

Voz e vídeo sobre redes sem Fio IEEE 802.11

ARLINDO FLÁVIO DA CONCEIÇÃO

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO DE DOUTOR
EM
CIÊNCIA DA COMPUTAÇÃO.

ORIENTADOR: PROF. DR. FABIO KON

São Paulo, maio de 2006

Voz e vídeo sobre redes sem Fio IEEE 802.11

Este exemplar corresponde à redação final da tese de doutorado de Arlindo Flávio da Conceição, que foi defendida e aprovada pela comissão julgadora.

São Paulo, 15 de maio de 2006.

BANCA EXAMINADORA:

PROF. DR. FABIO KON

IME-USP

PROF. DR. FÁBIO MOREIRA COSTA

INF-UFG

PROF. DR. OTTO CARLOS MUNIZ BANDEIRA DUARTE

COPPE-UFRJ

PROF. DR. SERGIO TAKEO KOFUJI

PSI-EPUSP

PROF. DR. ALFREDO GOLDMAN

IME-USP

Dedicatória

Ao meu pai, *in memoriam*, de quem me lembro quando preciso de bom senso e equilíbrio.

À minha mãe, em quem me baseio quando preciso de força.

À Daniela, que me faz lembrar que no mundo não existe nada superior ao Amor.

E a todos aqueles que moram no meu coração.

Agradecimentos

Gostaria de agradecer, primeiramente, ao professor Fabio Kon, pela sua valiosa orientação e pelo seu sempre pronto apoio nos momentos mais difíceis.

Agradeço à minha família, que entendeu minha ausência e a supriu, permitindo que eu pudesse me afastar para a elaboração desta tese.

Agradeço também ao CNPq, à Ericsson Telecomunicações do Brasil e à Microsoft Research, cujo apoio financeiro permitiu a realização deste trabalho.

Agradeço aos amigos das listas `aabalados`, `vamosla` e `cortico`, que aliviaram o fardo do trabalho com piadas — algumas politicamente incorretas — e permanente companheirismo. Ainda dentre os amigos, agradeço aos colegas da Sociedade Filosófica Nova Acrópole, que me mostraram através de exemplos o verdadeiro significado da palavra voluntariado.

Agradeço também aos professores e pesquisadores que, por meio de valiosas sugestões, contribuíram para este trabalho. Em particular, agradeço ao professor Fábio Costa, da Universidade Federal de Goiás, pelas sugestões feitas na fase inicial e aos pesquisadores Jin Li e Dinei A. Florêncio, da Microsoft Research, pelas discussões sobre a tecnologia de Voz sobre IP. Agradeço também aos colegas do Grupo de Sistemas Distribuídos, do IME-USP, pelo convívio e discussões semanais.

Por fim, e de modo especial, agradeço à Daniela pelo apoio incondicional.

Resumo

Devido à crescente utilização das redes sem fio IEEE 802.11 (Wi-Fi) e das aplicações multimídia, esta tese analisa a transmissão de voz e vídeo sobre redes Wi-Fi. Apresentamos, baseados em experimentos reais, os principais problemas e propomos estratégias adaptativas para a sua solução. Nas transmissões de voz, os principais problemas observados foram a maior incidência de tráfegos em rajada e os problemas associados à execução de *handoffs*. Por sua vez, nas aplicações de fluxos de vídeo, os principais problemas observados foram a perda de pacotes associada aos *handoffs* e o baixo isolamento entre fluxos. Devido a este último, a transmissão de vídeo para uma unidade móvel pode afetar negativamente a transmissão para todas as unidades móveis atendidas por um ponto de acesso. No entanto, apesar desses problemas, as aplicações oferecem boa qualidade, desde que algumas restrições sejam obedecidas, tais como: não exceder a capacidade efetiva de vazão da rede, utilizar a rede apenas sob boa qualidade de sinal e manter mobilidade moderada. Dentre as principais contribuições desta tese podemos destacar:

- levantamento dos principais problemas para a utilização de voz e vídeo sobre redes Wi-Fi;
- avaliação dos cenários viáveis para o uso de VoIP sobre redes IEEE 802.11 atuais;
- proposta de implementação de pontos de acesso capazes de lidar mais eficientemente com a transmissão de fluxos de vídeo H.264/AVC;
- proposta de algoritmo adaptativo capaz de utilizar a qualidade de sinal das redes IEEE 802.11 como métrica de adaptação;
- uso do mecanismo de pares de pacotes para, no nível da aplicação, monitorar a taxa efetiva de transmissão exercida entre o ponto de acesso e as unidades móveis.

Abstract

Considering the growing usage of IEEE 802.11 (Wi-Fi) wireless networks and multimedia applications, this work analyzes empirically the behavior of voice and video traffic over Wi-Fi networks. We present, based on real experiments, the major problems and propose some adaptive strategies to solve them. The major problems for voice traffic were: packet loss and increased latency due to bursts and handoffs. In the video applications, the main problems were packet loss related to handoffs and the weak isolation among multiple streams. In IEEE 802.11 networks, a video stream can affect negatively the transmission to other mobile units served by an access point. Despite these problems, the utilization of voice and video applications over IEEE 802.11 networks offered good quality when used moderately, in other words, when used with low mobility, under good signal strength conditions and not exceeding the network capacity. The main contributions of this thesis are:

- a survey of the major problems related to voice and video traffic over IEEE 802.11 wireless networks;
- evaluation of feasible scenarios for VoIP over IEEE 802.11 using the technologies currently available;
- proposal of adaptive IEEE 802.11 access points to deal with H.264/AVC streaming efficiently;
- proposal of a novel algorithm that uses IEEE 802.11 signal quality as an adaptation metric;
- usage of the packet-pair mechanism to monitor, at the application layer, the real transmission rate between access points and mobile units.

Sumário

1	Introdução	1
1.1	Objetivos	3
1.2	Metodologia	3
1.3	Principais referências e trabalhos relacionados	4
1.4	Escopo do trabalho	5
1.5	Ambiente experimental	5
1.6	Contribuições e perspectivas	6
1.7	Organização do texto	6
2	Introdução às redes IEEE 802.11	9
2.1	Cronologia de desenvolvimento do padrão IEEE 802.11	9
2.2	Modos de operação: infra-estruturado e <i>ad hoc</i>	11
2.3	A camada MAC de IEEE 802.11	12
2.4	Mecanismos e configuração de IEEE 802.11	13
2.4.1	Retransmissão automática	14
2.4.2	Fragmentação	14
2.4.3	Requisição e reserva do canal de transmissão: mecanismo RTS/CTS	15
2.4.4	Modo de economia de energia	15
2.4.5	Migração (<i>handoff</i>)	16
2.4.6	Controle automático de taxa de transmissão	16
2.4.7	Outros parâmetros configuráveis de IEEE 802.11	16
2.5	O futuro de IEEE 802.11	17

3	Voz sobre redes IEEE 802.11	18
3.1	Requisitos das aplicações de VoIP	20
3.1.1	Latência	20
3.1.2	<i>Jitter</i>	20
3.1.3	Taxa de perda de pacotes	21
3.1.4	Vazão	21
3.2	Metodologia para coleta dos traços de execução de VoWiFi	22
3.2.1	Método para estimar o atraso	22
3.3	Obstáculos para a implantação de VoWiFi	25
3.3.1	<i>Handoff</i>	25
3.3.2	Tráfego em rajadas	29
3.4	Problemas adicionais	32
3.4.1	Capacidade	32
3.4.2	Justiça	34
3.4.3	Interferências	35
3.4.4	Diversidade de implementação	35
3.4.5	Métricas de qualidade de conexão	36
3.5	Mais algumas considerações sobre VoWiFi	37
3.6	Trabalhos relacionados	38
3.7	Cenários viáveis para a utilização de VoWiFi	39
3.8	Conclusões	40
4	Vídeo sobre redes IEEE 802.11	42
4.1	Arquitetura das aplicações de fluxos de vídeo	43
4.2	Requisitos dos fluxos de vídeo	45
4.2.1	Vazão	46
4.2.2	Taxa de perda de pacotes	46
4.2.3	<i>Jitter</i>	47
4.3	Problemas de fluxos de vídeo sobre Wi-Fi	47
4.3.1	Baixo isolamento	48

4.3.2	Comportamento liga/desliga	48
4.4	Introdução ao padrão de codificação de vídeo H.264/AVC	50
4.4.1	Camada de abstração de rede	51
4.4.2	Empacotamento RTP dos NALUs	53
4.4.3	Conjunto de parâmetros	53
4.5	<i>Media Aware Network Elements</i> (MANEs)	54
4.5.1	Identificação do tráfego H.264/AVC	55
4.5.2	Exploração da qualidade de serviço oferecida por IEEE 802.11e	56
4.5.3	Políticas de descarte de pacotes	57
4.6	Revisão das políticas de descarte de pacotes de IEEE 802.11	58
4.7	Descartes preguiçosos	60
4.7.1	Algoritmo	61
4.8	Trabalhos relacionados e conclusões	61
5	Desenvolvimento de aplicações adaptativas para redes IEEE 802.11	63
5.1	Trabalhos relacionados	64
5.2	Novos desafios impostos pelas redes IEEE 802.11	65
5.2.1	Problemas intrínsecos às redes sem fio	65
5.2.2	Problemas intrínsecos às redes IEEE 802.11	66
5.2.3	Problemas relacionados à diversidade das implementações	67
5.3	Desenvolvimento para redes IEEE 802.11: práticas recomendadas e técnicas úteis .	69
5.3.1	Tamanhos dos pacotes	70
5.3.2	Novas métricas de adaptação	70
5.3.3	Recuperação de erros	72
5.3.4	Operação desconectada	72
5.4	Interação entre aplicações e interfaces IEEE 802.11	73
5.4.1	Em sistemas Windows: <i>Network Driver Interface Specification</i> (NDIS) . . .	74
5.4.2	Em sistemas Linux: <code>/proc/net/wireless</code>	75
5.5	Adaptação de fluxos contínuos baseada na qualidade do sinal	76
5.5.1	Estratégia adaptativa baseada nos limiares de adaptação α , β e γ	77

5.5.2	Implementação	81
5.5.3	Ambiente experimental	82
5.5.4	Comportamento do algoritmo adaptativo	82
5.6	Conclusões	84
6	Pares de pacotes: uma nova métrica de adaptação	86
6.1	Controle automático da taxa de transmissão	88
6.1.1	Problemas acarretados pelo uso de ARC	89
6.2	O mecanismo de pares de pacotes e trabalhos relacionados	91
6.3	Ambiente experimental	94
6.4	Comportamento dos IEPs mediante mobilidade e suas aplicações práticas	94
6.4.1	Aplicação 1: monitoração das alterações na taxa de transmissão	96
6.4.2	Aplicação 2: inferência da capacidade máxima da rede	98
6.5	Limitações do mecanismo de pares de pacotes	99
6.6	Conclusões	101
7	Conclusões	103
7.1	Contribuições	104
7.2	Cenários e trabalhos futuros	105
7.3	Perspectivas	106

Lista de Figuras

1.1	Ambiente de testes utilizado.	6
1.2	Representação dos ambientes utilizados para execução dos experimentos.	7
3.1	Cabeçalho estendido do protocolo RTP.	23
3.2	Marcas de tempo usadas para obtenção do atraso.	24
3.3	Implicações da operação de <i>handoffs</i> nos tráfegos de VoWiFi.	26
3.4	Impacto da atividade de busca em VoWiFi.	28
3.5	Exemplo do comportamento dos tráfegos em rajada.	30
3.6	Correlação da ocorrência de rajadas em tráfegos distintos.	31
3.7	Comportamento de tráfegos de VoIP (Philadelphia $\leftrightarrow$ Seattle).	33
4.1	Etapas da transmissão de multimídia.	43
4.2	Interferência entre tráfegos mediante mobilidade.	49
4.3	Formato do cabeçalho de um NALU.	52
5.1	O pseudo-arquivo <code>/proc/net/wireless</code>	75
5.2	Qualidade do sinal e taxa de transmissão na presença de mobilidade.	77
5.3	Algoritmo para atualização de α e β	79
5.4	Convergência de α e β (TPPs 1% e 25% e fator de amortização 16).	80
5.5	Comparação entre o algoritmo adaptativo e o não adaptativo.	83
6.1	Vazão TCP em uma rede IEEE 802.11b (com o mecanismo RTS/CTS habilitado).	89
6.2	Mecanismo de trem de pacotes.	91
6.3	Mecanismo de pares de pacotes.	93

6.4	Comportamento dos IEPs mediante mobilidade (pacotes de 1000 octetos).	95
6.5	Distribuição dos IEPs.	96
6.6	Resultados da estratégia para determinação das alterações na taxa de transmissão.	97
6.7	Comportamento dos IEPs mediante mobilidade (pacotes de 67 octetos).	101

Lista de Tabelas

3.1	Comportamento dos <i>handoffs</i>	27
3.2	Tráfego de VoIP.	32
5.1	Comparação entre o algoritmo adaptativo e o não adaptativo.	84

Glossário

3GPP	3rd Generation Partnership Project
AAC	Advanced Audio Coding
ACK	ACKnowledgement message
API	Application Programming Interface
ARC	Automatic Rate Control
ARQ	Automatic Repeat reQuest
AVC	Advanced Video Coding
CA	Collision Avoidance
CD	Collision Detection
CRC	Cyclic Redundancy Check
CSMA	Carrier Sense Multiple Access
DCF	Distributed Coordination Function
DIFS	Distributed Inter Frame Space
FCC	Federal Communications Commission
FEC	Forward Error Correction
FER	Frame Error Rate

IAPP	Inter Access-Point Protocol
IEEE	Institute of Electrical and Electronics Engineers
IEP	Intervalo Entre Pacotes
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISMA	Internet Streaming Media Alliance
LBS	Location Based Services
LSE	Least Square Error
MAC	Media Access Control
MANE	Media Aware Network Element
MIMO	Multiple-Input Multiple-Output
MTU	Maximum Transmission Unit
NAL	Network Abstraction Layer
NALU	NAL Unit
NDIS	Network Driver Interface Specification
NIC	Network Interface Card
P2T	Push-To-Talk
PCF	Point Coordination Function
PHY	Physical layer
PSM	Power Save Mode
QoS	Quality of Service

RTCP Real-Time Control Protocol

RTP Real-Time Protocol

SNR Signal-to-Noise Ratio

TCP Transport Control Protocol

TSF Time Synchronization Function

TXOP Transmission Opportunity

UDP User Datagram Protocol

VCL Video Coding Layer

VoD Video on Demand

VoIP Voice over IP

VoWiFi Voice over Wi-Fi

VoWLAN Voice over WLAN

WECA Wireless Ethernet Compatibility Alliance

WEP Wired Equivalent Privacy

WFA Wi-Fi Alliance

Wi-Fi Wireless Fidelity

WiMAX Worldwide Interoperability for Microwave Access

WLAN Wireless Local Area Network

WPA Wi-Fi Protected Access

Capítulo 1

Introdução

Multimedia applications, specially those that incorporate continuous time-sensitive data like video, will be a challenge for wireless systems.

Randy H. Katz

Em 1994, o pesquisador Randy H. Katz afirmou que a transmissão de multimídia seria um desafio para os sistemas sem fio [Kat94]. A preocupação de Katz, que continua presente nos dias atuais, advinha principalmente de um fator: as aplicações multimídia e os ambientes móveis nasceram pressionados por forças diametralmente opostas. Os ambientes móveis, devido aos requisitos de mobilidade e autonomia, caracterizam-se pela economia de recursos. Por outro lado, as aplicações multimídia, devido aos requisitos de qualidade, caracterizam-se pela alta demanda de recursos. Uns, sovinas; as outras, esbanjadoras.

Soma-se a este paradoxo o fato de as aplicações multimídia possuírem requisitos de tempo real — em especial as aplicações interativas de voz e videoconferência — enquanto as redes sem fio são sabidamente sujeitas a maior latência do que as redes cabeadas tradicionais. Além disso, como lembrou Menin [Men02], sistemas de transmissão de multimídia são, antes de tudo, sistemas distribuídos; ou seja, também estão sujeitos aos problemas tradicionalmente encontrados em sistemas distribuídos, tais como os congestionamentos de rede e a necessidade de coordenação entre os componentes do sistema.

Entretanto, em face dos últimos avanços tecnológicos em redes sem fio e codificação de voz e vídeo, a afirmação feita por Katz em 1994 hoje poderia ser outra. Foram lançadas novas tecnologias de rede sem fio — tais como IEEE 802.11 [IEE99], IEEE 802.16 (WiMax) [IEE04] e Bluetooth [Blu03] — e de redes celulares de dados [Int03, Qua03, NTT03]. As redes sem fio evoluíram não apenas em termos de capacidade de transmissão, mas também em confiabilidade e área de cobertura; atualmente, nas grandes cidades brasileiras, praticamente toda a área urbana é coberta por pelo menos uma alternativa de conectividade sem fio.

Além disso, em paralelo ao desenvolvimento das redes sem fio, as tecnologias de codificação e transmissão de voz e vídeo também evoluíram. A transmissão de vídeo, hoje, graças aos avanços em tecnologias de codificação, pode ser feita usando-se bem menos banda passante do que seria necessário há dez anos; as aplicações de voz, por sua vez, já estão presentes no cotidiano de muitos usuários de Internet.

Contudo, apesar do rápido avanço em relação a tecnologias isoladas, como em H.264/AVC e IEEE 802.11, a integração entre as tecnologias de multimídia e as de redes sem fio não avançou na mesma velocidade. Em outras palavras, mesmo dez anos depois de Katz tê-lo afirmado, a transmissão de voz e vídeo sobre redes sem fio continua sendo um dos principais desafios tecnológicos da atualidade; porém, o desafio principal passou a ser a integração dessas tecnologias.

Integrar as tecnologias de multimídia às redes sem fio implica construir arquiteturas adaptativas [KCN01, Bej02] capazes de otimizar o uso de recursos como banda passante, capacidade de processamento e de armazenamento, consumo de energia, latência de comunicação, qualidade de percepção visual e auditiva, entre outros [Men02]. Pode-se afirmar que ainda não existem estratégias adaptativas que funcionem para todos os cenários de utilização das aplicações multimídia sobre redes sem fio (soluções *one-size-fits-all*). Desse modo, decidimos analisar os cenários mais relevantes para a realidade atual dos usuários; e dentre esses cenários certamente destacam-se as redes IEEE 802.11 e as aplicações de Voz sobre IP (VoIP) e de fluxos de Vídeo sob Demanda (*Video on Demand* ou VoD).

As redes sem fio IEEE 802.11, também conhecidas como redes Wi-Fi, tornaram-se o padrão em conectividade sem fio para redes locais e já equipam a maioria dos computadores portáteis novos;

é possível afirmar que os acessos sem fio à Internet, em sua maioria¹, já são feitos através dessas redes. Além disso, as aplicações de VoIP e de VoD estão entre as aplicações que mais cresceram em utilização nos últimos anos. VoIP destacou-se por oferecer uma alternativa de baixo custo, ou mesmo gratuita, às ligações telefônicas. Por sua vez, o crescimento de VoD é apoiado por vários fatores, como por exemplo a popularização da produção de vídeo, a evolução dos formatos de codificação, a redução dos custos de armazenamento e o aumento da capacidade de transmissão desses vídeos. Ou seja, a motivação para escolher estudar essas três tecnologias, Wi-Fi, VoIP e VoD, está sobretudo relacionada à sua ampla e crescente utilização. Além disso, a escolha das aplicações de VoIP e VoD não se deve apenas ao apelo de mercado dessas aplicações, mas deve-se também à cobertura que ambas, juntas, oferecem dos problemas de distribuição de multimídia sobre redes IEEE 802.11. Sob a ótica dos interesses de investigação, as aplicações de VoIP e VoD oferecem desafios distintos e, desse modo, complementares: VoIP tem como principal restrição a latência e VoD, a banda passante. Sendo assim, ao estudarmos essas duas aplicações, cobrimos grande parte dos problemas relacionados à distribuição de multimídia sobre redes IEEE 802.11.

