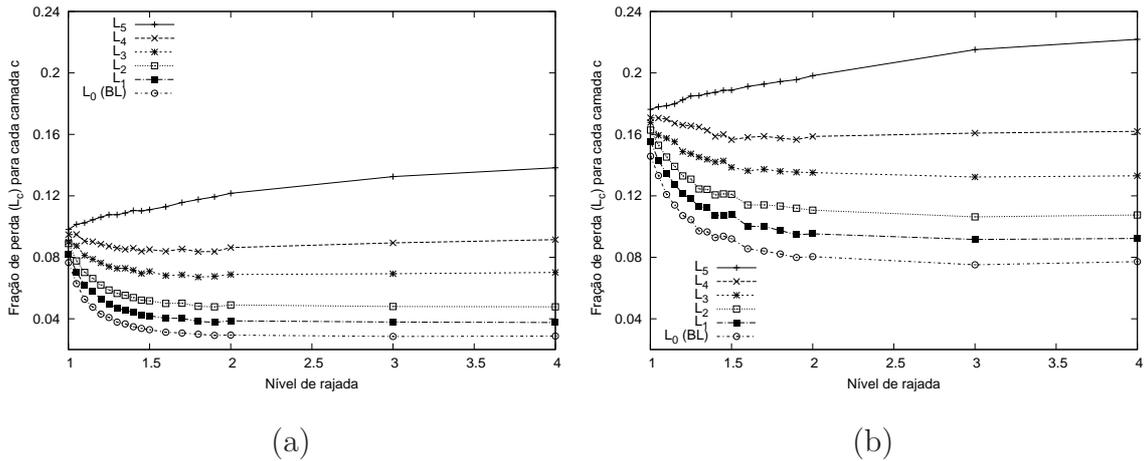


Figura 5.30: Perfil Football ($q=25$, $\rho \approx 0.99$): (a) Sem Disputa, (b) Com Disputa

(ex: 41%). Um comportamento que ocorreu de maneira bastante similar à observada através dos modelos anteriores é o aumento da perda observada nas últimas camadas.

Ainda no perfil Football, mas considerando o cenário com disputa (Figura 5.30 (b)), observa-se comportamento análogo, com a diferença de que o ponto de menor L_0 foi alcançado quando $b = 3$. Além disso, os valores absolutos obtidos foram aproximadamente dobrados, indicando que a disputa degradou consideravelmente a qualidade de serviço. Mesmo assim, o ajuste correto de b causa o mesmo tipo de efeito benéfico observado no caso anterior: redução da perda das camadas mais importantes, ao custo do aumento da perda na(s) última(s) camada(s).

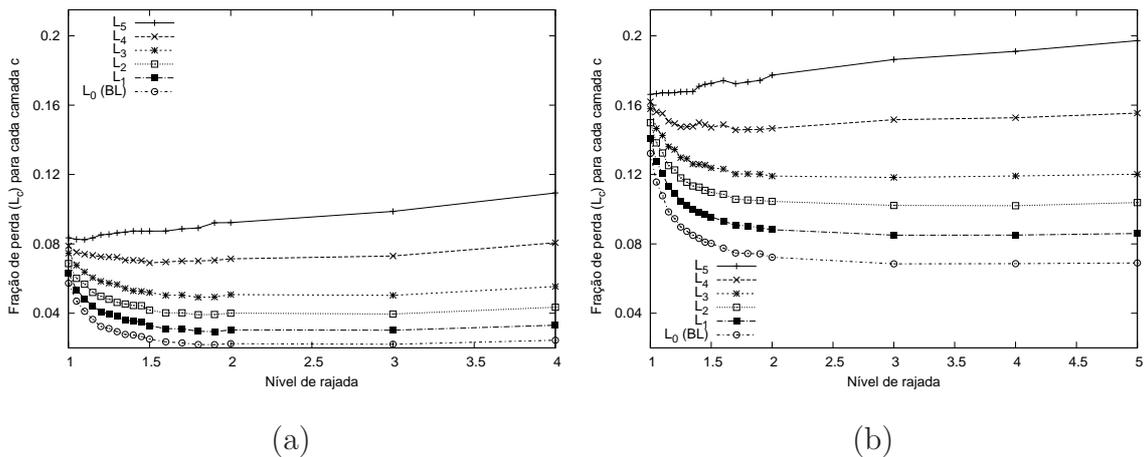


Figura 5.31: Perfil Crew ($q=25$, $\rho \approx 0.99$): (a) Sem Disputa, (b) Com Disputa

Resultados semelhantes para ambos os cenários de disputa são apresentados para os outros perfis de vídeo nas Figuras 5.31, 5.32 e 5.33.

Embora os resultados desta seção tenham apontado padrão de perda análogo ao observado nos modelos analítico e de simulação A, duas diferenças principais podem ser destacadas.

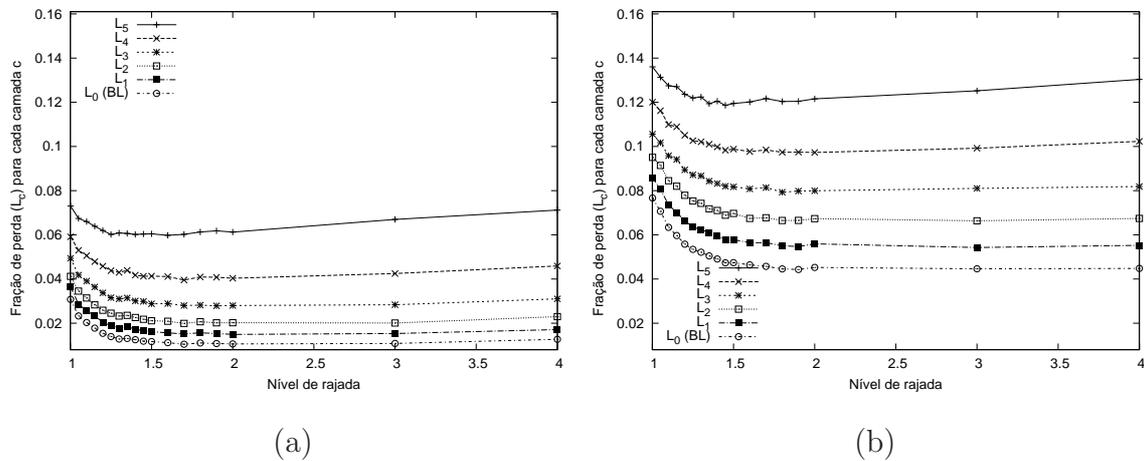
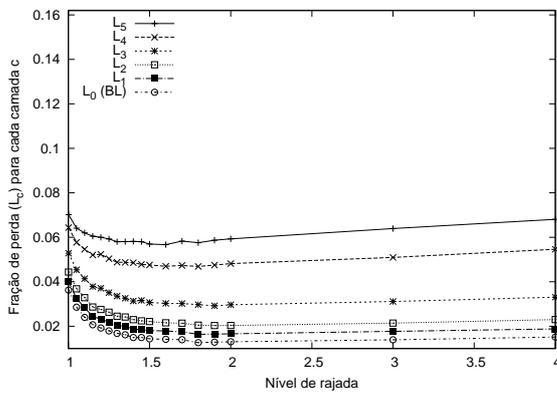


Figura 5.32: Perfil Mobile ($q=25$, $\rho \approx 0.98$): (a) Sem Disputa, (b) Com Disputa

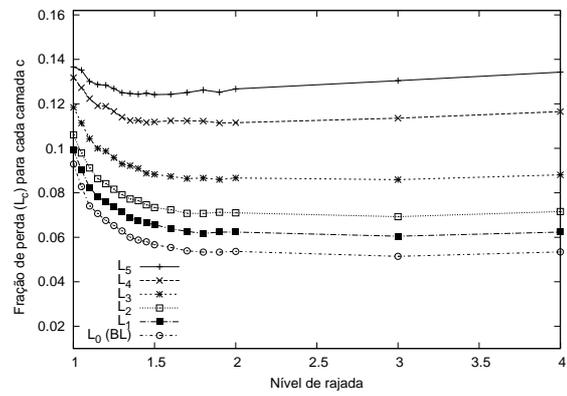
Primeiramente, se compararmos os resultados obtidos nesta seção com os do modelo de simulação A, observamos melhora na priorização das primeiras camadas. Houve até 61% de redução na fração de camadas indecodificáveis no cliente, contra até 40% para o modelo de simulação A.

Outra diferença interessante é que os resultados do modelo de simulação B indicam que há mais liberdade para ajuste do nível de rajada. Observe que é possível aumentar a priorização das primeiras camadas para valores de b que se aproximam de 2. No modelo de simulação A, obteve-se valores de b entre 1.1 e 1.6.

Atribuímos as duas diferenças comentadas acima ao fato de que, segundo o padrão EVDO, mesmo nos casos onde há disputa, cada cliente possui uma fila dedicada[49, 114, 140]. Como o tráfego emitido pelo servidor de vídeo é o único que chega à fila do cliente, não há interferência de tráfego de fundo na sua ocupação. Por isso, o ajuste do nível de rajada do tráfego de vídeo afeta mais significativamente o processo de perda observado na fila em questão.



(a)



(b)

Figura 5.33: Perfil Foreman ($q=25$, $\rho \approx 0.98$): (a) Sem Disputa, (b) Com Disputa

Capítulo 6

Conclusões e trabalhos futuros

O trabalho desenvolvido por esta tese foi dividido em duas partes principais, apresentadas nos Capítulos 4 e 5. Para manter o mesmo padrão de organização, sintetizamos separadamente as respectivas conclusões e trabalhos futuros nas Seções 6.1 e 6.2.

6.1 Sobre a justiça em redes sem fio de longa distância

No Capítulo 4 foi apresentada a parte da tese que trata do desempenho e da justiça do padrão EVDO, onde foram desenvolvidos modelos e estudos que, até onde sabemos, são os primeiros deste tipo a aparecerem na literatura. Um exemplo das diferenças de nosso trabalho em relação aos relacionados é o fato de termos como foco métricas como a vazão e o atraso considerando nível de usuário. Outro exemplo é o fato de conseguirmos quantificar as variações na qualidade de serviço e na justiça devido a variações da camada física para um cenário muito comum de uso de redes sem fio de longa distância: a coexistência de usuários fixos e usuários móveis.

Apesar de a vazão e o atraso do sistema disponibilizado por redes como o EVDO para serviços *web* terem se mostrado normalmente adequados, a justiça em relação ao *goodput* e ao atraso se mostrou inadequada para este tipo de aplicação. As contribuições obtidas até este ponto foram publicadas no Simpósio Brasileiro de Redes de Computadores, em 2006 [12].

Nossos experimentos também mostraram que a justiça é ainda mais degradada quando consideramos um cenário misto de usuários estáticos e móveis, em com-

paração aos cenários onde não há mobilidade de usuários. As contribuições trazidas por esta parte de nosso trabalho foram publicadas no *IEEE Personal Indoor Mobile Radio Communications Conference (PIMRC)*, ainda em 2006 [13], onde consideramos cenários de usuários estáticos. Um estudo com cenários mistos de usuários estáticos e móveis foi publicado *IEEE Wireless Communications and Networking Conference (WCNC)*, em 2008 [15].

Observamos que este problema é uma questão intrínseca ao sistema. Além disso, discutimos que é difícil contornar esta deficiência através do ajuste de parâmetros nos algoritmos de escalonamento sem que haja uma forte degradação na vazão agregada da rede, o que também não é desejado.

Discutimos que a justiça estudada em trabalhos relacionados é analisada através de métricas como o número de *slots* ou de oportunidades de transmissão alocados para cada usuário (justiça em nível de rede). A partir desta discussão, realizamos um estudo da justiça tomando como métricas o *goodput* e o atraso observados pelos usuários (justiça em nível de usuário). Através desta análise, observamos que a justiça quando analisada como tradicionalmente feito pelos trabalhos relacionados (nível de rede) não reflete, necessariamente, na justiça em relação ao *goodput* e o atraso (nível de usuário).