1.1 Objetivos

Desse modo, diante do contexto colocado acima, o objetivo deste trabalho é investigar o comportamento da transmissão de voz e vídeo sobre redes IEEE 802.11. Mais especificamente, nosso objetivo é apontar os principais problemas presentes na transmissão e propor as respectivas estratégias de solução. Além disso, nos concentramos na camada de aplicação, pois é nesta camada que encontram-se as principais possibilidades para construção de estratégias adaptativas.

1.2 Metodologia

Para identificar os problemas da transmissão de voz e vídeo sobre redes IEEE 802.11, fizemos uso de experimentos práticos e da coleta e análise de traços de execução. Para obtenção desses traços, implementamos geradores de tráfego nas linguagens C e Java — eles enviam tráfegos RTP para as

¹Veja os artigos “*Wi-Pie in the sky?*”, de 9 de março de 2006, e “*A brief history of Wi-Fi*”, de 10 de junho de 2004, publicados na revista Economist. Ambos os artigos estão disponíveis eletronicamente.

unidades móveis, reproduzindo as características dos fluxos reais de voz e vídeo. Utilizamos também as ferramentas *Netperf* e *Ethereal*, que são, respectivamente, um *benchmark* e um analisador de redes, para validar os resultados obtidos com os nossos geradores de tráfego.

Nos experimentos apresentados nesta tese não utilizamos tráfegos com conteúdo real de voz e vídeo, mas sim tráfegos sintéticos. Apesar de não agregar grandes complicadores para execução dos experimentos, a utilização de tráfegos reais não se fez necessária, pois nos concentramos na distribuição dos fluxos, e não na sua exibição, ou seja, o uso de tráfegos reais não implicaria diferentes valores de atraso e de perda de pacotes, que foram os objetos de estudo. Desse modo, o uso de tráfegos sintéticos não implicou a perda de fidelidade dos experimentos realizados.

No decorrer da leitura deste trabalho, transparecerá ao leitor que ele foi conduzido de maneira predominantemente empírica. Apesar da dificuldade de condução de experimentos práticos em redes móveis, essa abordagem foi importante, porque alguns dos problemas que levantamos não teriam sido identificados se fizéssemos uso apenas de avaliações analíticas e de simulações.

1.3 Principais referências e trabalhos relacionados

Os trabalhos relacionados a esta tese podem ser classificados em três grupos: os trabalhos dedicados a IEEE 802.11, os dedicados a VoIP e os dedicados a fluxos de vídeo e a H.264/AVC. Para cada um desses grupos existe vasta literatura, mas o principal referencial teórico utilizado foram as respectivas especificações, a saber: as especificações de IEEE 802.11 e de suas extensões [IEE99], as recomendações da ITU-T para os sistemas de VoIP [ITU03, ITU90], as especificações de H.264/AVC [Joi03] e as RFCs sobre empacotamento RTP [SCFJ03, WHS⁺05].

Além dessas referências, também devemos citar alguns trabalhos que se posicionam na intersecção entre as redes sem fio IEEE 802.11 e as aplicações multimídia, tais como o de Muñoz et al., que caracterizaram o comportamento da transmissão de fluxos UDP sobre redes IEEE 802.11 [MGCM01], e o de Luthra et al., que apresentaram detalhadamente as características de H.264/AVC [LSE03] e de sua distribuição sobre redes de pacotes tradicionais [Wen03] e sem fio [SHW03]. Sobre VoIP, a principal referência não foram artigos acadêmicos ou especificações, mas sim a própria arquitetura da plataforma Skype para VoIP [BS04].

Ao longo deste texto, muitas outras referências são citadas, mas esta seção resume o referencial

teórico para este trabalho, assim como as leituras fundamentais sobre este tema.

1.4 Escopo do trabalho

O ambiente estudado restringe-se às redes IEEE 802.11 infra-estruturadas. Concentramos-nos nesse modelo — e não no modelo *ad hoc* — porque é principalmente através do modo infra-estruturado que as redes IEEE 802.11 são mais utilizadas atualmente. O modo *ad hoc* tem as suas aplicações, mas é através do modo infra-estruturado, talvez em locais públicos e quiçá gratuitamente, que esperamos ver a grande maioria das utilizações de aplicações de VoIP e de VoD. Além disso, este trabalho concentra-se no comportamento da transmissão entre os pontos de acesso e as unidades móveis, ou seja, concentramo-nos no último salto (*hop*) da transmissão.

Esta pesquisa dedica-se, principalmente, ao desenvolvimento de técnicas adaptativas para a camada de aplicação; a razão para isso é que as propostas de melhorias nas camadas inferiores dificilmente poderiam ser adotadas, pois consistiriam em modificações em protocolos já amplamente estabelecidos. Desse modo, não propomos modificações nos protocolos de transporte ou na camada MAC de IEEE 802.11, mas eventualmente apontamos características desses protocolos que não se aplicam bem à distribuição de multimídia.

1.5 Ambiente experimental

O ambiente utilizado na execução de nossos experimentos, como ilustra a Figura 1.1, foi composto na maioria da vezes por um computador (*Streaming Server* - **SS**), um ponto de acesso (*Access Point* - **AP**) e uma ou mais unidades móveis (*Mobile Units* - **MU**). O computador SS foi normalmente utilizado como servidor dos fluxos e estava conectado ao ponto de acesso AP através de uma rede Ethernet cabeada; o computador SS e o ponto de acesso AP estavam em uma mesma sub-rede. As unidades móveis MU eram conectadas ao ponto de acesso AP através de interfaces IEEE 802.11. Foram utilizados pontos de acesso de diferentes fabricantes e diversos modelos de interfaces IEEE 802.11. Além disso, nos resultados apresentados nesta tese, os mecanismos de segurança de IEEE 802.11, por exemplo WEP e WPA, foram desabilitados para não influenciarem os resultados.

Além disso, para os experimentos que envolviam a execução de migrações (*handoffs*), foram

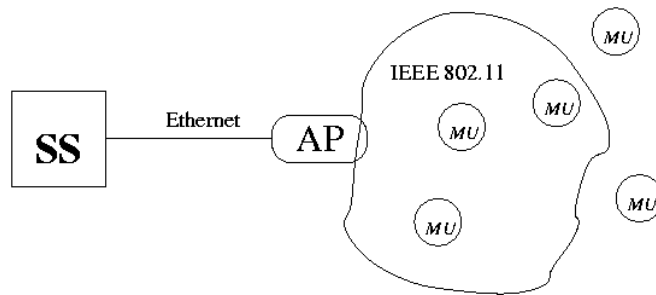


Figura 1.1: Ambiente de testes utilizado.

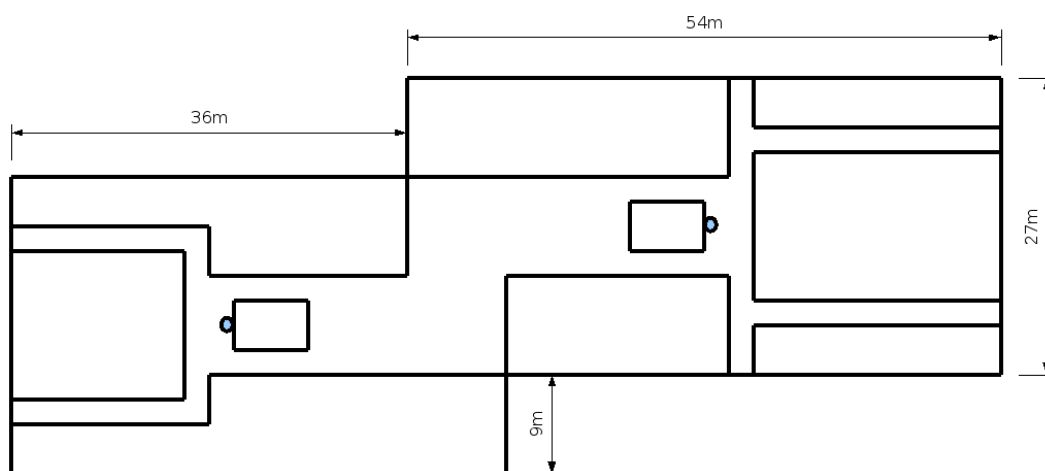
utilizadas as redes sem fio do Instituto de Matemática e Estatística da Universidade de São Paulo e da Microsoft Research. A Figura 1.2 ilustra as dimensões dos ambientes utilizados, a localização dos pontos de acesso é representada por pontos. Nos experimentos com ocorrência de migrações, um computador móvel era conduzido da área de cobertura de um ponto de acesso a outro. Na Figura 1.2(a), este movimento consistiu em ir e vir de um ponto de acesso ao outro; na Figura 1.2(b), consistiu em círculos ao redor dos corredores mais externos.

1.6 Contribuições e perspectivas

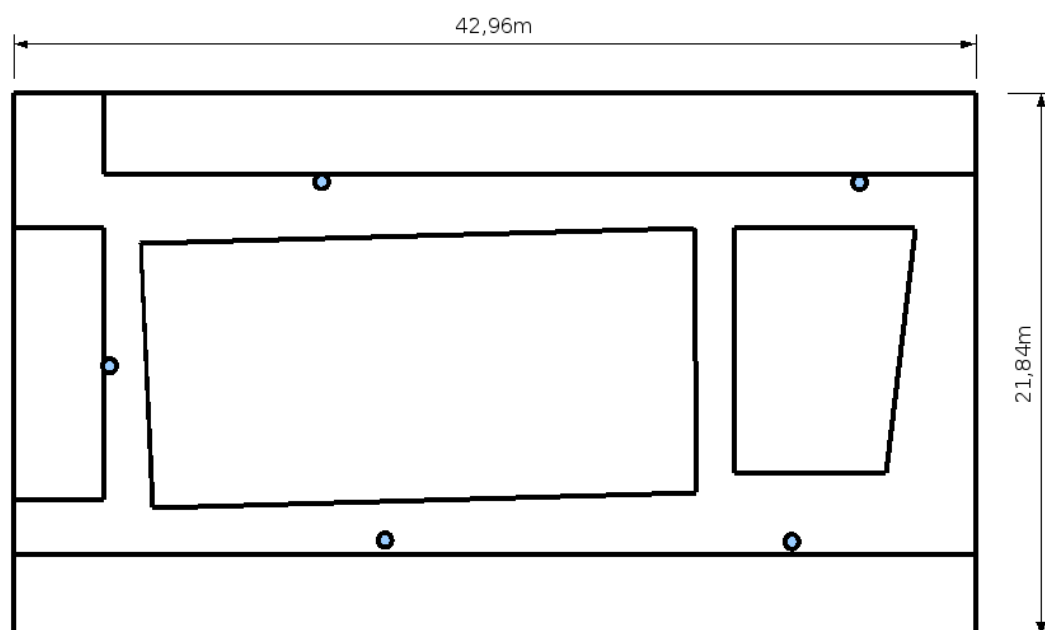
Por tratar-se de um tema recente, o número de referências a respeito de transmissão de multimídia sobre redes sem fio é pequeno. As fontes existentes são normalmente superficiais ou tratam de produtos específicos. Desse modo, temos a perspectiva de que este trabalho possa servir como referência aos interessados em desenvolver aplicações multimídia para redes sem fio e, em especial, para as redes IEEE 802.11. Além disso, esperamos contribuir com a caracterização dos problemas principais e a proposta de novos algoritmos adaptativos para detecção e solução desses problemas.

1.7 Organização do texto

A fim de atingir os objetivos definidos acima, o restante deste trabalho está organizado da seguinte maneira: o Capítulo 2, devido à importância das redes sem fio IEEE 802.11 para o tema desta tese, recapitula as principais características dessas redes e apresenta o funcionamento da camada MAC e dos principais mecanismos adaptativos presentes no protocolo.



(a) Instituto de Matemática e Estatística.



(b) Microsoft Research.

Figura 1.2: Representação dos ambientes utilizados para execução dos experimentos.

Os Capítulos 3 e 4 são dedicados, respectivamente, à transmissão de voz e de fluxos de vídeo sobre redes IEEE 802.11. O Capítulo 3 descreve os principais problemas relacionados ao uso de VoIP sobre essas redes. Por sua vez, o Capítulo 4 apresenta a arquitetura e o principal padrão de codificação utilizado para transmissão de fluxos de vídeo — o padrão H.264/AVC. Traz, também, uma proposta de adequação dos pontos de acesso IEEE 802.11 para a distribuição de vídeo H.264/AVC.

No Capítulo 5 são apresentados os principais problemas relacionados ao desenvolvimento de aplicações para redes IEEE 802.11 e também algumas das técnicas que podem ser utilizadas para resolvê-los. Além disso, um novo algoritmo para o uso da qualidade de sinal como métrica de adaptação é ali proposto.

O Capítulo 6 apresenta o uso do mecanismo de pares de pacotes — até então utilizado apenas em redes cabeadas — para a monitoração das taxas de transmissão efetivamente exercidas por IEEE 802.11.

Por fim, o Capítulo 7 traz nossas considerações finais e nossa visão de trabalhos futuros.

Capítulo 2

Introdução às redes IEEE 802.11

As redes sem fio IEEE 802.11, que também são conhecidas como redes Wi-Fi, foram uma das grandes novidades tecnológicas dos últimos anos. Atualmente, IEEE 802.11 é o padrão *de facto* em conectividade sem fio para redes locais. Como prova desse sucesso podemos citar o crescente número de *Hot Spots* e o fato de a maioria dos computadores portáteis novos já saírem de fábrica equipados com interfaces IEEE 802.11. Os *Hot Spots*, presentes nos centros urbanos e principalmente em locais públicos de lazer e de serviços, estão mudando o perfil de uso da Internet e — inclusive — dos usuários de computadores.

Desse modo, motivado pela crescente importância das redes IEEE 802.11 e considerando a recorrente referência a essa tecnologia no decorrer deste trabalho, este capítulo apresenta as principais características dessas redes, faz uma revisão cronológica do seu desenvolvimento e descreve seus principais mecanismos adaptativos e parâmetros de configuração. Além disso, para concluir, traçamos nossas considerações sobre o futuro dessas redes.

2.1 Cronologia de desenvolvimento do padrão IEEE 802.11

Antes de apresentar o padrão IEEE 802.11, vejamos uma breve cronologia do seu desenvolvimento:

- 1989: o *Federal Communications Commission* (FCC), órgão americano responsável pela regulamentação do uso do espectro de frequências, autorizou o uso de três faixas de frequência;
- 1990: o *Institute of Electrical and Electronics Engineers* (IEEE) instaurou um comitê para

definição de um padrão para conectividade sem fio;

- 1997: após sete anos de pesquisa e desenvolvimento, o comitê de padronização da IEEE aprovou o padrão IEEE 802.11; nessa versão inicial, as taxas de transmissão nominal eram 1 e 2 Mbps;
- 1999: foram aprovados os padrões IEEE 802.11b e 802.11a, que usam as frequências de 2,4 e 5 MHz e são capazes de atingir taxas nominais de transmissão de 11 e 54 Mbps, respectivamente. O padrão 802.11b, apesar de atingir taxas de transmissão menores, ganhou fatias maiores de mercado do que 802.11a; as razões para isso foram basicamente duas: primeiro, as interfaces 802.11b eram mais baratas do que as 802.11a e, segundo, as implementações de 802.11b foram lançadas no mercado antes do que as implementações de 802.11a. Além disso, nesse ano foi criada a *Wireless Ethernet Compatibility Alliance* (WECA), que se organizou com o objetivo de garantir a interoperabilidade entre dispositivos de diferentes fabricantes;
- 2000: surgiram os primeiros *hot spots*, que são áreas públicas onde é possível acessar a Internet por meio das redes IEEE 802.11. A WECA lançou o selo *Wireless Fidelity* (Wi-Fi) para atestar a aderência dos produtos às especificações; mais tarde o termo Wi-Fi tornou-se um sinônimo de uso abrangente das tecnologias IEEE 802.11;
- 2001: a companhia americana de cafeterias Starbucks implementou *hot spots* em sua rede de lojas. Os pesquisadores Scott Fluhrer, Itsik Mantin e Adi Shamir demonstraram que o protocolo de segurança *Wired Equivalent Privacy* (WEP) é inseguro [FMS01];
- 2002: a WECA passou a se chamar Wi-Fi *Alliance* (WFA) e lançou o protocolo Wi-Fi *Protected Access* (WPA) em substituição ao protocolo WEP;
- 2003: o comitê de padronização da IEEE aprovou o padrão IEEE 802.11g que, assim como 802.11b, trabalha na faixa de 2,4 MHz, mas alcança até 54 Mbps de taxa nominal de transmissão. Aprovou também, sob a sigla IEEE 802.11f, a recomendação de práticas para implementação de *handoff*;
- 2004: a especificação 802.11i aumentou consideravelmente a segurança, definindo melhores procedimentos para autenticação, autorização e criptografia [HM05];

- 2005: foi aprovada a especificação 802.11e, agregando qualidade de serviço (QoS) às redes IEEE 802.11. Foram lançados comercialmente os primeiros pontos de acesso trazendo pré-implementações da especificação IEEE 802.11e;
- 2006: surgiram as primeiras implementações do padrão 802.11n, que usa múltiplas antenas para transmissão e recepção, *Multiple-Input Multiple-Output* (MIMO), atingindo taxa nominal de transmissão de até 600 Mbps.

Ademais, o desenvolvimento continua. Ainda estão em pleno desenvolvimento as seguintes especificações: 802.11k, métricas para avaliação da conectividade; 802.11p, redes sem fio para ambientes veiculares (alta mobilidade); 802.11r, otimização do *handoff*; 802.11s, interconexão automática entre pontos de acesso (*Mesh Networks*); 802.11t, avaliação de desempenho de redes sem fio; 802.11u, uniformização da interconexão entre redes; 802.11v, gerenciamento de redes; e, por fim, 802.11w, cuja finalidade é aumentar a proteção dos quadros de transmissão contra interferências externas. Além disso, estima-se que neste ano as vendas de dispositivos com interface IEEE 802.11 atinjam dezenas de milhões de unidades (revista *Economist*, de 8 de dezembro de 2005). Além dos computadores portáteis, as interfaces IEEE 802.11 também devem equipar telefones celulares, câmeras digitais, impressoras, aparelhos tocadores de MP3 etc. Portanto, as redes IEEE 802.11 certamente alcançarão novos nichos de mercado.

2.2 Modos de operação: infra-estruturado e *ad hoc*

Existem dois modos de operação para as redes IEEE 802.11: o modo *ad hoc* e o infra-estruturado. O modo *ad hoc* caracteriza-se pela ausência de coordenação central: as unidades móveis podem formar suas redes, associando-se livremente umas às outras. Um exemplo de utilização de IEEE 802.11 em modo *ad hoc* ocorre durante as reuniões de negócios, em que os participantes formam redes *ad hoc* entre seus computadores para disponibilizar e trocar arquivos.

Mas, apesar da aplicabilidade das redes *ad hoc*, atualmente são as redes infra-estruturadas as principais responsáveis pela grande utilização das redes IEEE 802.11. O modo infra-estruturado caracteriza-se pela presença de um nó central — chamado de ponto de acesso — para a coordenação e distribuição dos tráfegos; dessa forma, todo o tráfego de dados passa através do ponto de acesso;

os *Hot Spots*, por exemplo, normalmente disponibilizam o acesso à Internet usando o modo infraestrutura.