Realizamos, então, diversos experimentos onde mostramos que a justiça em nível de usuário pode ser significativamente melhorada se uma pequena fração dos usuários (localizados nos anéis mais distantes) adotarem antenas direcionais. Para este novo cenário considerado, apresentamos resultados onde não só os usuários dos últimos anéis, mas outros usuários, como os usuários móveis, se beneficiam desta solução, tendo suas medidas de *goodput* e atraso melhoradas.

Nossos experimentos nos permitiram, também, mostrar que as limitações intrínsecas à camada física de redes sem fio são importantes componentes da justiça em nível de usuário alcançável em um sistema. Quando utilizamos antenas direcionais, atacamos o problema sob um ponto de vista diferente em relação aos trabalhos relacionados. Esta fase de nosso trabalho resultou em um artigo no periódico *IEEE Transactions on Broadcasting: Special Issue in Multimedia Broadcasting* [16], publicado em 2008.

Finalmente, é importante ressaltar que os experimentos realizados até então fo-

ram feitos com foco em redes EVDO. Porém, como mostrado no Capítulo 2, diversos padrões de redes sem fio de longa distância, como, por exemplo, o Wimax, o HSDPA e o WCDMA, entre outros, também estão analogamente sujeitos ao mesmo problema no que diz respeito ao uso de modulação adaptativa e algoritmos de escalonamento *cross-layer*. Sob este prisma, os estudos, os problemas e as discussões apresentadas nesta parte da tese não são úteis especificamente para um padrão, mas podem contribuir a qualquer tecnologia de rede sem fio de longa distância.

Trabalhos Futuros

Um trabalho futuro seria investigar o uso de antenas inteligentes que adaptem dinamicamente seu ganho em diferentes direções, de acordo com a movimentação dos terminais de acesso, com o objetivo de melhorar o equilíbrio entre as vazões dos usuários, alcançando, assim, maior justiça.

6.2 Sobre a qualidade de serviço de fluxos de vídeos escaláveis em redes sem fio

O forte crescimento do uso de fluxos de vídeo, sobretudo sobre cenários de redes sem fio, tem gerado grande interesse na pesquisa de novos, e complexos, algoritmos que melhorem a qualidade de serviço.

Partindo desta motivação, apresentamos uma nova maneira de transmitir fluxos de vídeo de forma altamente benéfica: e sem a necessidade de se empregar complexos algoritmos para priorização. Obtivemos níveis significativamente reduzidos de perda para pacotes mais importantes para a qualidade do vídeo, quando comparamos nosso método com o método de transmissão tradicionalmente usado.

Para embasar nossa proposta, através de um modelo analítico proposto nesta tese, estudamos o processo aleatório de perda sob um novo ponto de vista. Com isso, fomos capazes de capturar um fenômeno que, até onde sabemos, era desconhecido pela literatura. Chamamos a nova característica capturada pelos nossos modelos de "processo de perda *intra-rajada*". Os resultados fundamentais obtidos pelo estudo realizado através deste modelo trouxeram contribuições publicadas no SBRC 2010 [28];

Criamos, também, dois modelos de simulação que são alimentados por tráfegos reais de vídeos escaláveis. Estes modelos buscam incrementar o estudo realizado através do modelo analítico, à medida que apresentam características realistas de aplicações de fluxo de vídeo escalável sob redes sem fio. Um dos modelos de simulação considera um serviço ditado por medições reais obtidas de uma grande rede EVDO. Para estudar o problema de forma mais geral, criamos outro modelo de simulação que considera serviço exponencial. Realizamos experimentos com ambos os modelos e obtivemos resultados análogos no que diz respeito ao processo de perda *intra-rajada*.

Além disso, não analisamos apenas métricas de nível de rede, como a perda de pacotes, mas também métricas do ponto de vista do usuário, como a perda de camadas de vídeo, bem como o PSNR do vídeo recebido. Para isso, vários perfis de vídeo, com diferentes características de tráfego foram usados em nossos experimentos. Entre os resultados obtidos, destacamos os que indicaram redução de 25% a 50% da fração de quadros de vídeo com qualidade ruim, além de significativo aumento na fração de quadros com qualidade excelente.

As contribuições obtidas através dos modelos de simulação estão sendo organizadas para serem submetidas a um periódico internacional.

Outros cenários de aplicação

É importante destacar a generalidade do Modelo Analítico e do Modelo de Simulação A que são capazes de capturar o fenômeno chave para este trabalho.

Embora a parametrização escolhida tenha o objetivo de representar cenários de redes sem fio, os modelos podem ser parametrizados para estudo de outros cenários como, por exemplo, redes ADSL de acesso à Internet.

Com isso, as ideias trazidas por esta parte de nosso trabalho são aplicáveis a cenários bem mais gerais, sejam de redes cabeadas ou sem fio, desde que o último enlace seja o gargalo do caminho entre a fonte e o destino.

Trabalhos Futuros

Nesta parte de nosso trabalho, mostramos que é possível atingir significativo nível de priorização de pacotes mais importantes para a qualidade de fluxos escaláveis de

vídeo mesmo sem o uso de algoritmos complexos, como os disponíveis na literatura. No entanto, algumas questões ainda sem resposta servem de motivação para possíveis trabalhos futuros relacionados a essa área.

Primeiramente, deixamos como trabalho futuro o estudo mais detalhado dos resultados de PSNR para os experimentos realizados com o modelo de simulação B (rede EVDO). Este estudo está em fase final de desenvolvimento.

Outro estudo, realizado através de experimentos práticos capazes de avaliar o processo de perda *intra-rajada* em aplicações reais de vídeos escaláveis, é um dos trabalhos de continuidade desta tese que já está em andamento. Temos um aplicativo, em fase final de desenvolvimento, capaz de realizar transmissão de fluxos escaláveis de vídeos, de acordo com a proposta trazida por esta tese.

A escolha do nível de rajada ainda é um processo empírico. Uma outra questão a ser considerada seria: Como escolher um nível de rajada que resulte em qualidade ótima de vídeo, dado um determinado cenário de rede e de parametrização de codificação de vídeo? A formulação deste problema como um problema de otimização capaz de maximizar a qualidade final do vídeo é um problema em aberto nessa área.

Além disso, embora sejamos capazes de apontar que um intervalo de nível de rajada ($1.3 \leq b \leq 1.8$) em geral resulta em bom nível de priorização de pacotes BL, ainda não temos um algoritmo que seja capaz de escolher, mesmo que empiricamente, um bom valor para b .

Referências Bibliográficas

- [1] Mun Choon Chan and Ramachandran Ramjee. TCP/IP Performance over 3G Wireless Links with Rate and Delay Variation. *Journal Wireless Networks*, 11:81–97, Jan 2005.
- [2] Z Fu, P. Zerfos, H. Luo, S. Lu, L. Zhang, and M. Gerla. The impact of multihop wireless channel on TCP throughput and loss. In *Proc. of IEEE INFOCOM 2003*, volume 3, pages 1744–1753, 2003.
- [3] G. Xylomenos, G.C. Polyzos, P. Mahonen, and M. Saaranen. TCP performance issues over wireless links. In *Communications Magazine, IEEE*, volume 39, pages 52–58, 2001.
- [4] David W. Paranchych and Mehmet Yavuz. Analytical expressions for 1xEV-DO forward link throughput. In *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1823–1828, April-May 2003.
- [5] Qi Bi and Stan Vitebsky. Performance Analysis of 3G-1X EVDO High Data Rate System. In *Proc. of IEEE Wireless Communications and Networking Conference*, pages 389–395, 2002.
- [6] Eduardo Esteves. On the reverse link capacity of cdma2000 high rate packet data system. In *Proc. Inter. Conf. Commun*, pages 1823–1828, April-May 2002.
- [7] Wee Lum Tan, Fung Lam, and Wing Cheong Lau. An Empirical Study on 3G Network Capacity and Performance. In *Proc. of IEEE INFOCOM*, May 2007.

- [8] B.P. Crow, I. Widjaja, LG Kim, and PT Sakai. IEEE 802.11 wireless local area networks. *Communications Magazine, IEEE*, 35(9):116–126, 2002.
- [9] Okumura-Hata. Fading channels: information-theoretic and communications aspects. In *IEEE Transactions on Information Theory*, volume 44, pages 2619–2692, 1998.
- [10] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda. Field Strength and its variability in VHF and UHF land-mobile service. In *Hata-Okumura*, volume 16(9-10), pages 825–873, 1968.
- [11] Okumura-Hata. Propagation prediction model for UHF range, in the Prediction methods for the terrestrial land mobile service in the VHF and UHF bands. In *ITU-R Recommendation P. 529-2, Geneva: ITU*, pages 5–7, 1995.
- [12] José Roberto Boisson de Marca, Flávio Pimentel Duarte, G. D. G. Jaime, Edmundo de Souza e Silva, Rosa Maria Meri Leão, and Patricia A. Berquó. Modelagem e Avaliação de Desempenho da Tecnologia CDMA2000 1xEV-DO para Aplicação no SBTvD. In *Simpósio Brasileiro de Redes de Computadores*. SBRC, 2006.
- [13] F.P. Duarte G.D.G. Jaime, R.M.M. Leao, E. de Souza e Silva, P.A.Berquo, and J. Roberto B. de Marca. On the Use of Cellular Technology for Digital TV Bi-Directional Return Channel Services, Proc. In *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*, number 9188240, pages 1–6. IEEE, sep 2006.
- [14] J. Le Boudec and M. Vojnovi. Perfect simulation and stationarity of a class of mobility models. In *IEEE - INFOCOM*, 2005.
- [15] G.D.G. Jaime, R.M.M. Leão, E. de Souza e Silva, and J. Roberto B. de Marca. The effect of mobility on the user-level fairness of a 3G wireless Technology (EV-DO). In *IEEE Wireless Communications & Networking Conference (WCNC)*. IEEE, 2008.

- [16] G.D.G. Jaime, R.M.M. Leão, E.S. e Silva, and J.R.B. de Marca. User Level Fairness of a Multirate 3G Technology (EVDO) With Application to Digital TV Return Channel Technology. *IEEE TRANSACTIONS ON BROADCASTING*, 54(3), 2008.
- [17] I.E.G. Richardson. *H. 264 and MPEG-4 video compression: video coding for next-generation multimedia*. John Wiley & Sons Inc, 2003.
- [18] H. Radha, M. van der Schaar, and Y. Chen. The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP. *IEEE Transactions on multimedia*, 3(1):53–68, 2001.
- [19] W. Li. Overview of fine granularity scalability in MPEG-4 video standard. *IEEE Transactions on circuits and systems for video technology*, 11(3):301–317, 2001.
- [20] Hayder M. Radha, Mihaela Van der Schaar, and Yingwei Chen. The mpeg-4 fine-grained scalable video coding method for multimedia streaming over ip. In *IEEE Transactions on Multimedia*, volume 3, pages 53–67, mar 2001.
- [21] Thomas Schierl, Thomas Stockhammer, and Thomas Wiegand. Mobile video transmission using scalable video coding. In *IEEE Transactions on Circuits and Systems For Video Technology*, volume 17, pages 1204–1217, Sep 2007.
- [22] Y. Shan. Cross-layer techniques for adaptive video streaming over wireless networks. *EURASIP journal on applied signal processing*, 2005(2):220–228, 2005.
- [23] Ivaylo Haratcherev, Jacco Taal, Koen Langendoen, Reginald Lagendijk, and Henk Sips. Optimized video streaming over 802.11 by cross-layer signaling. In *IEEE Communications Magazine*, pages 115–121, jan 2006.
- [24] T.P. Chen and T. Chen. Fine-grained rate shaping for video streaming over wireless networks. *EURASIP Journal on Applied Signal Processing*, 2004(2):176–191, 2004.