2.3 A camada MAC de IEEE 802.11

A camada MAC de IEEE 802.11 define dois métodos de acesso: *Distributed Coordination Function* (DCF) e *Point Coordination Function* (PCF). Sendo DCF amplamente mais difundido do que PCF, a seguir apresentamos apenas o método de acesso DCF.

O DCF consiste na implementação de um mecanismo de acesso bem conhecido na indústria, o mecanismo *Carrier Sense Multiple Access* (CSMA). A aplicação mais popular de CSMA são as redes Ethernet. Mais especificamente, as redes Ethernet implementam CSMA/CD (*Carrier Sense Multiple Access/Collision Detection*), enquanto as redes IEEE 802.11, em função das limitações que veremos a seguir, implementam CSMA/CA (*Carrier Sense Multiple Access/Collision Avoidance*).

O protocolo CSMA funciona da seguinte maneira: a estação “ouve” o meio de comunicação antes de iniciar a transmissão; se o meio está livre, a estação transmite; senão, a estação espera até que o meio de transmissão esteja livre. Esse modelo de protocolo é eficiente quando o meio de transmissão não está sobrecarregado, mas, mesmo se o meio de comunicação estiver subutilizado, ainda existe a possibilidade de que duas ou mais transmissões ocorram ao mesmo tempo.

Se duas estações perceberem o meio ocioso ao mesmo tempo, existe a possibilidade de ambas as estações transmitirem ao mesmo tempo, daí as colisões. Para que o protocolo funcione corretamente, essas colisões devem ser detectadas e os pacotes, retransmitidos. Porém, a detecção de colisões, apesar de funcionar bem em redes cabeadas, é de difícil implementação em redes sem fio por duas razões:

- para implementar a detecção de colisões, é preciso que o dispositivo de rádio possa transmitir e receber ao mesmo tempo (*Full Duplex*), o que aumentaria significativamente os custos de produção de interfaces, pois implica a utilização de duas antenas que não produzam interferência entre si;
- em redes sem fio, não se pode garantir que todos os nós se ouçam (princípio fundamental para detecção de colisões). Ou seja, o fato de uma estação não perceber atividade no meio

de comunicação não significa que o meio esteja completamente livre, pode ser que estejam ocorrendo transmissões, mas essas transmissões simplesmente não estão alcançando a estação. Essas estações, cuja transmissão não alcança todos os nós, são conhecidas como nós escondidos (*hidden nodes*).

A fim de contornar as limitações acima, ao invés de implementar a detecção de colisões, IEEE 802.11 implementa mecanismos para evitar colisões e — somados a isso — mecanismos para identificação positiva da entrega de mensagens.

O mecanismo de *Collision Avoidance* (CA) possui a seguinte estrutura de funcionamento: uma estação, antes de transmitir, ouve o meio; se este está ocupado, a estação adia a sua transmissão usando um mecanismo de *Backoff* aleatório exponencial [Gas02]. Se o meio permanece livre por certo período de tempo — chamado de *Distributed Inter Frame Space* (DIFS) —, a estação transmite. A estação que recebe a transmissão verifica o pacote recebido usando *Cyclic Redundancy Check* (CRC) [HV02a] e, caso o pacote tenha sido recebido sem erros, envia uma mensagem ACK de confirmação de recebimento para o transmissor.

A recepção da mensagem ACK indica ao transmissor que a transmissão foi bem sucedida e que, portanto, no seu decorrer, não houve colisões ou interferências. Caso a mensagem não seja recebida, o transmissor tentará reenviar a mensagem, até que a recepção seja confirmada; ou até que seja excedido o limite máximo de tentativas de reenvio, como descreve a Seção 2.4.1.

2.4 Mecanismos e configuração de IEEE 802.11

O protocolo IEEE 802.11, com a finalidade de otimizar o uso do meio de transmissão, define alguns mecanismos adaptativos e parâmetros, cuja configuração permite a maior adequação do protocolo a certas aplicações e cenários de uso. Estes mecanismos têm grande relevância nas discussões conduzidas nesta tese e por essa razão serão apresentados a seguir; também serão apresentados os parâmetros de configuração mais relevantes.

2.4.1 Retransmissão automática

Um dos mecanismos de IEEE 802.11 mais referenciados neste texto é o de retransmissão automática. Em IEEE 802.11, antes de um pacote ser dado por perdido, a camada MAC tenta reenviá-lo. Existem dois limites para o número de tentativas de reenvio de um pacote: *short retry limit* e *long retry limit*. O primeiro, *short*, aplica-se aos quadros menores que o parâmetro *RTS Threshold*. O segundo, *long retry limit*, reserva-se aos quadros maiores. Os valores-padrão de *short retry limit* e de *long retry limit* são, respectivamente, sete e quatro; e o de *RTS Threshold* é de 2.347 octetos.

Uma das implicações do mecanismo de retransmissão automática é que a taxa de perda de pacotes não reflete a realidade do canal: afinal, um pacote é reenviado pelo menos quatro vezes antes da sua perda ser notificada ao protocolo de nível superior. Desse modo, uma métrica mais apropriada do que a taxa de perda de pacotes seria a taxa de perda de quadros.

2.4.2 Fragmentação

Os pacotes enviados sobre IEEE 802.11 podem ser divididos em partes menores — chamadas de quadros de transmissão — antes de serem transmitidos pela camada MAC. O objetivo da fragmentação dos pacotes em quadros é aumentar a probabilidade de entrega de um pacote, pois, quanto menores os pacotes, maiores as chances de sucesso da transmissão [Sch00]. Além disso, a perda de um quadro exige a retransmissão apenas do quadro perdido e não do pacote inteiro, reduzindo o tempo gasto em retransmissões.

A fragmentação é controlada pelo parâmetro *Fragmentation Threshold*, cujo valor máximo é de 4.096 octetos, sendo que o valor-padrão é de 2.346 octetos. Ou seja, o tamanho máximo de um quadro de transmissão é de 4.096 octetos e pacotes maiores do que 2.346 octetos são normalmente fragmentados em dois ou mais quadros de transmissão.

A configuração do parâmetro de fragmentação tem implicações diretas na taxa de perda de pacotes, pois quanto maiores os quadros, maiores serão as chances de ocorrência de colisões e de interferências. Por outro lado, quanto menores os quadros, menor será a eficiência de utilização do meio de transmissão, pois a fragmentação, ao aumentar o número de pacotes, aumenta também as sobrecargas: tais como o número de cabeçalhos e de mensagens ACK transmitidas. Na prática, o tamanho dos pacotes usados nas aplicações também deve levar em consideração o tamanho máximo

dos quadros de transmissão da camada Ethernet, como veremos na Seção 6.5.

2.4.3 Requisição e reserva do canal de transmissão: mecanismo RTS/CTS

Devido à existência de *hidden nodes*, pode ser necessária a requisição de uso exclusivo do meio de comunicação por parte de uma certa unidade móvel. Essa requisição é feita usando-se um mecanismo conhecido como *Request To Send/Clear To Send* (RTS/CTS). Esse recurso é especialmente útil em ambientes sujeitos a altas taxas de colisões ou na transmissão de pacotes grandes.

Nas versões preliminares de IEEE 802.11, o uso do mecanismo RTS/CTS era obrigatório, mas após algumas revisões o seu uso tornou-se opcional, sendo empregado apenas quando necessário ou requisitado. Seu funcionamento ocorre da seguinte maneira: o parâmetro *RTS Threshold*, cujo valor-padrão é 2.347 octetos, define o tamanho máximo de um quadro que pode ser enviado sem a execução do mecanismo RTS/CTS. Para enviar quadros maiores do que *RTS Threshold*, o transmissor envia uma mensagem *RTS* (*Request To Send*) e aguarda uma mensagem *CTS* (*Clear To Send*), para — só então — enviar o quadro.

Quanto menor o valor de *RTS Threshold*, maior a utilização do mecanismo *RTS/CTS* e maior a troca de mensagens, o que implica sobrecargas que reduzem o aproveitamento da capacidade da rede. Por outro lado, a utilização do mecanismo reduz a incidência de colisões, diminuindo a perda de pacotes.

2.4.4 Modo de economia de energia

O uso de redes sem fio está atrelado ao uso de computadores móveis, que por sua vez dependem de baterias para alimentação de energia. Por essa razão, IEEE 802.11 definiu um modo de operação para economia de energia: o *Power Save Mode* (PSM). Antes de uma unidade móvel “dormir”, ela notifica o ponto de acesso, que passa a armazenar os pacotes que forem enviados para essa unidade móvel, de tal forma que estes não sejam perdidos.

O ponto de acesso periodicamente divulga quais estações possuem pacotes armazenados, através de pacotes especiais de sinalização; estes são chamados de *Beacons*. Então, ao “acordarem”, as unidades móveis podem verificar o conteúdo desses *Beacons* e, caso constatem a existência de pacotes armazenados para elas no ponto de acesso, podem recuperar esses pacotes. Mas, obviamente, esses

pacotes armazenados nos pontos de acesso provavelmente não serão mais úteis para as aplicações de tempo real, como as aplicações de VoIP.

2.4.5 Migração (*handoff*)

Assim como nas redes celulares, o processo de migração (*handoff*) em redes IEEE 802.11 consiste em mover-se de uma área de cobertura para outra, ou seja, consiste em mover-se de um ponto de acesso para o outro. Idealmente, os *handoffs* deveriam ser realizados sem a perda de conexão ou de dados, mas veremos no Capítulo 3 que essas perdas são expressivas.

A política e os passos para execução de *handoffs* estão delineados no padrão IEEE 802.11f. Os passos básicos para execução de um *handoff* são: *scanning* e reassociação. O *scanning* consiste na busca por um novo ponto de acesso, podendo ser passivo (unidade móvel apenas “escuta”) ou ativo (unidade móvel manda uma mensagem requisitando resposta dos pontos de acesso e então “escuta” as respostas). Esses procedimentos, no entanto, estão sendo aperfeiçoados na especificação 802.11r (*fast handoff*), que ainda está em desenvolvimento.

2.4.6 Controle automático de taxa de transmissão

O mecanismo para controle automático da taxa de transmissão (*Automatic Rate Control* ou ARC) é o mecanismo responsável por escolher a modulação adequada às condições do canal, com o objetivo de explorar o compromisso entre a taxa de erros e a capacidade de transmissão. O princípio de atuação do mecanismo é: quanto pior a situação do canal, mais robusta deve ser a modulação utilizada (menor a taxa de transmissão); quanto melhor o canal, menos robusta pode ser a modulação utilizada (maior a taxa de transmissão). Esse mecanismo será descrito em detalhes no Capítulo 6.

2.4.7 Outros parâmetros configuráveis de IEEE 802.11

Além dos mecanismos adaptativos e parâmetros de configuração descritos acima, nas especificações do protocolo [IEE99] ou nos inúmeros livros dedicados ao assunto [Gas02, Fli03, Rey03], podem ser encontrados outros parâmetros configuráveis. Sobre os parâmetros e a forma de configuração de IEEE 802.11, recomendamos a leitura do capítulo 17 de [Gas02] e da documentação das espe-

cificações da *Network Driver Interface Specification* (NDIS, vide <http://msdn.microsoft.com/library>).

2.5 O futuro de IEEE 802.11

Por fim, tendo em vista os avanços tecnológicos e as tendências de uso de redes de computadores, gostaríamos de apresentar alguns fatores que apontam para a crescente utilização das redes 802.11:

- redução gradual dos custos e maior aderência dos produtos aos padrões;
- disponibilização de interfaces para programação e controle dos dispositivos (NDIS) e transferência de parte das funcionalidades da camada MAC, hoje geralmente implementadas em *hardware*, para as camadas de *software* (Native Wi-Fi);
- crescimento do número de *hot spots* gratuitos em locais públicos e também em locais onde a circulação de pessoas é restrita, tais como escritórios e condomínios;
- lançamento de dispositivos móveis — principalmente telefones celulares e computadores de mão — com capacidade para multiconexão transparente, o que significa capacidade de estar conectado sempre utilizando a melhor conexão dentre as diferentes tecnologias de rede disponíveis;
- implementação de redes sem fio de alta velocidade e de cobertura metropolitana (redes WiMax [IEE04]);
- lançamento de novos serviços com forte apelo de utilização sobre redes sem fio, tais como: serviços baseados em localização (*Location Based Services* ou LBS), VoIP, vídeo email, TV digital e interativa com canais de programação independente, dinheiro eletrônico etc.

Além disso, podemos contar com o avanço tecnológico das redes IEEE 802.11. Segundo os comitês de especificação da IEEE, os padrões que ainda estão em desenvolvimento estarão todos finalizados até 2008. Desse modo, em poucos anos, contaremos com redes com capacidade nominal de transmissão de até 600 Mbps, seguras e presentes em todas as regiões metropolitanas.

Capítulo 3

Voz sobre redes IEEE 802.11

Homens de países remotíssimos falarão uns com os outros e se responderão.

Leonardo da Vinci

Na última década, Voz sobre IP (VoIP) deixou de ser uma tecnologia usada apenas em determinados nichos de mercado para tornar-se uma tecnologia à disposição da grande massa de usuários finais. Empresas como Vonage, AT&T e Skype já oferecem serviços baseados em VoIP para milhões de clientes ao redor do mundo. Durante esse mesmo período, em paralelo à popularização de VoIP, a utilização das redes locais sem fio baseadas no padrão IEEE 802.11 (Wi-Fi) cresceu gradualmente e este tornou-se o padrão *de facto* para redes locais sem fio. A cada dia, seja na cafeteria ou no centro de compras, essas redes estão cada vez mais presentes no nosso cotidiano. Desse modo, apesar das redes IEEE 802.11 ainda não oferecerem suporte adequado às aplicações de tempo real, seu relativo baixo custo e sua ampla utilização fizeram de IEEE 802.11 a alternativa natural para se agregar mobilidade às aplicações de VoIP.

Voz sobre Wi-Fi, também conhecido como VoWiFi ou VoWLAN, atraiu o interesse da comunidade acadêmica e da indústria, principalmente por fornecer uma alternativa de baixo custo para comunicações móveis. Tanto a academia quanto a indústria propuseram extensões às especificações originais de IEEE 802.11, com o intuito de melhorar o seu desempenho e adequá-lo aos requisitos das aplicações de tempo real; algumas soluções proprietárias de *hardware* já foram inclusive implementadas, como os produtos advindos da parceria entre Linksys e Vonage.

No entanto, apesar do uso de soluções proprietárias permitirem a criação de soluções interessantes, neste trabalho estamos particularmente interessados em investigar a utilização de VoWiFi que se faz possível usando-se apenas a tecnologia atualmente disponível, ou seja, com base nos padrões consolidados. Ou seja, queremos saber, dados os cenários atuais, quais são os principais problemas que afetam negativamente as sessões de comunicação de voz sobre IEEE 802.11 e como esses problemas limitam o uso de VoWiFi. Com esse objetivo, este capítulo apresenta um resumo desses problemas e, além disso, propõe cenários — limitados, mas importantes — nos quais VoWiFi poderia ser implementado usando-se as soluções de rede sem fio atualmente disponíveis.

Antes de prosseguir, cabe dizer que este trabalho concentra-se no uso de IEEE 802.11 no último salto de acesso à Internet. Mais especificamente, concentra-se no comportamento dos pacotes de voz sobre redes IEEE 802.11 infra-estruturadas. Não discutimos os padrões e pilhas de protocolos utilizados. Os problemas não diretamente relacionados à entrega de pacotes, tais como as questões de consumo de energia e de interconexão com a rede pública de telefonia, não fazem parte do escopo de nossa investigação.

Este capítulo está organizado da seguinte maneira: a fim de familiarizar o leitor com as aplicações de voz, a Seção 3.1 resume os requisitos das aplicações de VoIP. A Seção 3.2 apresenta a metodologia usada para obtenção dos traços de execução de VoWiFi. Na Seção 3.3, esses traços são analisados, dando atenção especial a dois problemas que nos pareceram ser os principais: os *handoffs* e os tráfegos em rajada. Apresentados os problemas principais, a Seção 3.4 expõe alguns problemas adicionais que podem dificultar a ampla implementação de VoWiFi. A Seção 3.5 discorre sobre aspectos de VoWiFi que devem ser levados em conta por pesquisadores e desenvolvedores. A Seção 3.6 recapitula os trabalhos relacionados. A Seção 3.7, baseada nas constatações e observações feitas nas seções anteriores, propõe cenários onde acreditamos que VoWiFi pode ser utilizado com as tecnologias atuais e também oferece sugestões de como os problemas atuais podem ser contornados. Por fim, apresentamos as conclusões sobre o uso de VoWiFi e apontamos alguns trabalhos futuros que seriam relevantes.

3.1 Requisitos das aplicações de VoIP

Antes de apresentar como as características das redes IEEE 802.11 afetam VoIP, façamos uma revisão das métricas normalmente usadas para se definir a qualidade de uma sessão de VoIP; são elas: latência, *jitter*, taxa de perda de pacotes e vazão.

3.1.1 Latência

A latência, também conhecida como atraso unidirecional (*one-way delay*), é a métrica mais crítica para as aplicações de VoIP; trata-se do tempo necessário para que uma mensagem de voz seja transmitida do emissor para o receptor. Se a latência é longa, a conversação é comprometida e a comunicação pode tornar-se não natural, não interativa, apresentando cortes e correções audíveis. De acordo com as recomendações da ITU-T [ITU03], a latência pode acumular até 150 milissegundos sem que haja grandes impactos na qualidade da ligação. Porém, a partir de 150ms de latência, as conseqüências negativas começam a ser gradualmente perceptíveis. Atrasos de até 200ms ainda são definidos na literatura como aceitáveis [JCG04].

Em redes IEEE 802.11, considerando-se apenas um salto de conexão, a latência é normalmente, mesmo com a necessidade de retransmissões, menor do que 10ms [AAMM01]. Sendo assim, a latência da conexão IEEE 802.11, que é da ordem de poucos milissegundos, é muito menor do que a latência total de uma comunicação pela Internet, que é da ordem de dezenas de milissegundos [Cis02]. Ou seja, apesar de a latência ser uma métrica crítica, a latência média de IEEE 802.11 não é um problema para o uso de VoWiFi.

3.1.2 Jitter

O *jitter* é a variação da latência. A maioria dos sistemas multimídia modernos utilizam mecanismos adaptativos baseados em *buffers* para suavizar os efeitos dessa variação [LFG03]. Porém, o uso de *buffers* aumenta a latência média e pode reduzir a interatividade da comunicação e introduzir artefatos¹ na voz. Em redes IEEE 802.11, o *jitter* é normalmente pequeno [AAMM01]. Apesar

¹Artefato é o termo usado para denotar as falhas perceptíveis que podem ocorrer quando uma transmissão de voz e de vídeo é reproduzida. Os artefatos mais comumente percebidos nas aplicações de voz são os cortes na fala — aquelas pequenas interrupções. Nas aplicações de vídeo, o artefato mais comum é a exibição de pequenos quadrados

disso, a Seção 3.3.2 mostrará que podem ocorrer variações extremamente altas na latência, as quais possuem impacto relevante na qualidade da voz transmitida.

3.1.3 Taxa de perda de pacotes

A taxa de perda de pacotes expressa a quantidade de pacotes não entregues. Em aplicações de VoIP, a taxa de perda de pacotes é considerada aceitável até a razão de 1% [ITU03, JCG04]. No entanto, acima de 1%, os impactos na qualidade da voz normalmente tornam-se perceptíveis, surgem artefatos audíveis que dificultam a comunicação. Em redes IEEE 802.11, apesar de as taxas reais de perda de pacotes serem mascaradas pelo mecanismo de ARQ [Gas02, dCK03], a taxa de perdas de pacotes observada é normalmente menor do que 1% [AAMM01], sendo assim aceitável para as aplicações de VoIP. Porém, se uma unidade móvel se afasta da área de cobertura de um ponto de acesso, a taxa de perdas de pacotes pode crescer abruptamente, impossibilitando a conectividade e a comunicação.