- [25] A. Fiandrotti, D. Gallucci, E. Masala, and E. Magli. Traffic Prioritization of H. 264/SVC Video over 802.11 e Ad Hoc Wireless Networks. In *Computer Communications and Networks, 2008. ICCCN'08. Proceedings of 17th International Conference on*, pages 1–5, aug 2008.
- [26] Honghai Zhang, Yanyan Zheng, and ohammad A. (Amir) Khojastepour. Scalable Video Streaming over Fading Wireless Channels. *IEEE Wireless Communications and Networking Conference (WCNC'09)*, apr 2009.
- [27] M. van der Schaar and H. Radha. Unequal packet loss resilience for fine-granular-scalability video. *IEEE Transactions on Multimedia*, 3(4):381–394, 2001.
- [28] G.D.G. Jaime, R.M.M. Leão, E.S. e Silva, and J.R.B. de Marca. On the reduction of scalable video base-layer packet loss rate on droptail/FIFO queues XXVIII SBRC - Brazilian Symposium on Computer Networks and Distributed Systems. *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*, 2010.
- [29] G.D.G. Jaime, R.M.M. Leão, E.S. e Silva, and J.R.B. de Marca. Effect of varying the intra-frame packet burstiness on the performance of wireless-video streaming (**aceito para publicação**). *IEEE Globecom*, 2010.
- [30] ITU-T. Reference Algorithm for Computing Peak Signal to Noise Ration (PSNR) of a Processed Video Sequence with Constant Spatial Shifts and a Constant Delay. *New Recommendation*, 2009.
- [31] Theodore S. Rappaport and Theodore Rappaport. *Wireless Communications: Principles and Practice (2nd Edition)*. Prentice Hall PTR, December 2001.
- [32] John Proakis and Masoud Salehi. *Digital Communications*. McGraw-Hill, 2007.
- [33] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communications*. Cambridge, 2005.

- [34] J. B. Andersen, T.S. Rappaport, and S. Yoshida. Propagation Measurements and Models for Wireless Communications Channels. In *IEEE Communications Magazine*, volume 33, pages 42–49, 1995.
- [35] Bernard H. Fleury and Peter E. Leuthold. Radiowave Propagation in Mobile Communications: An overview of European Research. In *IEEE Communications Magazine*, volume 34, pages 70–81, 1996.
- [36] T.K. Sarkar, Ji Zhong, Kim Kyungjung, A. Medouri, and M. Salazar-Palma. A survey of various propagation models for mobile communication. In *IEEE Antennas and Propagation Magazine*, volume 45, pages 51–82, 2003.
- [37] Bernard Sklar. Rayleigh Fading Channels in Mobile Digital Communications Systems Part I: Characterization. In *IEEE Communications Magazine*, volume 35, pages 90–100, 1997.
- [38] Jianhua Lu, K. B. Letaif, Justin C-I Chuang, and Ming L. Liou. M-PSK and M-QAM BER Computation Using Signal-Space Concepts. In *IEEE Transactions on Communications*, volume 47, pages 181–184, 1999.
- [39] Arunabha Ghosh, David r. Wolter, Jeffrey G. Andrews, and Runhua Chen. Broadband Wireless Access with Wimax/802.16: Current Performance benchmarks and Future Potential. In *IEEE Communications Magazine*, pages 129–136, Feb 2005.
- [40] Noriyuki Fukui. Study of channel quality feedback in umts hsdpa. In *IEEE Personal Indoor Mobile Radio Communications 2008*, pages 336–340, Oct 2003.
- [41] Erik Dhalman, Per Beming, Jens Knutsson, fredrik Oversjö, Magnus Persson, and Christiaan Roobol. Wcdma - the radio interface for future mobile multimedia communications. In *IEEE Transactions on Vehicular Technology*, volume 47, pages 1105–1118, Nov 1998.
- [42] Erik Dhalman, Per Beming, Jens Knutsson, fredrik Oversjö, Magnus Persson, and Christiaan Roobol. Hybrid arq and adaptive modulation and coding schemes for high speed downlink packet access. In *Personal, Indoor*

and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on, volume 3, pages 1073–1077, Sep 2002.

- [43] R. Knopp and PA Humblet. Information capacity and power control in single-cell multi-user communications. *Communications, 1995. ICC 95 Seattle, Gateway to Globalization, 1995 IEEE International Conference on*, 1, 1995.
- [44] DNC Tse. Multiuser diversity and proportional fair scheduling. *Bell Labs presentation*, 1999.
- [45] D. Gesbert and M.S. Alouini. Selective multi-user diversity. *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology*, pages 162–165, 2003.
- [46] S. Shakkottai, TS Rappaport, and PC Karlsson. Cross-layer design for wireless networks. *Communications Magazine, IEEE*, 41(10):74–80, 2003.
- [47] D. Gesbert and M.S. Alouini. How much feedback is multi-user diversity really worth? *Communications, 2004 IEEE International Conference on*, 1, 2004.
- [48] M. Realp, AI Perez-Neira, and C. Mecklenbrauker. A cross-layer approach to multi-user diversity in heterogeneous wireless systems. *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, 4, 2005.
- [49] 3rd Generation Partnership Project 2. Cdma2000 High Rate Packet Data Air Interface Specification, C.S2002.4, V3.0, December 2001.
- [50] E. Esteves, P.J. Black, and M.I. Gurelli. Link Adaptation Techniques for High-Speed Packet Data in Third Generation Cellular Systems. In *Proc. European Wireless Conference*, 2002.
- [51] E. de Souza and Silva e R.R. Muntz. *Métodos Computacionais de Solução de Cadeias de Markov: Aplicações a Sistemas de Computação e Comunicação*. VIII Escola de Computação, 1992.

- [52] KS Trivedi. *Probability and Statistics with Reliability, Queuing, and Computer Science Applications. 1982*, volume 7. Prentice-Hall, 1982.
- [53] E. de Souza e Silva and H.R. Gail. *Computational Probability*, chapter Transient Solutions for Markov Chains. Kluwer Academic Publishers, 2000.
- [54] N. Feamster, D. Bansal, and H. Balakrishnan. On the interactions between layered quality adaptation and congestion control for streaming video. In *11th International Packet Video Workshop*. Citeseer, 2001.
- [55] Q. Li and M. Van Der Schaar. Providing adaptive QoS to layered video over wireless local area networks through real-time retry limit adaptation. *IEEE Transactions on Multimedia*, 6(2):278–290, 2004.
- [56] R. Denda, A. Banchs, and W. Effelsberg. The fairness challenge in computer networks. In *Quality of Future Internet Services*, pages 208–220. Springer, 2000.
- [57] W. Ogryczak, A. Wierzbicki, and M. Milewski. A multi-criteria approach to fair and efficient bandwidth allocation. *Omega*, 36(3):451–463, 2008.
- [58] A. Jalali, R. Padovani, and R. Pankaj. Data throughput of Cdma-HDR a high efficiency-high data rate personal communication wireless system. In *Proc. of IEEE Vehicular Technology Conference*, volume 3(7), pages 1854–1858, 2000.
- [59] Jong Hun Rhee, Tae Hyung Kim, and Dong Ku Kim. A wireless fair scheduling algorithm for 1xev-do system, 2001.
- [60] Kuenyoung, Hoon Kim, and Youngnam Han. A proportional fair scheduling algorithm with qos and priority in 1xev-do. In *Proc. of IEEE PIMRC*, pages 2239–2243, 2002.
- [61] Robert C. Elliot and Witold Krzymietn. Scheduling algorithms for the cdma2000 packet data evolution. In *Proc. of IEEE VTC*, pages 304–310, 2002.

- [62] Matthew Andrews. Instability of the proportional fair scheduling algorithm for hdr. In *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*, pages 1422–1426, sep 2004.
- [63] Taewon Ban. A new efficient scheduling algorithm for the support of qos in cdma2000 1xev-do. In *IEEE TRANSACTIONS ON COMMUNICATIONS E SERIES B*, 2004.
- [64] Andrews M., Kumaran K., Ramanan K., Stolyar A., Whiting P., and Vijayakumar R. Providing quality of service over a shared wireless link. In *Proc. of IEEE Communications Magazine*, pages 777–787, feb 2001.
- [65] S.H. Kang and A. Zakhor. Packet scheduling algorithm for wireless video streaming. *International Packet Video Workshop*, 2002.
- [66] YuLong Fan and ChungYao Huang. Adaptative scheduler algorithm for multicritarion designs in b3g wireless systems. In *IEEE Wireless Communications and Networking Conference (WCNC)*, march 2004.
- [67] Peng-Yong Kong and Keng-Hoe Teh. Performance of proactive earliest due date packet scheduling i wireless networks. In *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, volume 53, pages 1224–1234, July 2004.
- [68] Seungwan Ryu, Byung-Han Ryu, Hyunhwa Seo, Muyong Shin, and Seikwon Park. Wireless packet scheduling algorithm for ofdma system based on time-utility and channel state. In *International Symposium on Wireless Systems and Networks (ISWSN05)*, volume 7, pages 777–787, december 2005.
- [69] P. Black and M. Gurelli. Capacity Simulation of Cdma2000 1x-EV Wireless Access Internet System. In *IEEE MWCN*, 2001.
- [70] Bubhu Mohanty, Ramin Rezaiifar, and Rajesh Pankaj. Application layer capacity of the Cdma2000 1xEV wireless internet access system. In *Qualcomm Technical Memo*, 2001.

- [71] C. Huang, Qi Bi, Asif Gandhi, Ron Brown, and Dongzhe Cui. Forward and Reverse Link Capacity of 1xEV-DO: third generation wireless high speed data systems, Records. In *Proc. of IEEE Globecom*, pages 871–875, 2002.
- [72] Eduardo Esteves, Mehmet I. Gurelli, and Mingxi Fan. Performance of fixed wireless access with cdma2000 1xeV-do. In *IEEE Vehicular Technology Conference Proceedings*, 2003.
- [73] R. Leelahakriengkrai et al. Performance Analysis of 1xEV-DO Systems under Realistic Traffic Models and Limited Size IP Backhaul. In *10th Asia Pacific Conf. on Comm. and 5th Intl. Symposium on Multi-Dimensional Mobile Comm.*, pages 179–183, 2004.
- [74] W. Xiao, F. Wang, E. Love, A. Ghosh, and R. Ratasuk. 1xEV-DO System Performance: Analysis and Simulation. In *Proc. of IEEE Vehicular Technology Conference*, pages 5305–5409, 2004.
- [75] A. Scaglione and M. van der Schaar. Cross-layer resource allocation for delay constrained wireless video transmission. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05)*, volume 5, 2005.
- [76] ISO/IEC MPEG. Svc requirements specified by mpeg. jvt-n026. 2005.
- [77] H.F. Hsiao and J.N. Hwang. Layered FGS video over active network with selective drop and adaptive rate control. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, volume 5, 2003.
- [78] C.W. Huang and J.N. Hwang. An embedded packet train and adaptive FEC scheme for effective video adaptation over wireless broadband networks. *Journal of Zhejiang University-Science A*, 7(5):811–818, 2006.
- [79] C.H. Foh, Y. Zhang, Z. Ni, and J. Cai. Scalable video transmission over the IEEE 802.11 e networks using cross-layer rate control. In *IEEE International Conference on Communications, 2007. ICC'07*, pages 1760–1765, jun 2007.