3.1.4 Vazão

A vazão é o volume de dados por segundo necessário para o tráfego dos pacotes de voz. A vazão necessária para uma conexão de VoIP é muito menor do que a capacidade nominal das redes IEEE 802.11. Uma sessão típica de voz requer aproximadamente 64 Kbps, enquanto 802.11g oferece 54 Mbps. Entretanto, se um ponto de acesso é utilizado para transmitir múltiplas chamadas de voz, pode haver um problema de capacidade. A Seção 3.4.1 discutirá a questão de capacidade em mais detalhes.

Desse modo, as redes IEEE 802.11 parecem ser — com certa margem de folga — apropriadas para VoIP, pois os valores médios de latência, *jitter*, vazão e a taxa de perda de pacotes dessas redes indicam que VoWiFi deveria poder ser usado sem maiores problemas. Entretanto, as próximas seções enumeram características das redes IEEE 802.11 atuais que, apesar de não afetarem o comportamento “médio”, consistem em barreiras significativas para a ampla utilização de VoWiFi.

que nitidamente destoam da imagem.

3.2 Metodologia para coleta dos traços de execução de VoWiFi

Para a análise dos tráfegos de VoWiFi, implementamos programas geradores de tráfego nas linguagens C e Java que simulam o tráfego de uma sessão de VoIP. Os programas transmitem fluxos bidirecionais de pacotes à taxa de 50 pacotes por segundo; ou seja, um pacote a cada 20ms. A carga útil dos pacotes é de 60 octetos. Dessa forma, simulamos uma sessão de VoIP de 24 Kbps sem supressão de silêncio.

Além disso, para evitar que as especificidades de um ou outro dispositivo Wi-Fi atrapalhassem as medições, executamos os geradores de tráfego em diferentes cenários e ambientes de rede. Também utilizamos diferentes interfaces IEEE 802.11 (*Network Interface Cards* ou NICs) — modelos WMP55AG, AIR-PCM340, WUSB54GP Ver. 4, WUSB54AG, WUSB54G Ver. 4, WPC54G Ver. 1.2, e PC24E-H-FC — e cinco redes IEEE 802.11 diferentes, sendo três dedicadas (usando os pontos de acesso: Netgear FWAG114, Dlink DI624+ e Linksys WRT54GS) e duas não dedicadas (uma da Microsoft Research e outra da Universidade de São Paulo).

Os programas geradores de tráfego, a cada pacote recebido, registram: o instante de recebimento do pacote, a potência ou força do sinal (*signal strength*), a perda de pacotes e o atraso. Do instante de recebimento, derivamos um atraso relativo, o Intervalo Entre Pacotes (IEP), que consiste na diferença de tempo entre a chegada de dois pacotes consecutivos. Se a diferença entre o número de sequência de dois pacotes recebidos é k , o intervalo esperado entre esses dois pacotes é $k \times 20ms$. A força do sinal foi obtida através da API de NDIS (vide Capítulo 5). A perda de pacotes foi detectada comparando-se o número de sequência de dois pacotes consecutivos (o número de sequência está registrado no cabeçalho RTP de cada pacote). A medição do atraso, porém, foi não trivial e será explicada a seguir.

3.2.1 Método para estimar o atraso

Para obter o valor absoluto do atraso de cada pacote são necessários os instantes de tempo em que cada pacote foi enviado e recebido. Marcar o instante (*timestamping*) de envio e de chegada de um pacote é simples; o problema é que essas marcas de tempo pertencem a computadores diferentes e, para serem corretamente interpretadas, precisam ser sincronizadas.

Além disso, nos computadores atuais, a contagem de tempo é feita por cristais de quartzo

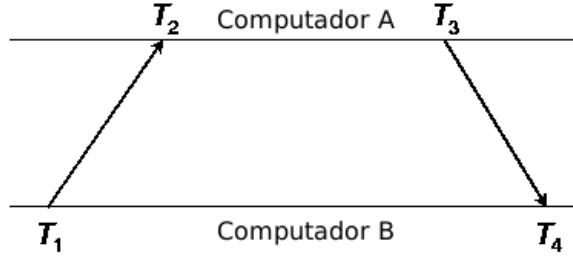


Figura 3.2: Marcas de tempo usadas para obtenção do atraso.

drifting.

Sendo as marcas de tempo T_1 , T_2 , T_3 e T_4 tomadas como ilustra a Figura 3.2, ou seja, T_1 e T_4 obtidas em B e T_2 e T_3 obtidas em A , queremos determinar o tempo de transmissão — o atraso — do computador A para o computador B . Lembrando que os pacotes de VoIP são continuamente enviados de A para B , podemos assumir que o i -ésimo pacote sai de A no instante $T_{i,3}$ e chega em B em $T_{i,4}$. Além disso, podemos assumir que o computador A reporta para o computador B , através do cabeçalho estendido de RTP, a marca de tempo $T_{i,2}$ e o último pacote recebido de B . Observe que, em função do número de seqüência do último pacote recebido de B no computador A , o computador B pode determinar $T_{i,1}$.

Para o i -ésimo pacote recebido em B , calculamos o deslocamento de tempo θ_i entre os relógios dos computadores A e B da seguinte maneira:

$$\theta_i = 1/2 [(T_{i,2} + T_{i,3}) - (T_{i,1} + T_{i,4})], \quad (3.2)$$

Além disso, para o i -ésimo pacote recebido em B , calculamos também o tempo total de ida e volta de uma mensagem (*round trip delay*) δ_i entre os computadores A e B :

$$\delta_i = (T_{i,4} - T_{i,1}) - (T_{i,3} - T_{i,2}). \quad (3.3)$$

Pesquisas sobre o protocolo NTP [Mil95] mostram que o valor de θ_i é tão preciso quanto menores forem os valores do atraso δ_i . Sendo assim, armazenamos os 32 pares θ_i e δ_i mais significativos, ou seja, os 32 pares dentre os últimos 100 que possuem o menor valor de δ_i .

Esses 32 pares mais significativos e suas respectivas marcas de tempo T_3 e T_4 são usados para encontrar, através do método de erros quadráticos mínimos [PTVF02] (*Least Square Error* ou

LSE), os coeficientes θ_0 , β e χ da Equação 3.1, os quais são utilizados para converter a marca de envio $T_{i,3}$ em uma marca de tempo no computador B , calculando assim o atraso absoluto referente à transmissão de um pacote. Mais detalhes sobre essa abordagem de conversão podem ser consultados em [dCLFK06].

3.3 Obstáculos para a implantação de VoWiFi

A partir da análise dos traços de execução, identificamos os principais obstáculos para a implantação de VoWiFi: a migração de ponto de acesso (*handoffs*) e o comportamento bloqueio/rajada das redes IEEE 802.11.

3.3.1 *Handoff*

Espera-se que os usuários de dispositivos IEEE 802.11 possam mover-se durante uma chamada de voz. Sendo assim, dado que o raio de cobertura de um ponto de acesso em espaços abertos é de aproximadamente 100m, espera-se também a ocorrência de *handoffs* durante as chamadas. Entretanto, os traços de execução mostraram que a migração em redes IEEE 802.11 não é suficientemente suave a ponto de não ser percebida durante uma sessão de VoWiFi.

A Figura 3.3 demonstra o comportamento típico de um *handoff*. Nessa execução, o traço foi registrado enquanto a unidade móvel era movida através de uma rede sem fio corporativa composta por vários pontos de acesso IEEE 802.11. As Figuras 3.3(a) e 3.3(b) mostram, respectivamente, o comportamento dos atrasos e das perdas de pacotes. Na Figura 3.3(a), o eixo vertical esquerdo demonstra a variação nos atrasos, o eixo horizontal mostra o tempo do experimento e o eixo vertical direito, com o objetivo de ilustrar o movimento da unidade móvel, mostra a força do sinal. Por sua vez, na Figura 3.3(b), o eixo vertical esquerdo dimensiona as perdas de pacotes.

Nas figuras, podem ser observadas quatro operações de *handoff*: aos 251s, 310s, 336s e 382s, aproximadamente. O momento em que os *handoffs* são executados caracteriza-se por um salto na força do sinal, pois o NIC desconecta-se do ponto de acesso de sinal mais fraco e conecta-se a outro de sinal mais forte — e provavelmente mais próximo. Também se podem observar duas consequências principais dos *handoffs*: a incidência de perda de pacotes e o aumento no atraso. Por sua vez, a consequência da perda de pacotes e do aumento nos atrasos é a ocorrência de artefatos

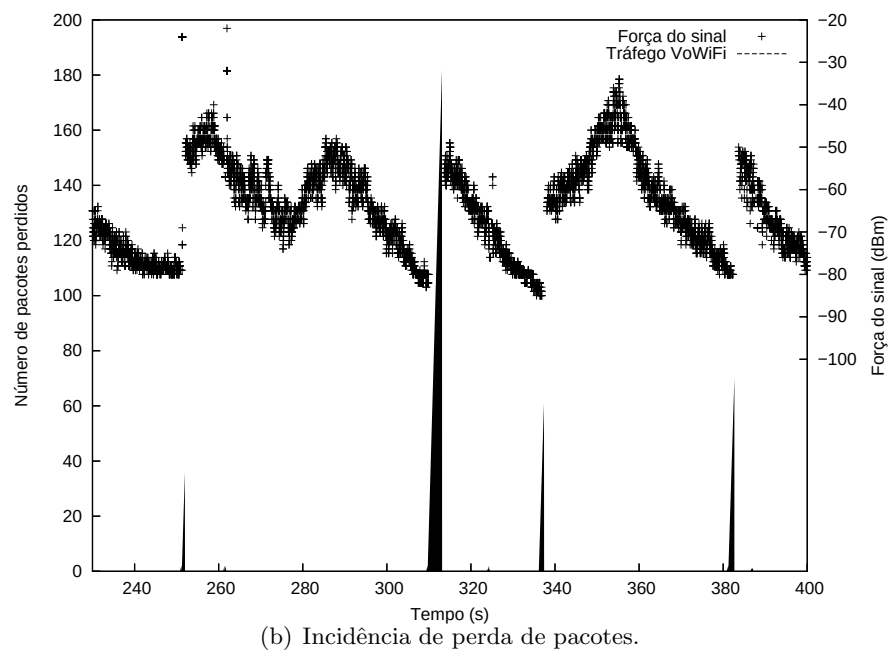
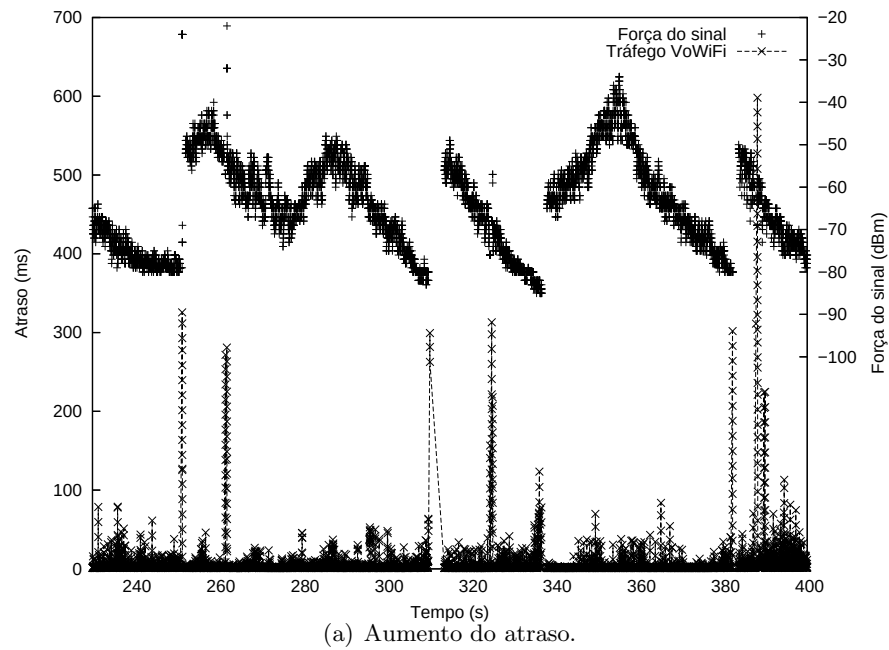


Figura 3.3: Implicações da operação de *handoffs* nos tráfegos de VoWiFi.

Tabela 3.1: Comportamento dos *handoffs*.

Ambiente	Número de <i>Handoffs</i>	Duração Média (<i>s</i>)	Desvio Padrão	Mín. (<i>s</i>)	Máx. (<i>s</i>)	Perda Média Pacotes	Mín.	Máx.
MSR Lab.	12	2,55	1,41	1,13	4,28	118,67	36	208
USP	10	8,4	4,12	4,25	16,19	386,1	187	799

audíveis nesses períodos de conversação.

Se estabelecermos que o início de um *handoff* acontece quando a perda de pacotes atinge 1% e a latência atinge 200ms, e que o fim desse *handoff* acontece quando a latência volta aos níveis normais, a duração dos períodos de *handoff* mostrados na Figura 3.3 é de 1.18, 3.75, 1.25 e 1.82s, respectivamente. Esses valores estão consistentes com os reportados na literatura [RS05].

A Tabela 3.1 resume o comportamento observado em dois traços de execução registrados enquanto as unidades móveis eram movidas e estavam, portanto, sujeitas a *handoffs*. Para cada traço, a tabela apresenta: 1) o ambiente onde o traço foi coletado; 2) o número de *handoffs* executados durante a coleta do traço; 3) o tempo médio de duração dos *handoffs* (segundo os critérios de início e fim definidos no parágrafo anterior); 4) desvio padrão da duração média dos *handoffs*; 5 e 6) o tempo mínimo e máximo de duração dos *handoffs*; 7) número médio de pacotes perdidos por *handoff*; 8 e 9) o número mínimo e máximo de pacotes perdidos por *handoff*.

Da análise dos traços podemos extrair três considerações importantes. A primeira era esperada: a duração dos *handoffs* varia em função dos ambientes onde foram executados. A segunda: todos os *handoffs* em nossos experimentos gastaram pelo menos um segundo para serem completados; esse tempo é uma forte restrição para aplicações de VoIP, pois durante esse período não há comunicação entre os pares de uma chamada de voz. Por fim, a terceira: a duração dos *handoffs* é muito variável, podendo variar significativamente de uma execução para outra. Por exemplo, no traço colhido na rede IEEE 802.11 da Universidade de São Paulo, os tempos para execução dos *handoffs* variaram de 4.25 a 16.19s.

Esse comportamento — indesejável — dos *handoffs* deve-se às recomendações feitas no padrão IEEE 802.11f, ou *Inter Access-Point Protocol* (IAPP), como é mais conhecido. O protocolo IAPP define que a unidade móvel deve conduzir o processo de migração sem o auxílio dos pontos de

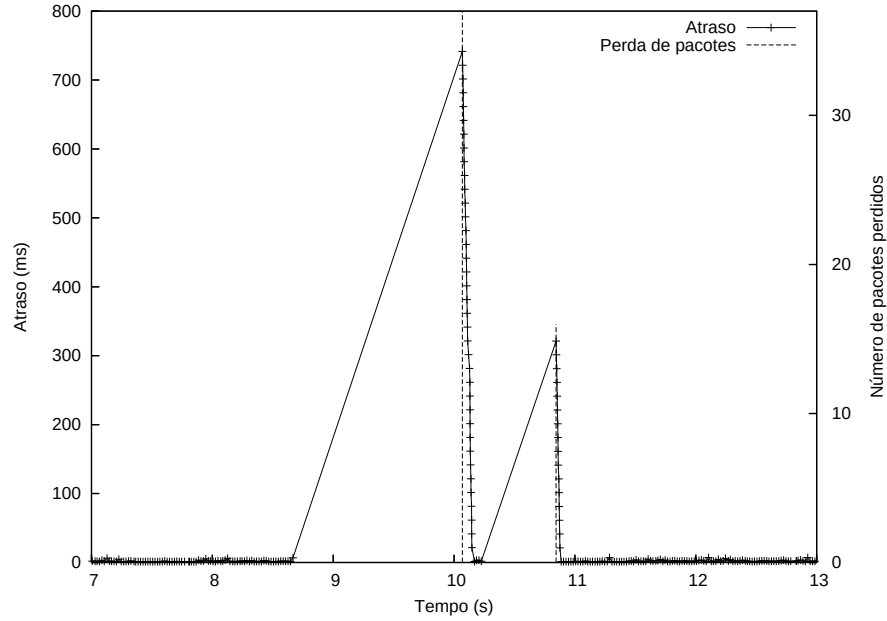


Figura 3.4: Impacto da atividade de busca em VoWiFi.

acesso. Além disso, com o objetivo de alcançar maior simplicidade e segurança, IAPP define que uma unidade móvel deve estar associada apenas a um — e apenas um — ponto de acesso. Em função dessas restrições, a migração consiste em um processo seqüencial composto de quatro fases: busca (*scanning*), autenticação, associação e reassociação. Dentre esses passos, a busca é a fase mais demorada, podendo consumir cerca de 90% do tempo total de migração; ou seja, alguns segundos [VK04].

A fim de estudar os efeitos da fase de busca em tráfegos de VoWiFi, forçamos um *scanning* durante uma sessão de VoWiFi. A Figura 3.4 mostra o resultado do experimento. A requisição de busca por um ponto de acesso foi implementada usando-se a API de NDIS, mais precisamente a requisição `OID_802_11_BSSID_LIST_SCAN`. O experimento foi preparado de forma que o NIC podia associar-se apenas a um dos pontos de acesso, de tal modo que o NIC sempre se reassociava a um deles, sempre o mesmo, após a fase de busca. A figura mostra que mais do que um segundo é gasto com o *scanning*.

Existem vários esforços com o objetivo de otimizar o processo de migração em redes IEEE 802.11 [VK04], principalmente na fase de busca [RS05]. O processo de *handoff* é reconhecidamente um dos pontos fracos de IEEE 802.11 e por isso está sendo revisto na especificação IEEE 802.11r. Porém, os primeiros dispositivos compatíveis com o padrão IEEE 802.11r devem levar alguns anos

para serem lançados e chegarem aos usuários.

3.3.2 Tráfego em rajadas

Além dos problemas relacionados à execução de *handoffs*, os traços de execução revelaram outro problema de VoWiFi: a alta incidência de tráfegos em rajada. A ocorrência de uma rajada caracteriza-se pelo fato de vários pacotes de voz ficarem bloqueados na rede por algumas centenas de milissegundos e então serem entregues praticamente ao mesmo tempo. A entrega de pacotes em rajada também pode ser acompanhada da perdas de pacotes, mas não necessariamente.

As rajadas não são exclusividade das redes sem fio, pois também podem ocorrer em redes Ethernet cabeadas; observamos a ocorrência de rajadas tanto em conexões Internet tradicionais, quanto em conexões Wi-Fi. No entanto, elas são mais comuns nas redes IEEE 802.11. A Figura 3.5 mostra um experimento no qual se pôde observar a incidência de rajadas durante uma sessão de VoWiFi: o transmissor está conectado ao ponto de acesso e o receptor é a unidade móvel; as rajadas podem ser observadas aos 608, 629, 639, 644, 646 e 658s. Durante as rajadas, um certo número de pacotes fica bloqueado na rede e então são todos entregues praticamente ao mesmo tempo; mas, é claro, com atraso. No traço apresentado na Figura 3.5, os atrasos vão de $200ms$ a $380ms$ (de 10 a 19 pacotes ficam bloqueados na rede). Pode-se observar também que o atraso decresce em aproximadamente $20ms$ para cada pacote recebido, isso ocorre porque os pacotes são transmitidos a cada $20ms$ e os pacotes atrasados são entregues praticamente ao mesmo tempo.