- [80] S. Mohamed and G. Rubino. A study of real-time packet video quality using random neural networks. *IEEE Transactions On Circuits and Systems for Video Technology*, 12(12):1071–1083, 2002.
- [81] J. Klaue, B. Rathke, and A. Wolisz. Evalvid-a framework for video transmission and quality evaluation. *Lecture notes in computer science*, pages 255–272, 2003.
- [82] MH Pinson and S. Wolf. A new standardized method for objectively measuring video quality. *IEEE Transactions on broadcasting*, 50(3):312–322, 2004.
- [83] Z. Wang, L. Lu, and A.C. Bovik. Video quality assessment based on structural distortion measurement. *Signal processing: Image communication*, 19(2):121–132, 2004.
- [84] W. Li, F. Ling, and X. Chen. Fine granularity scalability in MPEG-4 for streaming video. In *Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on*, volume 1, 2000.
- [85] S. Mohamed and G. Rubino. A study of real-time packet video quality using random neural networks. *IEEE Transactions On Circuits and Systems for Video Technology*, 12(12):1071–1083, 2002.
- [86] Q. Qu, Y. Pei, and J.W. Modestino. Robust H. 264 video coding and transmission over bursty packet-loss wireless networks. *IEEE 58th Vehicular Technology Conference (VTC)*, 5:3395–3399, 2003.
- [87] M. Wien, R. Cazoulat, A. Graffunder, A. Hutter, and P. Amon. Real-time system for adaptive video streaming based on SVC. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(9):1227–1237, 2007.
- [88] M. Van Der Schaar, H. Radha, and C. Dufour. Scalable MPEG-4 video coding with graceful packet-loss resilience over bandwidth-varying networks. In *Proc. ICME 2000*.
- [89] G. Liebl, T. Schierl, T. Wiegand, and T. Stockhammer. Advanced wireless Multiuser Video Streaming using the Scalable Video Coding Extensions

- of H. 264/MPEG4-AVC. In *2006 IEEE International Conference on Multimedia and Expo*, pages 625–628, 2006.
- [90] Y.S. Baguda, N. Faisal, S.H. Syed, S.K. Yusof, M.S. Abdullah, A. Mohd, and A. Zulkarmawan. Mobile streaming of H. 264 video over Gilbert-Elliotts Channel. *constraints*, 3:4.
- [91] T. Pliakas, G. Kormentzas, and S. Tsekeridou. Joint scalable video coding and packet prioritization for video streaming over IP/802.11 e heterogeneous networks. In *Proceedings of the 3rd international conference on Mobile multimedia communications*, page 31. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007.
- [92] S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz, and L. Stibor. IEEE 802.11e Wireless LAN for Quality of Service. In *Proc. European Wireless*, volume 18, pages 32–39. Citeseer, 2002.
- [93] Y. Bernet, P. Ford, R. Yavatkar, F. Baker, L. Zhang, M. Speer, R. Braden, B. Davie, J. Wroclawski, and E. Felstaine. A framework for integrated services operation over diffserv networks. *RFC2998*, November, 2000.
- [94] D. Vassiss, G. Kormentzas, A. Rouskas, and I. Maglogiannis. The IEEE 802.11g standard for high data rate WLANs. *IEEE network*, 19(3):21–26, 2005.
- [95] J.F. Kurose and K.W. Ross. *Computer networking: a top-down approach featuring the Internet*. Boston: Addison-Wesley, 2003.
- [96] A.C. Feng, A.C. Kapadia, W.C. Feng, and G.G. Belford. Packet spacing: An enabling mechanism for delivering multimedia content in computational grids. *The Journal of Supercomputing*, 23(1):51–66, 2002.
- [97] V. Sivaraman, H. Elgindy, D. Moreland, and D. Ostry. Packet pacing in short buffer optical packet switched networks. In *Proceedings of IEEE Infocom*. Citeseer, 2006.
- [98] Y. Cai, S. Hanay, and T. Wolf. Practical Packet Pacing in Small-Buffer Networks. *IEEE International Conference on Communications - ICC'09*, 2009.

- [99] YJ Liang, JG Apostolopoulos, and B. Girod. Analysis of packet loss for compressed video: Does burst-length matter? In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, volume 5, 2003.
- [100] Qiang Wu and Eduardo Esteves. The cdma2000 High rate Packer Data System, March 2002. Qualcomm.
- [101] Lucent Technologies Inc. Engineering Guidelines. 401-614-040.
- [102] K.S. Gilhousen, I.M. Jacobs, R. Padovani, A.J. Viterbi, L.A. Weaver, and Wheatley. On the capacity of a cellular CDMA system. *IEEE Transactions on Vehicular Technology*, 40:303–312, 1991.
- [103] R. M. M. Leão, E. de Souza e Silva, T. R. T. Fudoli, and G. D. G. Jaime. Um modelo para avaliação do desempenho do mecanismo de acesso múltiplo do protocolo docsis. In *Simpósio Brasileiro de Telecomunicações*, 2001.
- [104] Hyoung-Kee Choi and J. O. Limb. A Behavioral Model of Web Traffic. In *Proceedings of the Seventh Annual International Conference on Network Protocols*, 1999.
- [105] Mark E. Crovella and Azer Bestavros. Self-Similarity in World Wide Web Traffic. *IEEE/ACM Transactions on Networking*, 05:835–846, 1997.
- [106] S. Deng. Empirical Model of WWW Document Arrivals at Access Link. In *Proceedings of IEEE ICC*, 1996.
- [107] É. Tamás and S. Molnár. On the distribution of round-trip delays in tcp/ip networks. In *Proceedings of the 24th Annual IEEE Conference on Local Computer Networks*, 1999.
- [108] Jean-Yves Le Boudec. On the stationary distribution of speed and location of random waypoint. In *IEEE Transactions on Mobile Computing*, Jul-Ago 2005.
- [109] W. Navidi and T. Camp. Stationary distributions for the random waypoint mobility model. *Mobile Computing, IEEE Transactions on*, 3(1):99–108, 2004.

- [110] C. Bettstetter, G. Resta, and P. Santi. The node distribution of the random waypoint mobility model for wireless ad hoc networks. *Mobile Computing, IEEE Transactions on*, 2(3):257–269, 2003.
- [111] J. Yoon, M. Liu, and B. Noble. Random waypoint considered harmful. *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, 2.
- [112] E. de Souza e Silva, D.R. Figueiredo, and R.M.M. Leão. The TANGRAMII integrated modeling environment for computer systems and networks. *ACM SIGMETRICS Performance Evaluation Review*, 36(4):64–69, 2009.
- [113] E. de Souza e Silva, A.P.C. da Silva, A.A. de A. Rocha, R.M.M. Le ao, F.P. Duarte, F.J.S. Filho, G.D.G. Jaime, and R.R. Muntz. Modeling, analysis, measurement and experimentation with the tangram-ii integrated environment. In *International Conference on Performance Evaluation Methodologies and Tools (ValueTools '06)*, page 7, New York, NY, USA, 2006. ACM Press.
- [114] Qiang Wu and Eduardo Esteves. The Cdma2000 High rate Packet Data System. In *Advances in 3G Enhanced Technologies for Wireless Communications*, chapter 4. Jiangzhou Wang and Tung-Sang Ng., march 2002.
- [115] W.J. Goralski. *ADSL and DSL technologies*. McGraw-Hill Professional, 1998.
- [116] S. Jaiswal, G. Iannaccone, C. Diot, J. Kurose, and D. Towsley. Measurement and classification of out-of-sequence packets in a tier-1 IP backbone. *IEEE/ACM Transactions on Networking (TON)*, 15(1):54–66, 2007.
- [117] J. Cao, W.S. Cleveland, D. Lin, and D.X. Sun. Internet traffic tends toward Poisson and independent as the load increases. *LECTURE NOTES IN STATISTICS-NEW YORK-SPRINGER VERLAG-*, pages 83–110, 2003.
- [118] T. Karagiannis, M. Molle, M. Faloutsos, and A. Broido. A nonstationary Poisson view of Internet traffic. In *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 3, 2004.

- [119] G. Latouche and V. Ramaswami. *Introduction to matrix analytic methods in stochastic modeling*. Society for Industrial Mathematics, 1999.
- [120] Rosa M.M. Leão Edmundo de Souza e Silva, Daniel R. Figueiredo. The TANGRAMII Integrated Modeling Environment for Computer Systems and Networks. *Performance Evaluation Review*, pages 64–69, 2009.
- [121] E. de Souza e Silva, et al. Modeling, Analysis, Measurement and Experimentation with the Tangram-II Integrated Environment. In *International Conference on Performance Evaluation Methodologies and Tools, 2006*, volume 180,, pages 1–10, Pisa,, 2006. Proceedings of the First International Conference on Performance Evaluation Methodologies and Tools, 2006.
- [122] S. Floyd. Van Jacobson, 201c. *On traffic phase effects in packet-switched gateways*, 201d *Internetworking: Research and Experience*, 3(3):115–156, 1992.
- [123] S. Floyd. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking (TON)*, 1(4):397–413, 1993.
- [124] D.F. Towsley, V. Firoiu, J. Padhye, and J.F. Kurose. Modeling TCP Reno performance: a simple model and its empirical validation. *IEEE/ACM Transactions on Networking (ToN)*, 8(2):133–145, 2000.
- [125] G. Van der Auwera, P.T. David, M. Reisslein, and L.J. Karam. Traffic and quality characterization of the H. 264/AVC scalable video coding extension. *Advances in Multimedia*, 2008(2):1, 2008.
- [126] Z. Avramova, D. De Vleeschauwer, K. Laevens, S. Wittevrongel, and H. Brueneel. Modelling H.264/AVC VBR video traffic: comparison of a Markov and a self-similar source model. In *Journal of Telecommunication Systems (Springer Netherlands)*, volume 39, pages 91–102. Springer Netherlands, october 2008.

- [127] M. Dai. Analysis and modeling of MPEG-4 and H. 264 multi-layer video traffic. In *Proceedings of IEEE INFOCOM*, pages 2257–2267. Citeseer, 2005.
- [128] W. Hui, SHA Ji-chang, SUN Xiao, and TAO Jun. MPEG-4 FGS Video Traffic Model and Its Application in Simulation. *Computer*, page 12, 2006.
- [129] HE Wei and W. Hui. A Video Traffic Model for MPEG-4 FGS. *Computer*, page 10, 2004.
- [130] J.A. Zhao, B. Li, and I. Ahmad. Traffic model for layered video: an approach on markovian arrival process. In *Packet Video*, pages 28–30. Citeseer, 2003.
- [131] W. Li and Z. Xiao. System Modeling and Implementation of MPEG-4 Encoder under Fine-Granular-Scalability Framework. 2002.
- [132] M. Wien, H. Schwarz, and T. Oelbaum. Performance analysis of SVC. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(9):1194–1203, 2007.
- [133] J. Reichel, H. Schwarz, and M. Wien. Joint scalable video model 11 (JSVM 11). *Joint Video Team, doc. JVT-X202, Geneva, Switzerland*, 2007.
- [134] Joint Video Team (JVT). Jsvm (joint scalable video model) software manual. <http://ftp.lahtermaher.org/pub/video/jsvm/current/jsvm/SoftwareManual.doc>.
- [135] D. De Vera, P. Rodríguez-Bocca, and G. Rubino. Automatic quality of experience measuring on video delivering networks. 2008.
- [136] K.D. Singh, J. Orozco, D. Ros, and G. Rubino. Improving Perceived Streaming-Video Quality in High Speed Downlink Packet Access. In *IEEE Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008*, pages 1–6, 2008.
- [137] S. Mohamed, G. Rubino, F. Cervantes, and H. Afifi. Real-time video quality assessment in packet networks: A neural network model. *RAPPORT DE*

- [138] A. Lo, G. Heijenk, and I. Niemegeers. Performance evaluation of MPEG-4 video streaming over UMTS networks using an integrated tool environment. In *Proc. of*, pages 24–28. Citeseer, 2005.
- [139] R. Attar, D. Ghosh, C. Lott, M. Fan, P. Black, R. Rezaiifar, and P. Agashe. Evolution of cdma2000 cellular networks: multicarrier EV-DO. *IEEE Communications Magazine*, 44(3):46–53, 2006.
- [140] M.I. Gurelli Esteves and M. Fan. Performance of Fixed Wireless Access with Cdma2000 1xEV-DO, Records. In *Proc. of IEEE Vehicular Technology Conference*, 2003.