A ocorrência de rajadas foi consistentemente observada nas redes IEEE 802.11, mesmo usando-se diferentes ambientes e configurações; o que inclui o uso de diferentes NICs, pontos de acesso e ferramentas geradoras de tráfego. Pode-se, portanto, afirmar que as rajadas não ocorrem em função das peculiaridades de um ou de outro dispositivo, mas são uma característica dessas redes.

As rajadas podem ser originadas por diferentes fatores, tais como: *scannings*, interferências externas, interrupções de processamento etc. Além disso, pode-se afirmar que as rajadas não ocorrem devido a uma razão única, mas sim devido a um conjunto de fatores. As rajadas não podem ocorrer apenas devido à colisão de pacotes, pois, se um pacote colide e os esforços de retransmissão também falham, em atrasos da ordem de $100ms$, ou mais, o pacote seria dado como perdido; e não seria simplesmente atrasado, como acontece nas rajadas. Em outras palavras, uma

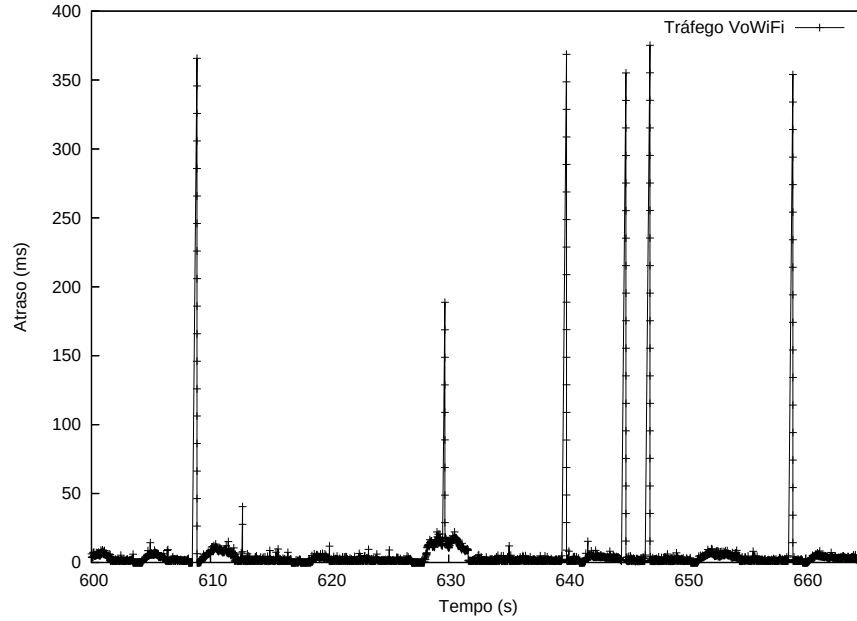


Figura 3.5: Exemplo do comportamento dos tráfegos em rajada.

perda não seria confundida com uma rajada. As rajadas também não ocorrem apenas em função da carga excessiva das redes, pois foram observadas rajadas mesmo em redes ociosas. E também não podem ser atribuídas apenas aos *scannings*, pois foram observadas em locais onde a força do sinal era máxima e, portanto, os *scannings* não deveriam acontecer.

Inicialmente, acreditávamos que as rajadas ocorressem devido à ação de interferências externas. Essa suposição nos levou a pesquisar e explorar o uso de duas ou mais conexões IEEE 802.11 [CBB04, BAPW04]. A idéia era reduzir o impacto das rajadas e dos *handoffs*, mantendo-se sempre duas conexões ativas em redes distintas [TXE⁺05]. Por redes distintas, entenda-se o uso de dois NICs diferentes (AIR-PCM340 e WUSB54GP) conectados a dois pontos de acesso diferentes (linksys WRT54GS e Netgear FWAG114), operando nos canais 1 e 11, respectivamente, ou seja, em canais diferentes e que não interferem entre si.

Ao usar duas conexões distintas, observamos, mesmo assim, a ocorrência de *rajadas correlacionadas*, isto é, rajadas que ocorrem exatamente ao mesmo tempo em ambas as conexões. A Figura 3.6 mostra a ocorrência de rajadas correlacionadas aproximadamente aos 143s e 185s. A ocorrência dessas rajadas não elucida a razão de sua existência, mas deixa claro que as razões não estão restritas às questões de rede: podem também estar associadas, por exemplo, à capacidade dos computadores atuais de lidar com aplicações de tempo real. Outra alternativa — porém me-

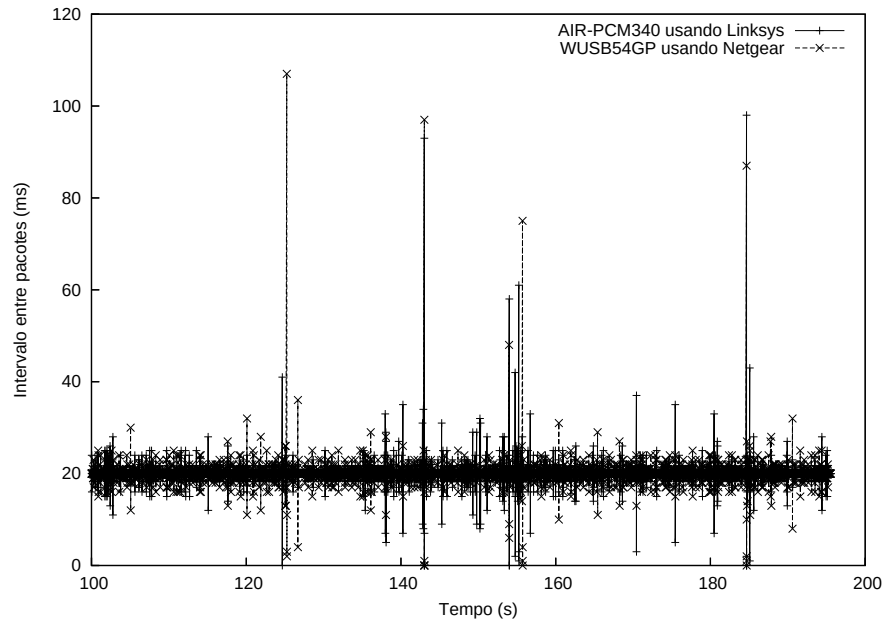


Figura 3.6: Correlação da ocorrência de rajadas em tráfegos distintos.

nos provável — é a existência de interferências externas capazes de bloquear todo o espectro de transmissão de IEEE 802.11.

A Tabela 3.2 apresenta um estudo mais extenso sobre as rajadas. Ali comparamos a execução de VoIP sobre a Internet tradicional e sobre IEEE 802.11. Os traços possuem de 2 a 30 horas de tráfego e os resultados são apresentados em nove colunas, sendo que cada coluna informa: 1) locais de coleta dos traços; 2) duração da coleta do traço, em segundos; 3) atraso médio, em milissegundos; 4) desvio padrão; 5) taxa de perda de pacotes (TPP), em porcentagem; 6) número de ocorrências de perdas de pacotes em rajada por segundo (PPR/s), sendo considerada perda em rajada a perda de 3 ou mais pacotes consecutivamente; 7) número de ocorrências de atrasos em rajada por segundo (AR/s), sendo considerado atraso em rajada atrasos de mais de 100ms e mais de 5 pacotes entregues quase ao mesmo tempo; e 8) tamanho da rajada (aumento médio do atraso em caso de rajada), em milissegundos.

Note que os traços IEEE 802.11 apresentados na Tabela 3.2 são formados por um trecho IEEE 802.11g (último salto) e pelo trecho restante, que é composto pela Internet tradicional; ou seja, o trecho Internet também pode afetar as características da conexão sem fio. Apesar disso, podemos fazer algumas observações interessantes: em geral, a qualidade das conexões IEEE 802.11 é pior do que da Internet sobre Ethernet; em média, a taxa de perda de pacotes das conexões IEEE 802.11 é

Tabela 3.2: Tráfego de VoIP.

Local	Duração (s)	Atraso (ms)	Desvio Padrão	TPP (%)	PPR/s	AR/s	Tamanho da Rajada (ms)
Internet							
Beijing ↔ Seattle (1)	21.998	108	114	0,02	0,000	0,000	260
Beijing ↔ Seattle (2)	11.455	148	148	0,07	0,032	0,001	237
Philadelphia ↔ Seattle	107.761	63	75	0,76	0,313	0,013	170
IEEE 802.11							
Beijing ↔ Seattle	28.592	216	220	1,54	0,426	0,005	249
Philadelphia ↔ Seattle	6.529	54	70	0,71	0,095	0,016	184
Amsterdam ↔ Seattle	10.426	93	93	1,10	0,506	0,002	154
New York City ↔ Seattle	49.567	110	203	0,53	0,025	0,064	308

0.97% versus 0.28% da Internet, a frequência de perdas em rajada é 0.26 versus 0.12 por segundo e a frequência de atrasos em rajada é 0.022 versus 0.005 por segundo.

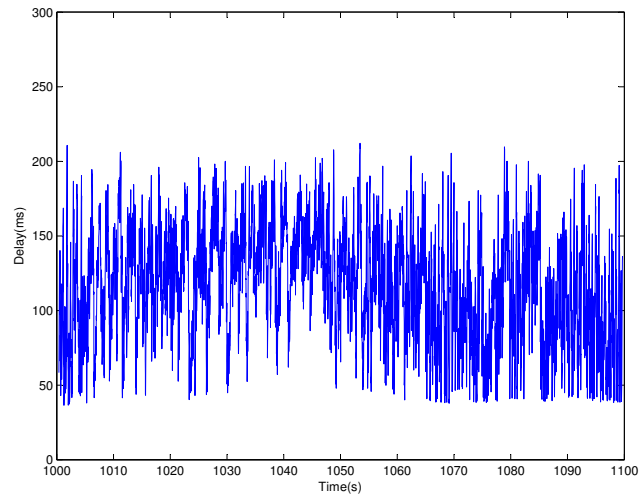
Mesmo quando as estatísticas da conexão Internet parecem similares às da conexão IEEE 802.11, os traços revelam diferentes características de comportamento entre as conexões. Por exemplo, a Figura 3.7 mostra duas partes de traços de VoIP (um usando apenas Internet e outro usando também Wi-Fi) referentes à conexão entre Philadelphia e Seattle; observe que a diferença de comportamento entre os tráfegos é bem maior do que revelam as estatísticas.

3.4 Problemas adicionais

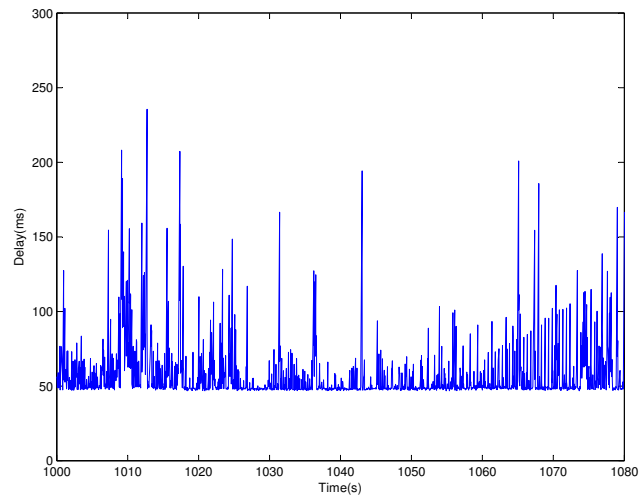
Além dos dois problemas principais, os *handoffs* e as rajadas de tráfego, também identificamos outros problemas que poderiam dificultar a implantação de VoWiFi.

3.4.1 Capacidade

Apesar de a capacidade nominal de vazão das redes IEEE 802.11 ser maior do que a vazão usada pelas aplicações de VoIP, na prática, um ponto de acesso mal pode acomodar uma dezena de chamadas de voz [MGFK04]. Na verdade, uma rede sem fio IEEE 802.11b consegue acomodar bem apenas 6 sessões de VoIP [GK03].



(a) Tráfego Internet.



(b) Tráfego Wi-Fi.

Figura 3.7: Comportamento de tráfegos de VoIP (Philadelphia $\leftrightarrow$ Seattle).

Essa diferença entre a capacidade nominal e a capacidade real deve-se principalmente às sobrecargas relacionadas ao uso de cabeçalhos e troca de mensagens de controle. Por exemplo, ao se usar RTP para transmitir 60 octetos de voz, transmitem-se também 40 octetos de cabeçalhos (sendo 12 octetos de RTP, 8 de UDP e 20 de IP). Considerando-se apenas essa sobrecarga, a utilização efetiva da rede pelos tráfegos de VoIP já se reduz para apenas 60%. Quando são considerados outros fatores, tais como os cabeçalhos usados pela camada MAC e PHY de IEEE 802.11 e os tempos gastos com sincronização e trocas de mensagens, essa eficiência se reduz ainda mais, podendo chegar a menos de 3% [WLL05]. Na prática, uma sessão de VoIP a 64 Kbps toma cerca de 900 Kbps da vazão total efetiva do sistema [GK03], que é de 6 a 7 Mbps em redes IEEE 802.11b [dCK03].

3.4.2 Justiça

Equilíbrio, justiça, ou ainda *fairness*, é um conceito difícil de ser obtido em redes IEEE 802.11. Isso porque, primeiro, os pontos de acesso têm capacidade limitada e essa capacidade é compartilhada entre as estações móveis; conseqüentemente, unidades móveis “gulosas” podem afetar negativamente o tráfego de outras unidades. Segundo, a camada MAC de IEEE 802.11 foi projetada com o objetivo de prover a mesma probabilidade de acesso ao meio a todas as unidades móveis, independentemente do tamanho dos pacotes, da qualidade do sinal ou da taxa de transmissão. Desse modo, uma unidade móvel transmitindo a 1 Mbps pode afetar negativamente outras unidades que transmitem a 11 Mbps [HRBSD03]. Por fim, os dispositivos IEEE 802.11 aplicam estratégias locais para otimização de seus fluxos, sem considerar o impacto que essas estratégias possam ter no desempenho do sistema como um todo [TG05].

No centro do problema de justiça, está um mecanismo interno de IEEE 802.11: o Controle Automático da Taxa de Transmissão (*Automatic Rate Control* ou ARC). Ele é responsável por escolher a modulação adequada às condições do canal, com o objetivo de explorar o compromisso entre a taxa de erros e a capacidade de transmissão [HTL⁺05]. Seu princípio de atuação é: quanto pior a situação do canal, mais robusta deve ser a modulação utilizada (e menor a taxa de transmissão); quanto melhor o canal, menos robusta pode ser a modulação utilizada (e maior a taxa de transmissão). O padrão IEEE 802.11g define o uso de até 14 diferentes modulações, que resultam nas seguintes taxas de transmissão: 54, 48, 36, 33, 24, 22, 18, 12, 11, 9, 6, 5.5, 2 e 1 Mbps.

Porém, como veremos no Capítulo 6, o padrão deixa em aberto o algoritmo e os critérios usados pelo mecanismo de ARC para escolher dentre essas possíveis modulações. Como consequência dessa não-definição, os mecanismos de ARC são implementados pelos fabricantes de interfaces de rede IEEE 802.11 de maneira proprietária e “fechada”; isso não seria um problema não fosse o impacto direto do mecanismo de ARC na capacidade da rede e na distribuição justa dessa capacidade [TG05, Gup05]. Em poucas palavras, o mecanismo de ARC faz com que as redes IEEE 802.11 sejam menos previsíveis e o compartilhamento da capacidade de vazão menos justa.

3.4.3 Interferências

As redes IEEE 802.11 são susceptíveis a interferências externas. As redes IEEE 802.11b e 802.11g são particularmente susceptíveis porque utilizam a mesma frequência (2.4 GHz) de alguns outros dispositivos, tais como fornos de microondas, telefones sem fio (*cordless phones*) e dispositivos Bluetooth. Além disso, a propagação do sinal é afetada por fatores como a disposição física de objetos, a orientação das antenas e, inclusive, as condições climáticas. A força do sinal pode variar significativamente, mesmo em períodos curtos de tempo (veja a Figura 3.3), devido à mobilidade e ao efeito *multi-path* [Sch00]. Em suma, são tantos os fatores capazes de afetar e interferir negativamente em uma conexão IEEE 802.11, que é muito difícil dimensionar essas interferências.

3.4.4 Diversidade de implementação

Definimos como diversidade de implementação a diferença de comportamento que pode ser observada nas transmissões quando são utilizados diferentes dispositivos IEEE 802.11. Chamamos a atenção para essas diferenças em trabalhos anteriores [dCK03] e para o fato de que os trabalhos baseados em simulações não consideravam esses problemas. Nossos últimos experimentos com VoWiFi ajudaram a reforçar essas questões, pois em tráfegos de tempo real, como os de VoIP, a questão de diversidade de implementação deixa de ser um problema sutil, como o é em determinados tráfegos TCP. De maneira geral, o uso de diferentes NICs, computadores e configurações implicam características diferentes de transmissão, ou seja, implicam transmissões sujeitas a diferentes taxas de perdas de pacotes e valores médios de atraso e *jitter*. Além disso, outro aspecto da diversidade de implementação pode-se notar na Figura 3.3(a).

Ali, dentre os valores de força de sinal, existem valores que nunca são apresentados, como se nunca ocorressem, por exemplo -61 dBm. Essa pequena “anomalia” acontece porque o valor da força do sinal é mapeado em uma escala de -10 a -200 dBm, como define NDIS, mesmo que para o NIC esses valores sejam, por exemplo, de 1 a 50. Além disso, como veremos no Capítulo 5, os sistemas Windows e Linux usam diferentes formas de mapeamento para a força do sinal. Mais informações sobre como — e quão diferentemente — a força do sinal é relatada pelos NICs para a camada de aplicação podem ser encontradas em [Wil02].

Uma vez que IEEE 802.11 não define detalhadamente como alguns de seus mecanismos devem atuar, estes são implementados de maneira proprietária e fechada. Sendo assim, fica restrita nossa capacidade de determinar com precisão seus comportamentos. Conseqüentemente, é impossível prever ou calcular o comportamento e as sobrecargas das redes IEEE 802.11 [GK03, HTL⁺05]. A diversidade de implementação, em outras palavras, dificulta a interpretação das métricas de adaptação e da real qualidade da conexão.

3.4.5 Métricas de qualidade de conexão

Existem várias métricas a partir das quais podemos inferir a qualidade de uma conexão, como por exemplo a taxa de perda de pacotes, a força do sinal, o *jitter* etc. Entretanto, nas redes IEEE 802.11, as métricas usuais não são suficientes para definir a qualidade de uma conexão; ou, no mínimo, são métricas menos confiáveis.

A taxa de perda de pacotes, por exemplo, é afetada pelo mecanismo de retransmissão automática de IEEE 802.11. Recapitulando, o mecanismo de retransmissão automática, antes de notificar a perda de um pacote, tenta retransmitir o pacote até sete vezes [Gas02]. Dessa forma, a camada de aplicação percebe a perda de pacotes de maneira imprecisa e, mais do que isso, pode percebê-la apenas quando a taxa dessas perdas for excessivamente alta, de forma tal que pouco pode ser feito para recuperar-se dos erros ocorridos.

A taxa de perda de quadros de transmissão (*Frame Error Rate* ou FER), ou seja, o número de quadros cuja transmissão não foi confirmada através de uma mensagem ACK, é uma métrica melhor do que a taxa de perda de pacotes. Entretanto, nem todo NIC fornece essa informação. A implementação da requisição `OID_802_11_STATISTICS`, a requisição de NDIS que retorna FER,

é apenas recomendada. Ou seja, os NICs não são obrigados a implementá-la. Em nossos experimentos, testamos a requisição em quatro NICs diferentes, mas ela funcionou em apenas um deles (WUSB54GP).

As métricas baseadas na qualidade de sinal, tais como força do sinal e SNR, são afetadas negativamente pela alta variabilidade das conexões IEEE 802.11. Além disso, o ruído, uma das informações necessárias para o cálculo de SNR, não está disponível na API de NDIS.

Em suma, não existem métricas apropriadas para determinar a qualidade da conexão. As métricas existentes são instáveis e podem não refletir a real qualidade da conexão. Além disso, as métricas existentes são locais, isto é, não há uma métrica de adaptação capaz de prover o nível de utilização da rede como um todo; ou quanto está sendo utilizado da capacidade de um ponto de acesso, ou quanto ainda poderia ser utilizado. Se fossem oferecidas às unidades móveis métricas de adaptação capazes de indicar o desempenho global da rede, essas unidades poderiam ser mais proativas com relação à adaptação e à solução dos problemas de conexão.

3.5 Mais algumas considerações sobre VoWiFi

É importante notar que os problemas apresentados nas seções anteriores não podem ser considerados separadamente, pois eles podem acontecer simultaneamente, por vezes tornando-se ainda piores. Por exemplo, as rajadas podem ficar mais intensas com o aumento do tráfego e da mobilidade e os *scannings* são mais freqüentes em ambientes com maior incidência de interferências.

Sob a ótica do desenvolvimento de sistemas, os problemas de VoWiFi representam, grosso modo, as razões porque é tão complexo desenvolver aplicações adaptativas para redes IEEE 802.11. A imprevisibilidade e vulnerabilidade a interferências das redes IEEE 802.11 fazem com que o projeto de estratégias adaptativas para essas redes seja uma tarefa “artesanal”, única para cada circunstância. Além disso, é muito difícil, se não impossível, determinar a causa dos problemas e, conseqüentemente, prover as reações adaptativas apropriadas que acabem não afetando negativamente outros usuários do sistema.

Do ponto de vista da pesquisa, a diversidade de comportamento entre implementações e a imprevisibilidade de IEEE 802.11 dificultam a proposta e a validação de algoritmos e modelos. Existem tantas variáveis — e variância(s) — que se torna excessivamente trabalhoso produzir

experimentos representativos, especialmente experimentos relacionados à mobilidade.

3.6 Trabalhos relacionados

Os principais trabalhos relacionados aos problemas de VoWiFi são especificações em andamento². Por exemplo, a especificação IEEE 802.11r otimiza o processo de *handoff*, 802.11n aumenta a taxa de transmissão, 802.11k especifica melhores métricas para se definir a qualidade da conexão e 802.11w reduz as interferências.

A melhoria do problema de tráfego em rajadas, porém, não deve ser tratada por uma, mas sim por várias especificações conjuntamente: IEEE 802.11r deve reduzir os efeitos dos *scannings*, 802.11e deve permitir o controle de QoS e permitir a gerência de filas. As especificações IEEE 802.11n e 802.11w também devem contribuir indiretamente para a redução das rajadas.

Outra iniciativa que deve ter impacto em VoWiFi foi a recente inclusão, em abril de 2005, do comando `OID_802_11_MEDIA_STREAM_MODE` na API de NDIS, o que deve permitir que as aplicações entrem em modo protegido de operação durante transmissões de tráfego de tempo real, evitando assim que operações e mecanismos da rede sem fio, como *scannings* e PSM, interfiram nas transmissões. A inclusão desse comando na API de NDIS, antes de tudo, demonstra a sensibilização da comunidade para a importância do tráfego multimídia para as redes IEEE 802.11.

No decorrer deste trabalho, referiremos propostas para a efetiva melhora do desempenho de VoWiFi, tais como estratégias mais inteligentes de migração [RS05, VK04], algoritmos adaptativos [LFG03], melhorias na camada MAC de IEEE 802.11 [TG05, Gup05], o uso de multi-conexões [BAPW04, CBB04] e de multi-caminhos [TXE⁺05]. Esses trabalhos, porém, ainda estão em desenvolvimento e provavelmente levarão anos para chegarem aos usuários. Isso leva a pergunta: O que pode ser feito agora?

²Veja <http://grouper.ieee.org/groups/802/11/802.11-Timelines.htm>

3.7 Cenários viáveis para a utilização de VoWiFi

Nas seções anteriores, apresentamos as razões porque a implementação de VoWiFi é um desafio. Além disso, também apresentamos as extensões que podem ser aplicadas para viabilizar VoWiFi — mesmo que algumas dessas extensões só venham a estar disponíveis daqui a alguns anos. Nesta seção, contudo, nosso objetivo é discutir em que cenários VoWiFi é viável atualmente.

Nossos experimentos mostraram que VoWiFi provê qualidade satisfatória quando as seguintes restrições são observadas:

- o usuário fica conectado a um único ponto de acesso durante a chamada. Isso requer que o usuário tenha pouca mobilidade; movimentar-se em regiões cobertas por vários pontos de acesso levará à realização de migrações, que afetam fortemente a qualidade da ligação;
- a força do sinal a qual o usuário está conectado é alta. Apesar de reduzir a necessidade de *scannings* e de retransmissões, isso limita a área de utilização de VoWiFi para as proximidades do ponto de acesso. Em nossos experimentos, as taxas de erros em tráfegos VoWiFi foram consistentemente baixas para qualidades de sinal superiores a -70 dBm, mas esse valor pode variar em função, entre outros fatores, do ambiente e da carga da rede;
- um número reduzido de chamadas é feito em cada ponto de acesso. Para redes, IEEE 802.11b, por exemplo, o limite é de aproximadamente 6 conexões. Aproximar-se desse limite pode reduzir a qualidade das chamadas;
- se possível, não se deve fazer *scannings* durante as chamadas, o que pode ser feito, em alguns NICs, chamando-se o comando `OID_802_11_STATISTICS` das APIs de NDIS.

Acreditamos que existem basicamente dois cenários para o uso de VoIP. No primeiro cenário, a chamada é roteada através do sistema de telefonia padrão e VoWiFi é usado apenas para agregar mobilidade; isto seria equivalente a telefones sem fio (*cordless phones*). Este cenário poderá ser comum em ambientes corporativos e é menos complexo, pois o atraso inserido pelo uso de IEEE 802.11 deve ser insignificante.

No segundo cenário, a chamada é toda roteada através da Internet. Apesar de desafiador, esse cenário é mais próximo das necessidades dos usuários finais e, por isso, é nossa principal preocupação. Nesse caso, o usuário tem um “número” de telefone VoIP e pode migrar entre redes durante a execução de uma chamada. Os atrasos e perdas de pacotes, nesse cenário, somar-se-ão aos da Internet. Esta é uma aplicação muito interessante e acreditamos que, colocadas as restrições acima, VoWiFi é uma alternativa de boa qualidade, barata e flexível para os sistemas de comunicação de voz em tempo real. Estamos também investigando mecanismos capazes de fazer a transição entre o modelo tradicional de comunicação bidirecional para modelos *push-to-talk* (P2T), mais resilientes a perdas de pacotes e atrasos. A transição para o modelo P2T poderia ser uma boa alternativa quando a qualidade da conexão viesse a ficar inaceitável para a comunicação bidirecional.

Observamos que, respeitadas as restrições definidas acima, a qualidade de VoWiFi é ligeiramente inferior à de VoIP. Mesmo assim, achamos que a flexibilidade de uma conexão sem fios é atrativa para muitos usuários, fazendo com que VoWiFi seja uma evolução natural de VoIP, assim como o foram outras tecnologias de telefonia sem fio. Além disso, uma das razões do sucesso da telefonia celular foram os serviços agregados que se tornaram possíveis pela digitalização. De maneira análoga, novas aplicações poderiam se tornar viáveis e VoWiFi, desse modo, poderia oferecer serviços de valor agregado ao usuário.

3.8 Conclusões

Este capítulo mostrou os dois principais problemas para a ampla implantação de VoWiFi: os *handoffs* excessivamente longos e instáveis e a alta incidência de rajadas. Além disso, apontamos outros problemas que podem dificultar a utilização de VoWiFi. Apesar desses problemas, os experimentos realizados mostraram que VoWiFi oferece boa qualidade quando a capacidade de vazão da rede não é excedida, a qualidade do sinal é boa e a mobilidade é moderada.

Trabalhamos também em algoritmos para reduzir o problema de diversidade de implementação; esses algoritmos assemelham-se a processos de calibração das métricas de adaptação e serão mencionados nos Capítulos 5 e 6.

No horizonte de trabalhos a realizar, gostaríamos de usar os mecanismos de sincronização pre-

sentos em IEEE 802.11, *Time Synchronization Function* (TSF), para aproximadamente sincronizar o tempo de transmissão entre as unidades móveis. A idéia é distribuir estatisticamente essas transmissões de VoIP, a fim de reduzir as colisões. Isso é possível porque em tráfegos de VoIP as transmissões se dão em intervalos regulares de tempo, por exemplo, a cada 20ms.

Por fim, resta observar que as tecnologias de suporte à VoWiFi estão amadurecendo; os seus problemas estão sendo resolvidos à medida que evoluem os produtos baseados em IEEE 802.11. Se hoje ainda há algo de artesanal em VoWiFi, em alguns anos VoWiFi deve estar plenamente disponível para os usuários.

Capítulo 4

Vídeo sobre redes IEEE 802.11

Movies are a fad. Audiences really want to see live actors on a stage.

Charlie Chaplin

Nos últimos anos, a utilização das aplicações de vídeo cresceu fortemente entre os usuários de computadores. Esse crescimento foi possível devido a alguns progressos tecnológicos, tais como o desenvolvimento de tecnologias de codificação de vídeo, a popularização das câmeras digitais, o barateamento dos meios de armazenamento e a maior disponibilidade de redes de alta velocidade para troca de conteúdos.

Os avanços tecnológicos reduziram gradativamente as limitações para a utilização das tecnologias de vídeo e, desse modo, levaram as aplicações de vídeo de nichos específicos para o alcance dos usuários finais. Nesse cenário, as aplicações de transmissão de fluxos contínuos de vídeo (*video streaming*) têm lugar de destaque.

As tecnologias de *streaming* permitem que os usuários assistam aos vídeos enquanto o conteúdo é transmitido; desse modo, o início da exibição é antecipado e reduz-se também a necessidade de armazenamento local. Outra razão para o uso de fluxos de vídeo é a redução da latência de comunicação, o que é fundamental para aplicações de comunicação em tempo real, tais como as aplicações de videoconferência.

Desse modo, devido à importância das aplicações de vídeo, este capítulo apresenta a utilização de fluxos de vídeo sobre redes IEEE 802.11. Para isso, descreve a arquitetura usual dessas

Coordenação	Produção de conteúdo	SS
	Codificação	
	Armazenamento	
	Transmissão	
	Distribuição	AP
	Recepção	MU
	Decodificação	
	Exibição	

Figura 4.1: Etapas da transmissão de multimídia.

aplicações, apresenta o padrão de codificação de vídeo atualmente considerado o padrão mais apropriado para implementação de fluxos de vídeo — o padrão H.264/AVC — e aponta os principais problemas para distribuição desses fluxos sobre redes Wi-Fi. Por fim, elencamos algumas características das redes IEEE 802.11 que, se revisadas, poderiam melhorar a distribuição de vídeo nessas redes.

4.1 Arquitetura das aplicações de fluxos de vídeo

A transmissão de vídeo sobre redes se dá através de etapas bem definidas, que vão da produção de conteúdo à sua exibição. No contexto desta pesquisa, que envolve redes IEEE 802.11 infra-estruturadas, as fases de produção, codificação, armazenamento e transmissão do vídeo são executadas em um servidor de vídeo (*Streaming Server* ou SS). A distribuição ocorre com o auxílio de um ponto de acesso (*Access Point* ou AP). As fases finais — recepção, decodificação e exibição — ocorrem na unidade móvel (*Mobile Unit* ou MU). Além disso, também se faz necessária a coordenação entre essas atividades. A Figura 4.1 ilustra o processo de transmissão de vídeo; vejamos em que consiste cada uma dessas etapas:

1. **Produção do conteúdo.** Consiste na geração de sinais digitais com o uso de dispositivos captadores, como câmeras e microfones, ou sintetizadores.

2. **Codificação.** Transformação do conteúdo bruto para um formato mais conveniente. Em geral, a transformação envolve compactação espacial e temporal. A compressão espacial minimiza a quantidade de dados necessários para a representação de uma cena ou quadro, como faz o padrão JPEG de compactação de imagens. Por sua vez, a compressão temporal elimina a redundância entre quadros. Os padrões de codificação de vídeo, pelo menos os mais modernos [HV02b, Mov03, Joi03], permitem definir parâmetros de quantização (*Quantization Parameters*), de tal modo que o tamanho e a qualidade dos vídeos codificados sejam adequados à capacidade de recepção e de exibição do cliente.
3. **Armazenamento.** Nesta fase, o conteúdo é armazenado em mídias não voláteis, tais como discos rígidos ou discos ópticos. Algumas aplicações dispensam esta fase; exemplo disso são as aplicações de videoconferência. Outras aplicações, devido à variedade de clientes, armazenam um mesmo vídeo em diferentes formatos de codificação e qualidade.
4. **Transmissão.** Consiste em enviar conteúdo de vídeo de um ponto a outro da rede, o que pode ser feito por: *unicast*, *multicast* ou *broadcast*. Além disso, a transmissão pode usar diferentes meios e protocolos de transporte. A forma mais comum de transmissão de fluxos de vídeo sobre a Internet é realizada usando-se o conjunto de protocolos RTP/UDP/IP. A escolha de UDP, um protocolo não confiável, deve-se principalmente a dois fatores. Primeiro, UDP, por não fazer retransmissões, pois não implementa recuperação de erros, é mais leve do que os protocolos confiáveis e, sendo assim, alcança maiores taxas de transmissão, o que é importante para aplicações de vídeo. Segundo, em fluxos contínuos de vídeo, pequenas taxas de perdas de dados são aceitáveis e não necessariamente comprometem a qualidade de exibição do vídeo.
5. **Distribuição.** Nesta camada, onde residem elementos de rede, *proxies* e *middleware*, pode-se implementar mecanismos adaptativos para garantir, transparentemente, os requisitos de tempo real dos fluxos de vídeo [BCD⁺97] e da mobilidade [ACKL98]. O ponto de acesso, por exemplo, é um elemento de rede que encapsula mecanismos adaptativos e pode ser configurado e otimizado para tráfegos específicos. A Seção 2.4 apresentou alguns parâmetros configuráveis de IEEE 802.11. A Seção 4.5, por sua vez, apresentará mais alguns mecanismos que podem

melhorar a distribuição de fluxos de vídeo se implementados nos pontos de acesso.

6. **Recepção.** Nesta fase, a unidade móvel recebe os pacotes RTP, verificando seu conteúdo, sua ordem de chegada e prioridade de exibição. É nesta etapa que as perdas e atrasos de pacotes podem ser observados. Pequenas taxas de perda de pacotes são toleradas pelas aplicações de vídeo; no entanto, taxas altas podem inviabilizar a exibição, como veremos na Seção 4.4.
7. **Decodificação.** Consiste na transformação do formato recebido para o formato a ser exibido.
8. **Exibição.** Reprodução da mídia recebida nos respectivos dispositivos. Pode ser necessário sincronizar as trilhas de transmissão, tais como as trilhas de áudio e vídeo.
9. **Coordenação.** De maneira ortogonal às fases anteriores, são necessárias atividades de coordenação que sincronizem todas as etapas, levando sempre em consideração os requisitos de tempo real das aplicações multimídia.

Dadas as fases definidas acima, este trabalho concentra-se nos problemas de distribuição; mais especificamente, este capítulo concentra-se em como a distribuição de vídeo pode ser otimizada através da implementação de pontos de acesso inteligentes. Antes disso, porém, vejamos: os requisitos das aplicações de fluxos de vídeo (Seção 4.2), os principais problemas relacionados ao uso de fluxos de vídeo sobre Wi-Fi (Seção 4.3) e o padrão de codificação de vídeo H.264/AVC (Seção 4.4) — atualmente considerado o mais apropriado para implementação de fluxos de vídeo.

4.2 Requisitos dos fluxos de vídeo

Os fluxos de vídeo podem ser conversacionais ou não conversacionais. Para os fluxos conversacionais, tais como os utilizados em videoconferências, o principal requisito é a latência, que não deve superar 250 ms [SHW03].

Nas aplicações de vídeo não conversacionais, também classificadas como aplicações de Vídeo sob Demanda (*Video on Demand* ou VoD), que atualmente representam a maior fatia de utilização de tecnologias de vídeo digital, o principal requisito é a vazão. Alguns exemplos de sistemas e serviços de vídeo não conversacional são: TV Digital (não interativa), *Google Video* (<http://video.google.com>), *YouTube* (<http://www.youtube.com>), *iTunes* (<http://www.apple.com/>

itunes) etc. Essas aplicações permitem aos usuários escolher e receber vídeos através da Internet, seja através de *downloads* ou de fluxos contínuos — que permitem iniciar a exibição de um vídeo durante a sua transmissão.

Os principais requisitos para as aplicações de fluxos contínuos são, nesta ordem: a vazão, a taxa de perda de pacotes e o *jitter*.

4.2.1 Vazão

A banda passante necessária para trafegar um fluxo de vídeo pode variar — sobretudo — em função do vídeo transmitido. As tecnologias de vídeo atuais permitem limitar a vazão necessária durante a fase de codificação; por exemplo, um mesmo vídeo pode ser codificado para 128 Kbps ou 56 Kbps, porém essa redução na taxa de transmissão se dá em detrimento da qualidade visual.

Desse modo, para as redes IEEE 802.11, é necessário verificar se a taxa de transmissão de um vídeo não supera a capacidade de vazão disponível na rede; se superar, certamente haverá problemas na exibição do vídeo. Isso não é trivial de verificar por duas razões: primeiro, como veremos no Capítulo 6, no padrão não existem mecanismos que permitam inferir a capacidade disponível na rede [dCK06b] e, segundo, a carga da rede pode variar substancialmente em curtos intervalos de tempo. Deve-se observar também que as taxas nominais das redes IEEE 802.11 não são alcançáveis; por exemplo, uma rede IEEE 802.11b, com taxa nominal de 11 Mbps, alcança efetivamente apenas cerca de 6 Mbps [dCK03].

4.2.2 Taxa de perda de pacotes

Baixas taxas de perdas de pacotes podem não comprometer a exibição de um vídeo, mas altas taxas de erro simplesmente tornam a exibição inviável. Nas tecnologias de vídeo analógicas, as taxas de erros atuam gradativamente na qualidade de exibição, gerando, por exemplo, chuviscos; porém, nas tecnologias digitais a atuação não é gradual; quando sujeitas a pequenas taxas de erros, pode-se verificar a presença de artefatos, mas, após certo nível de erros, a exibição se torna inviável, restando congelar a imagem ou exibir mensagens de erro.

Os padrões mais avançados de vídeo digital possuem um amplo conjunto de ferramentas para a recuperação de erros e aumento da robustez da transmissão [LSE03, WHV⁺02]. Desse modo,

é difícil determinar quais são as taxas de perdas de pacotes aceitáveis para fluxos de vídeo, pois para isso é preciso considerar como o vídeo em questão foi codificado e quais ferramentas estão sendo utilizadas. Por exemplo, os fluxos de vídeo são normalmente gerados seguindo um padrão de codificação, que pode conter quadros I, P e B [Mov03]. Fluxos contendo mais quadros I atingem menores taxas de compressão, mas são mais robustos em caso de perda de pacotes. Por outro lado, os fluxos contendo poucos quadros I são mais eficientes em termos de compressão, mas os erros referentes à perda de um pacote podem se propagar por vários quadros de exibição. Além disso, sabe-se que as métricas utilizadas para avaliar o impacto da perda de um pacote não estão diretamente relacionadas com a percepção humana [MCP03], mas sim ao volume de ruídos gerados na imagem. Em outras palavras, não existe um mapeamento preciso entre a taxa de perda de pacotes e o quanto essas perdas afetam a percepção humana.

Fluxos de vídeo H.264 podem comportar-se aceitavelmente, mesmo sujeitos a taxas de perdas de pacotes entre 3 e 20%, dependendo da forma de codificação e das ferramentas para recuperação de erros utilizadas [CMM06, CM03]. Por exemplo, vídeos codificados para mídias confiáveis, tais como DVDs, comportam-se aceitavelmente mesmo sujeitos a 3% de perdas de pacotes. Por sua vez, codificações para redes sem fio, que incluem dados redundantes e mecanismos para recuperação de erros (FEC), são mais robustas e podem suportar até 20% de perdas [CMM06].

4.2.3 *Jitter*

O *jitter* afeta as políticas de gerenciamento de *buffers*: quanto maior o *jitter*, maior o *buffer*; quanto menor o *jitter*, menor o *buffer* [LFG03]. Observe que a latência (principal requisito em aplicações conversacionais) não é um limitante para fluxos não conversacionais, pois nessas aplicações valores altos de latência implicam apenas alguns instantes a mais para o início da exibição de um vídeo.

4.3 Problemas de fluxos de vídeo sobre Wi-Fi

Os principais problemas encontrados para a transmissão de fluxos de vídeo sobre redes IEEE 802.11 ocorrem quando se extrapola a sua capacidade de transmissão. Porém, além desses problemas, também podemos citar: o baixo isolamento dessas redes e, na presença de falhas, seu comportamento “liga/desliga”.

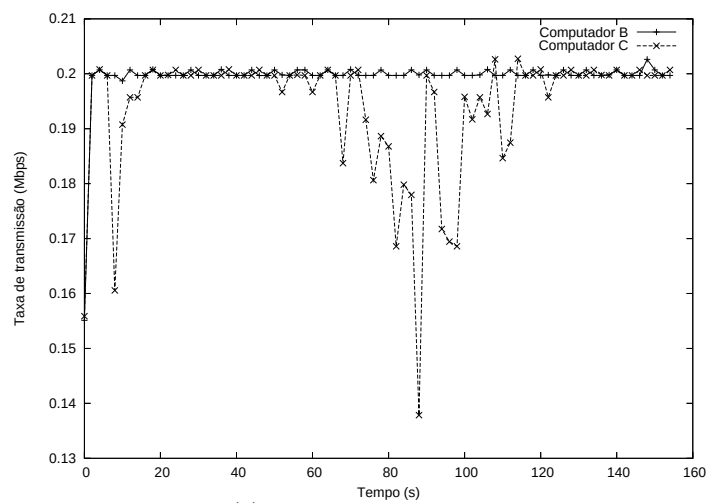
4.3.1 Baixo isolamento

As redes IEEE 802.11 oferecem baixo isolamento entre transmissões, isto é, a transmissão de um fluxo para uma unidade móvel pode afetar negativamente as transmissões para outras unidades. Isto ocorre por várias razões, duas das principais razões são: o compartilhamento da frequência de rádio entre as unidades móveis e os próprios mecanismos internos de IEEE 802.11, tais como ARQ e ARC. Vejamos um experimento que ilustra essa questão.

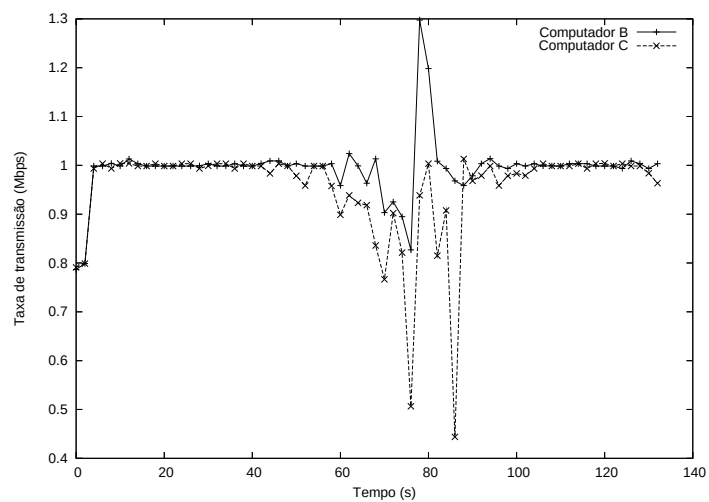
A Figura 4.2 apresenta três experimentos nos quais fluxos UDP, a diferentes taxas de transmissão, são concorrentemente enviados de um computador fixo **A** para unidades móveis **B** e **C** sobre uma rede IEEE 802.11b. No início da execução dos experimentos, os computadores **B** e **C**, ambos exatamente com as mesmas configurações, estão localizados em uma área próxima ao ponto de acesso, ou seja, estão bem cobertos pela rede sem fio. Então, o computador **C** é deslocado para uma região afastada, ou seja, mal coberta pelo ponto de acesso. Observe que existe forte influência entre os fluxos, mesmo tendo sido o computador **B** mantido fixo em uma área bem coberta pelo ponto de acesso. Note também que essa influência é maior, quanto maiores forem as taxas de transmissão. No experimento apresentado na Figura 4.2(c), para maior generalidade, além de mover o computador **C**, posteriormente também foi movido o computador **B**.

4.3.2 Comportamento liga/desliga

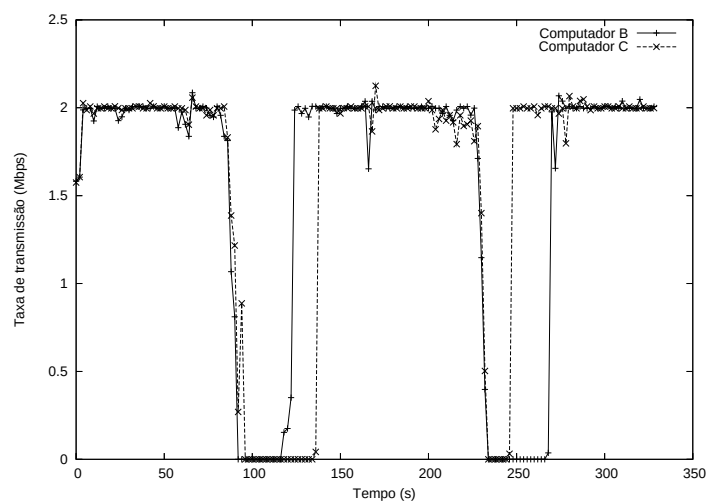
Outro problema para fluxos de vídeo sobre redes IEEE 802.11 é que essas redes assumem um comportamento “liga/desliga” em caso de falhas. Na Figura 4.2(c), por exemplo, nota-se que devido a uma ação externa — referente ao deslocamento da unidade móvel **C** — o computador **B** sofre fortes consequências; de fato, sofre um colapso total e abrupto de sua transmissão. O mesmo comportamento pôde ser observado em boa parte dos experimentos apresentados neste trabalho, mas foi especialmente notório nos experimentos referentes a fluxos de vídeo, em que as taxas de transmissão são mais elevadas. Devido ao comportamento liga/desliga, é bastante difícil propor estratégias adaptativas para transmissão de fluxos de vídeo sobre redes IEEE 802.11, pois as consequências negativas de alguns fatores, tais como mobilidade, *handoffs* ou interferências, são notadas pelas aplicações apenas quando a comunicação está completamente comprometida, restando apenas, para a maioria dos casos, aguardar o restabelecimento da conexão.



(a) Fluxos de 200 Kbps.



(b) Fluxos de 1 Mbps.



(c) Fluxos de 2 Mbps.

Figura 4.2: Interferência entre tráfegos mediante mobilidade.

4.4 Introdução ao padrão de codificação de vídeo H.264/AVC

O padrão de codificação de vídeo H.264/AVC possui várias denominações, dentre elas: ITU-T H.264, ISO/IEC 14496-10, MPEG-4 parte 10 e *Advanced Video Coding* (AVC). Por ser mais conhecido como H.264/AVC, vamos usar essa denominação.

H.264/AVC foi desenvolvido conjuntamente pela ITU-T *Video Coding Experts Group* (VCEG) e pela ISO/IEC *Moving Picture Experts Group* (MPEG), que, unidas, formaram a *Joint Video Team* (JVT). Atualmente, H.264/AVC é reconhecidamente o padrão de codificação de vídeo capaz de atingir as mais altas taxas de compressão. Para uma mesma qualidade relativa de percepção visual, a transmissão de um vídeo H.264/AVC usa 50% menos banda passante do que seu principal antecessor [EY05], o padrão MPEG-4 parte 2. Por exemplo, uma transmissão de vídeo H.264/AVC com qualidade de TV digital via satélite consome aproximadamente 1,5 Mbps de banda passante, enquanto o mesmo vídeo gastaria 3,5 Mbps usando-se MPEG-2 [LSE03]. Principalmente devido ao menor uso de banda passante, H.264/AVC é considerado o padrão de codificação de vídeo mais adequado para as aplicações de fluxo de vídeo; mas há também outras razões para isso, as quais veremos adiante.

A especificação de H.264/AVC foi concluída¹ em maio de 2003 e seu objetivo principal foi criar um padrão de codificação capaz de conciliar alta qualidade de vídeo e de compressão, sem, no entanto, aumentar demasiadamente a complexidade dos algoritmos, pois isso poderia inviabilizar sua implementação. Em 2005, H.264/AVC firmou-se como a principal tecnologia de codificação de vídeo, tendo sido adotada em uma série de produtos de destaque comercial, como por exemplo o sistema MediaFlo para distribuição de multimídia (<http://www.qualcomm.com/mediaflo/index.shtml>) e os tão populares iPods (<http://www.apple.com/ipod/ipod.html>).

Dentre as principais melhorias introduzidas por H.264/AVC estão: melhoria das técnicas de predição de movimento e dos métodos de codificação de entropia, uso de blocos menores de in-

¹Cabe mencionar que, a exemplo de MPEG, a especificação H.264/AVC define apenas a sintaxe de representação dos vídeos (os formatos de armazenamento e de transmissão) e a forma de decodificação, mas deixa em aberto muitas das técnicas de codificação, especificando apenas linhas mestras que devem ser obedecidas para garantir a compatibilidade entre codificadores e decodificadores. A adoção dessa estratégia promoveu ampla concorrência na pesquisa de técnicas de codificação, tendo sido esse um dos fatores que contribuíram para o desenvolvimento de algoritmos de codificação mais eficientes e de maior qualidade.

formação e a inclusão de um filtro para suavizar a transição entre os blocos. Além disso, em H.264/AVC, as relações de dependência entre os dados são muito mais flexíveis do que nos padrões anteriores. Por exemplo, em H.264/AVC, os quadros P podem usar mais de um quadro de referência anterior (em MPEG era apenas um quadro) e os quadros B também podem ser usados como referência (em MPEG apenas os quadros P e I podiam ser usados como referência). Essas melhorias, entre muitas outras [LSE03], fizeram de H.264/AVC um padrão de alta qualidade e suficientemente flexível para poder ser usado em uma vasta gama de aplicações e dispositivos, que vão de vídeo *emails* em aparelhos celulares a TVs de alta definição.

Cabe dizer, porém, que não existe um consenso sobre o uso de H.264/AVC em aplicações de baixa latência, tais como as aplicações de videoconferência e de vídeo telefone. Isso ocorre porque as inovações presentes em H.264/AVC aumentaram o custo computacional dos algoritmos de codificação em cerca de cinco vezes e de decodificação em duas vezes [EY05], impactando negativamente a latência. Atualmente, as codificações de vídeo usadas nessas aplicações são predominantemente baseadas em MPEG-4 e H.324. Contudo, o amadurecimento das tecnologias para codificação e decodificação de H.264/AVC, como por exemplo o desenvolvimento de circuitos integrados dedicados a esse fim, deve reduzir a latência relacionada ao uso desse padrão. Destarte, H.264/AVC deverá ser cada vez mais utilizado também nas aplicações de baixa latência.

Uma das principais inovações introduzidas no projeto de H.264/AVC foi a sua divisão conceitual em duas camadas, cada uma delas com funções bem específicas: a camada de codificação de vídeo (*Video Coding Layer* ou VCL), que encapsula as técnicas de compressão, e a camada de abstração de rede (*Network Abstraction Layer* ou NAL), que encapsula as técnicas para transmissão de vídeo. A camada NAL foi uma das principais inovações introduzidas em H.264/AVC, fazendo dele um padrão adequado para a transmissão de vídeo sobre redes, como veremos a seguir.

4.4.1 Camada de abstração de rede

A camada de abstração de rede (NAL) foi desenvolvida para permitir a adaptação do conteúdo gerado pelo VCL para diferentes aplicações e ambientes de rede. Desse modo, um mesmo vídeo H.264/AVC pode ser transmitido usando-se diferentes métodos de transporte, como RTP/UDP/IP para transmissões pela Internet, formatos ISO MP4 e MMS para armazenamento e MPEG-2 TS

para distribuição de TV digital.

H.264/AVC define também as NAL *Units* (NALUs), que são agrupamentos lógicos do conteúdo codificado. Cada NALU é composto por um cabeçalho de 1 octeto seguido de uma sequência de octetos de tamanho variável correspondente ao conteúdo codificado; essa sequência de octetos deve — idealmente — ser uma parte por si significativa do conteúdo de vídeo, como por exemplo um quadro completo ou uma parte de um quadro (*slice*) de vídeo. A Figura 4.3 apresenta o formato do cabeçalho de um NALU.

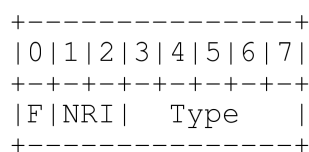


Figura 4.3: Formato do cabeçalho de um NALU.

O primeiro bit do cabeçalho, o bit *F*, também conhecido como *forbidden_zero_bit*, define se existem erros no NALU. Se *F* é igual a zero, então o NALU não contém erros nem violações de sintaxe; caso contrário, se *F* é igual a 1, o NALU pode conter erros. Esses erros são detectados pelos elementos de rede que compõem o caminho de transmissão do NALU.

Os dois bits seguintes, *NRI* ou *nal_ref_idc*, denotam a importância do NALU, ou seja, se o NALU é usado ou não como quadro de referência. *NRI* também denota a importância de entrega do NALU e pode assumir os valores 00, 01, 10 e 11, sendo essa a ordem de importância. Se *NRI* é igual a 00, o NALU não é usado como referência e, como veremos na Seção 4.5.3, poderia ser descartado sem prejuízo à decodificação de outros quadros. Por outro lado, se *NRI* é igual a 11, é importante que a entrega desse NALU seja garantida.

Por fim, os bits correspondentes ao campo *Type*, que podem assumir valores entre 0 e 31, denotam o tipo de informação carregada pelo NALU. Os valores 0, 30 e 31 são reservados para uso futuro; os valores de 1 a 23 são usados para os NALUs simples; valores de 24 a 27 são usados para a agregação de dois ou mais NALUs em um único; e os valores restantes, 28 e 29, são usados quando um NALU precisa ser fragmentado.

4.4.2 Empacotamento RTP dos NALUs

A camada de abstração de rede, com o objetivo de minimizar o impacto das falhas de transmissão, atua de maneira diferenciada para cada meio de transporte e armazenamento, para isso leva em consideração as características da rede a ser utilizada e do vídeo a ser transmitido. Para redes orientadas a pacotes, a camada NAL define os NALUs de tal maneira que cada um deles possa ser transportado em um pacote. Em outras palavras, em redes IP, que são redes orientadas a pacotes, os NALUs são definidos de modo tal que cada um deles deve ser — idealmente — transmitido em um pacote IP.

Observe que a utilização da camada NAL permite a simplificação da política de empacotamento RTP de conteúdo H.264/AVC. A regra é: cada NALU deve ser enviado em um pacote RTP. Exceções a essa regra ocorrem quando os NALUs, por alguma razão, excedem o tamanho máximo dos pacotes RTP ou são pequenos demais; nesses casos aplicam-se, respectivamente, os procedimentos de fragmentação e de agrupamento de NALUs, como está definido em [WHS⁺05].

4.4.3 Conjunto de parâmetros

O padrão H.264/AVC define uma classe especial de dados chamados de conjuntos de parâmetros (*Parameter Sets*) que consistem em informações — tais como o tamanho da imagem, a profundidade de cores, o modo de codificação etc. — fundamentais para a decodificação do vídeo. O uso dos conjuntos de parâmetros elimina redundâncias de dados, reduzindo o volume de dados transmitidos, mas por outro lado cria uma nova fragilidade, pois a perda de um conjunto de parâmetros durante a transmissão de um fluxo H.264/AVC pode comprometer toda a decodificação desse fluxo. Por essa razão, a entrega dos pacotes contendo conjuntos de parâmetros deve ser priorizada e, quando possível, garantida.

A seguir apresentamos como pode ser implementada a entrega dos conjuntos de parâmetros em ordem de preferência, ou seja, a partir da forma mais adequada — e segura — para a forma mais sujeita a perda de pacotes.

1. com o uso de protocolos confiáveis, baseados por exemplo em TCP, antes do início da transmissão dos fluxos de vídeo;

2. por meio de protocolos confiáveis durante a transmissão dos fluxos de vídeo, tomando-se o cuidado de sincronizar as transmissões a fim de obedecer a dependência entre os dados;
3. através da mesma conexão por onde trafegam os fluxos de vídeo; nesse caso, recomenda-se que os pacotes contendo os conjuntos de parâmetros sejam duplicados² na camada de aplicação [WHS⁺05].

Além disso, a probabilidade de entrega dos conjuntos de parâmetros também pode ser melhorada através da implementação de técnicas para recuperação de erros, tais como ARQ e FEC, que serão mencionadas na Seção 5.3.3.

4.5 *Media Aware Network Elements* (MANEs)

O padrão H.264/AVC incluiu mecanismos sofisticados para a recuperação de erros [BBC⁺05] (*error concealment tools*) e para o aumento da robustez da transmissão sobre redes IP [Wen03] e, inclusive, sobre redes sem fio [SHW03]. O funcionamento apropriado desses mecanismos, porém, não depende apenas da codificação H.264/AVC, mas também da implementação de estratégias adaptativas nos elementos presentes ao longo do tráfego. A literatura sobre H.264/AVC, em especial [WHS⁺05], refere-se à implementação de *Media Aware Network Elements* (MANEs), que são elementos de rede, tais como roteadores, *switches* e repetidores, capazes de lidar apropriadamente com o tráfego H.264/AVC. No entanto, apesar de a literatura prever a implementação desses elementos, até o presente momento eles não foram implementados ou — no mínimo — não foram lançados comercialmente. Além disso, se por um lado H.264/AVC levou em consideração as aplicações de *streaming*, as redes IEEE 802.11 ainda apresentam deficiências quando usadas para transmissão de fluxos contínuos de vídeo [dCK03]; essas deficiências se dão porque as redes IEEE 802.11 foram inicialmente projetadas para serem uma forma de acesso à Internet, em que a maioria dos tráfegos são do tipo TCP, e não para transmitir fluxos de vídeo sujeitos a requisitos de tempo real.

O fato de IEEE 802.11 não ser plenamente adequado para aplicações de *streaming* reforça a

²Na página 31 de [WHS⁺05], especifica-se que essa duplicação de conteúdo não deve ser feita simplesmente enviando-se dois pacotes RTP idênticos (com o mesmo número de sequência).

importância da implementação de MANEs 802.11, ou melhor, de pontos de acesso capazes de lidar adequadamente com os fluxos H.264/AVC. O restante deste capítulo apresenta as funcionalidades que um ponto de acesso IEEE 802.11 deve implementar para atuar como um MANE para conteúdo H.264/AVC. Nossa proposta baseia-se na avaliação dos códigos-fonte do ponto de acesso WRT54GS, da Linksys — cujo *firmware* está disponível em <ftp://ftp.linksys.com/opensourcecode> — e do *driver* HostAP — um *driver* para dispositivos IEEE 802.11 equipados com *chipsets* Intersil Prism — que implementa as funções de um ponto de acesso; ou seja, usando-se o HostAP, dispositivos IEEE 802.11 podem ser usados como pontos de acesso. O HostAP é distribuído com licença GPL e está disponível em <http://hostap.epitest.fi>.

A seguir apresentamos as funcionalidades necessárias a um MANE 802.11.

4.5.1 Identificação do tráfego H.264/AVC

O primeiro passo para criação de um MANE é a identificação do tráfego. Dado que a maioria das sessões de transmissão de multimídia usam RTP/UDP/IP como protocolo de transporte, é preciso detectar tráfegos do tipo H.264/AVC inspecionando-se o campo **Payload Type** dos cabeçalhos RTP (vide Figura 3.1). Porém, para a maioria dos novos tipos de tráfego — o que inclui os tráfegos do tipo H.264/AVC —, o valor de **Payload Type** não é predefinido. Sendo assim, seu valor deve ser definido dinamicamente, usando-se um protocolo confiável para controle de sessão, durante o estabelecimento da conexão entre as partes da transmissão [WHS⁺05].

Segundo os consórcios 3GPP e ISMA, dois dos mais influentes consórcios sobre tecnologias de redes sem fio e transmissão de multimídia, os protocolos SIP [RSC⁺02] e SDP [HJ98] devem ser usados para estabelecimento e controle de sessão. Apesar dessa recomendação, o protocolo mais utilizado pelas aplicações — tanto clientes quanto servidoras — para o estabelecimento de conexões RTP é o protocolo RTSP [SRL98].

Os fluxos RTP também podem ser estabelecidos de maneira *ad hoc*, ou seja, sem o uso de protocolos para o estabelecimento de conexão. Nesse caso, resta ao MANE “investigar” os tráfegos UDP e os respectivos cabeçalhos. As técnicas para identificação de tráfegos podem ser consultadas na documentação de alguns analisadores de rede, como o Ethereal, cujos códigos-fonte e documentação estão disponíveis em www.ethereal.org.

4.5.2 Exploração da qualidade de serviço oferecida por IEEE 802.11e

A primeira abordagem para obtenção de um ambiente IEEE 802.11 adequado à distribuição de vídeo é utilizar, quando disponíveis, as funcionalidades de IEEE 802.11e [Sta05]. O protocolo IEEE 802.11e, finalizado em julho de 2005, oferece resumidamente as seguintes funcionalidades para garantia de qualidade de serviço:

- definição de classes de tráfego;
- uso de filas diferentes para cada classe de tráfego;
- janelas privilegiadas de transmissão (*Transmission Opportunity* ou TXOP).

O padrão IEEE 802.11e permite o uso de oito diferentes classes de tráfego, cuja implementação é feita utilizando filas separadas, com diferentes prioridades de acesso ao meio, para cada classe. Para cada fila são atribuídos diferentes valores para os parâmetros que definem a probabilidade de acesso ao meio de transmissão, como por exemplo os intervalos de *backoff*, SIFS e DIFS, entre outros [Gas02].

Por sua vez, a janela privilegiada de transmissão, TXOP, consiste em um período de tempo no qual uma estação pode usar o meio de transmissão diretamente, ou seja, sem esperar pelo direito de transmissão a cada pacote transmitido. O período de tempo TXOP é limitado a um valor máximo *TXOPLimit* e deve ser usado especialmente para reduzir as filas com maior prioridade de transmissão.

Apesar dos avanços propostos, os pontos de acesso compatíveis com IEEE 802.11e foram lançados no mercado apenas recentemente. No entanto, alguns equipamentos anteriores à homologação do padrão, como por exemplo o roteador DLink modelo DI-634M, já trazem alternativas que, se não oferecem o nível de qualidade de serviço proposto em IEEE 802.11e, permitem priorizar os fluxos através de implementações proprietárias. Essas soluções proprietárias permitem, por exemplo, priorizar transmissões que usem certas faixas de portas UDP. Desse modo, pode ser necessária a configuração manual do ponto de acesso para utilização das funcionalidades de QoS.

Nos próximos anos, contudo, antes de IEEE 802.11e ser amplamente implantado, aparentemente um subconjunto de suas funcionalidades deverá ser adotado. A Wi-Fi *Alliance* (WFA) está

preparando o *Wireless Multimedia Extensions* (WME), também conhecido como Wi-Fi *Multimedia* (WMM) que, fortemente baseado em IEEE 802.11e, provê quatro categorias de tráfego a serem mapeadas em quatro classes de aplicações, a saber: voz, vídeo, *best effort* e *background*. Apesar de não prover garantias de vazão, WME deve ser adequado para a maioria das aplicações atuais e sua implementação, por ser mais simples, tem sido bem aceita pela indústria. Em WME, cada pacote é rotulado com o devido *DiffServ Code Point* (DSCP) no cabeçalho IP e então enviado para a respectiva fila no elemento de rede. Mais informações sobre WME podem ser encontradas em http://www.wi-fi.org/getfile.asp?f=WMM_QoS_whitepaper.pdf.

4.5.3 Políticas de descarte de pacotes

A principal vantagem de implementar um MANE 802.11 é que o descarte de pacotes, quando necessário, pode ser feito de maneira mais inteligente, levando-se em consideração as peculiaridades do conteúdo. Por exemplo, suponha um fluxo H.264/AVC sendo transmitido por um MANE 802.11 sujeito a uma situação de congestionamento. Nesse cenário, o MANE 802.11 pode avaliar quais são os pacotes que, caso sejam descartados, impliquem menor impacto na percepção visual do usuário. Deste modo, o MANE 802.11 pode descartar esses pacotes, em ordem crescente de impacto, até que o congestionamento seja contornado. Os descartes também podem ser aplicados com o objetivo de reduzir a latência de comunicação.

Descartes baseados no cabeçalho dos NALUs

A camada NAL presta um grande serviço aos MANEs 802.11 ao definir, no cabeçalho de cada NALU, os valores dos campos **F** e **NRI**. Com base nesses valores e também nos cabeçalhos RTP, recomenda-se a seguinte ordem de descarte de pacotes:

1. pacotes duplicados;
2. pacotes em que **F=1**;
3. pacotes atrasados³;

³Para definir exatamente quando um pacote pode ser considerado atrasado, deve-se levar em consideração a aplicação em questão.

4. pacotes em que NRI é igual a 00, 01, 10 e 11, nessa ordem.

Em complemento a essas regras de descarte, se um pacote contendo um NALU fragmentado for descartado, então todos os fragmentos restantes — e os respectivos pacotes — também devem ser descartados.

Além de monitorar os seus próprios descartes, um MANE também poderia monitorar os pacotes que não lhe foram entregues, ou seja, os pacotes que se perderam no caminho entre o transmissor e o ponto de acesso e que, obviamente, também não serão retransmitidos para a unidade móvel. Desse modo, em função dos pacotes perdidos, o MANE poderia identificar e descartar os NALUs subseqüentes que dependem dos pacotes perdidos para serem decodificados. Para implementar essa análise de dependência, no entanto, é necessário interpretar não só o cabeçalho dos NALUs, mas também o seu conteúdo, o que pode fugir da capacidade de processamento de um MANE, mas isso certamente otimizaria a carga da rede, pois seriam retransmitidos apenas os dados que efetivamente pudessem ser decodificados.

4.6 Revisão das políticas de descarte de pacotes de IEEE 802.11

A implementação de MANEs 802.11, como a definimos até então, não ameniza o impacto de uma das principais características da camada MAC de IEEE 802.11: o descarte de pacotes. Em redes IEEE 802.11, um pacote é descartado após um certo número de tentativas de envio, o que faz sentido quando o objetivo é manter o acesso ao meio sem fio disponível. Porém, as retransmissões e o respectivo descarte, feitos sem levar em consideração a importância dos pacotes, podem ter forte impacto negativo na exibição dos fluxos de vídeo.

Na prática, o descarte de um pacote afeta a exibição de um ou mais quadros de vídeo. Em aplicações de fluxos de vídeo sob demanda, as quais geralmente utilizam *buffers* para pré-armazenamento relativamente grandes (da ordem de segundos), o impacto desses descartes pode ser amenizado por estratégias de retransmissão. Em outras aplicações, como as de videoconferência, a retransmissão torna-se impraticável, pois o quadro retransmitido chegaria após o seu instante de exibição.

Entretanto, devido ao maior controle das filas de transmissão existente nos dispositivos atuais,

devemos repensar a eficácia da política de descartes realizada por IEEE 802.11. Suponha que um tráfego H.264/AVC, composto pelos pacotes $\{P_1, P_2, \dots, P_n\}$, esteja sendo enviado de um ponto de acesso PA para uma unidade móvel UM. Se PA, após várias tentativas de envio de P_i , não recebe de UM a confirmação de recepção do pacote, então P_i é descartado. Após o descarte de P_i , PA tentará enviar o próximo pacote da fila para UM, o pacote P_{i+1} . Porém, se P_i não pôde ser entregue, por que P_{i+1} poderá sê-lo?

Além disso, o Capítulo 3 mostrou que, após um período de falhas, quando o ponto de acesso consegue entregar um pacote para a unidade móvel, em geral, a entrega dos pacotes atrasados é feita em rajada. Ou seja, todos os pacotes, exceto aqueles que foram descartados, são entregues. São entregues com atraso, mas são entregues.

A importância de não descartar os pacotes reside no fato de que, para boa parte das aplicações, um pacote entregue com atraso é melhor do que um pacote perdido. Nas aplicações de fluxos de vídeo, por exemplo, a perda de um pacote pode afetar a decodificação dos pacotes subsequentes. Além disso, os pacotes atrasados podem ser exibidos de maneira adaptativa [LFG03], acelerando-se a sua exibição até que os fluxos retomem a sincronia desejada. Ademais, as aplicações de fluxos de vídeo também podem desacelerar a exibição quando notarem que os *buffers* não estão sendo devidamente preenchidos em função de atrasos na entrega dos pacotes; essas técnicas para aceleração e desaceleração de exibição de vídeo têm sido usadas com sucesso em algumas aplicações para exibição de vídeo, como por exemplo no *MediaPlayer*, da Microsoft.

O cenário em que o descarte de um pacote é preferível ao seu atraso se resume ao uso de aplicações de alta interação, tais como as aplicações de videoconferência. Mas mesmo para essas aplicações, em algumas circunstâncias, o atraso é melhor do que o descarte: suponha a existência de um tráfego de videoconferência, composto pelos pacotes $\{P_1, P_2, P_3, P_4, \dots\}$, em que o envio de quadros é feito a cada 40 ms (ou seja, são exibidos 25 quadros por segundo). Suponha também que os três primeiros pacotes (P_1, P_2 e P_3) teriam sido descartados, o que implicaria 120 ms de informação perdida. Ao invés de serem descartados, esses pacotes poderiam ser entregues atrasados; as avaliações apresentadas no Capítulo 3 indicam que P_1 poderia ser entregue aproximadamente no instante de recebimento de P_4 , e os pacotes P_2, P_3 e P_4 alguns poucos milissegundos depois. Em outras palavras, o que é melhor: P_1, P_2 e P_3 perdidos e P_4 entregue no momento certo? Ou

P_1 , P_2 , P_3 entregues e P_4 entregue com um pequeno atraso?

Sem dúvida, a entrega, mesmo com atrasos, é mais promissora para a aplicação de técnicas adaptativas de exibição. Na pior das hipóteses, ao custo de alguns milissegundos de atraso, delegamos para a aplicação a decisão de descartar ou não os pacotes atrasados [SRC84].

4.7 Descartes preguiçosos

Desse modo, com os avanços na gerência das filas de transmissão de IEEE 802.11, o envio de pacotes para uma unidade móvel poderia ser planejado de forma mais racional. Ao invés de descartar um pacote e passar para o próximo na fila, seria interessante que os pacotes fossem descartados apenas quando o espaço de armazenamento da fila estivesse próximo de atingir os seus limites⁴.

Em suma, o descarte de pacotes para aplicações de *streaming* poderia ser feito apenas quando um pacote fica velho demais para ser exibido ou quando a capacidade da fila de transmissão do ponto de acesso está próxima de ser excedida. Dado o aumento da capacidade de processamento e armazenamento dos pontos de acesso IEEE 802.11 e os avanços na gerência de filas, essas novas políticas poderiam ser implementadas com pouco ou nenhum prejuízo à operação dos pontos de acesso ou transmissões concorrentes. Apesar disso, infelizmente, nos pontos de acesso atuais, a política de retransmissão e descarte é implementada diretamente no *chipset* dos elementos de rede, o que representa uma forte restrição para a implementação da política de descartes aqui proposta. No entanto, existe uma tendência para que essa restrição seja superada num futuro próximo, pois os dispositivos IEEE 802.11 tendem a ser cada vez mais programáveis [CBB04]. Por exemplo, a tecnologia *Native Wi-Fi*, que será incorporada na próxima versão do sistema operacional da Microsoft, o Windows Vista, transfere para os sistemas operacionais parte das funcionalidades hoje implementadas nos *chipsets* dos dispositivos de rede, o que permitiria a implementação de estratégias adaptativas específicas e adequadas para cada sistema.

⁴Em um ponto de acesso WRT54GS v1.1, da Linksys, esse limite é de aproximadamente 32 MB, vide <http://en.wikipedia.org/wiki/WRT54G> para mais informações.

4.7.1 Algoritmo

Baseado no fato de que pode ser vantajoso adiar os descartes, seria interessante que as futuras versões de IEEE 802.11 implementassem um modelo de descarte “preguiçoso” que, através da gerência racional das filas de transmissão, reduzisse o volume de pacotes perdidos por descarte. O modelo de descarte “preguiçoso” deveria implementar os procedimentos definidos a seguir:

- suspensão da transmissão de pacotes para uma UM, se as transmissões para essa UM não estiverem sendo bem sucedidas; deve-se proceder normalmente com os tráfegos restantes;
- avaliação periódica das conexões suspensas (atividade de *probing*), o que pode ser feito usando-se os próprios pacotes do fluxo;
- a execução do descarte de pacotes deve ser feita, preferencialmente considerando a natureza dos fluxos, se as filas de transmissão estiverem próximas de alcançar o seu limite de armazenamento. Por exemplo, para fluxos H.264/AVC, a prioridade de descarte definida na Seção 4.5.3 deve ser observada;
- envio dos pacotes pendentes, se possível, usando o mecanismo de TXOP, caso a capacidade de transmissão para uma UM suspensa for restabelecida.

4.8 Trabalhos relacionados e conclusões

Até a escrita final deste capítulo, não encontramos nenhuma implementação de um MANE para conteúdo H.264/AVC, nem propostas de algoritmos para essa finalidade, como o descrito na Seção 4.7.1. No entanto, desenvolvidas a partir da parceria entre a operadora americana de VoIP Vonage e a Linksys, MANEs 802.11 para tráfegos de VoIP foram lançados em 2005. É natural, portanto, que MANEs 802.11 para H.264/AVC comecem a ser produzidos.

É importante observar que a estratégia de adequação dos elementos de rede a finalidades específicas tem ganhado força recentemente. Exemplos disso são as iniciativas Fon (www.fon.com) — apoiada financeiramente por empresas de grande porte, tais como Google e Skype — e WiFree (www.wifree.org), que disponibilizam *firmwares* projetados para garantir o compartilhamento seguro de pontos de acesso em locais públicos. Em outras palavras, as primeiras adequações de

firmwares — os precursores para os MANEs — já foram lançadas. Desse modo, pode ser apenas uma questão de tempo a chegada ao mercado dos primeiros MANEs, que deverão ser elementos de rede modulares e realmente configuráveis, que permitam, por exemplo, adicionar ou remover módulos para lidar adequadamente com determinadas classes de tráfego; ou seja, deverão ser elementos de rede realmente configuráveis às necessidades dos usuários.

Capítulo 5

Desenvolvimento de aplicações adaptativas para redes IEEE 802.11

*Vivemos numa realidade móvel à qual procuramos adaptar-nos
como as algas que se dobram sob o ímpeto das ondas do mar.*

Giuseppe Tomasi di Lampedusa

As redes sem fio IEEE 802.11 (Wi-Fi) firmaram-se como o padrão *de facto* para redes locais sem fio por oferecerem mobilidade e alta capacidade de vazão a custos relativamente baixos. Trazem, porém, novos desafios ao desenvolvimento de aplicações, como a ocorrência de migrações (*handoffs*) e a alta variabilidade da qualidade da conexão.

Durante a execução de *handoffs* em redes IEEE 802.11, as aplicações podem experimentar períodos de desconexão da ordem de alguns segundos, o que é uma forte restrição — especialmente para aplicações de tempo real. Por sua vez, a alta variabilidade da qualidade da conexão em IEEE 802.11 faz com que as aplicações experimentem diferentes condições de conexão ao longo de sua execução. No entanto, apesar dessas diferenças entre as redes sem fio e as redes cabeadas tradicionais, do ponto de vista do desenvolvimento de aplicações, as redes sem fio IEEE 802.11 têm sido muitas vezes tratadas como redes Ethernet.

Além disso, os protocolos UDP e TCP, que geralmente são utilizados na camada de transporte,

não foram projetados levando em consideração as características das redes sem fio, mas sim das redes Ethernet cabeadas. A fim de adequar esses protocolos às redes sem fio, foram feitas propostas para a otimização de TCP [BPSK97, PRR⁺03, SdMBC04] e de UDP [LDP⁺04], mas essas otimizações, em geral, não foram amplamente adotadas.

Dado que as abstrações providas por TCP e UDP não são plenamente capazes de lidar com as peculiaridades das redes IEEE 802.11, uma das alternativas é a implementação de técnicas adaptativas na camada de aplicação. Considerando essa necessidade, este capítulo tem dois objetivos: prover uma referência aos desenvolvedores interessados em desenvolver aplicações adaptativas para redes IEEE 802.11 e, principalmente, apresentar uma estratégia adaptativa que, implementada na camada de aplicação, seja capaz de aumentar a eficiência de transmissão em redes IEEE 802.11. Para cumprir esses objetivos, o conteúdo aqui exposto está organizado da seguinte maneira: a Seção 5.1 discute os trabalhos relacionados e contextualiza este trabalho em relação a essas referências. A Seção 5.2 apresenta os novos desafios e problemas referentes ao desenvolvimento para redes IEEE 802.11. A Seção 5.3 descreve algumas práticas e técnicas úteis para a solução desses problemas. A Seção 5.4 introduz os métodos que podem ser utilizados para interagir com as interfaces de rede sem fio a partir da camada de aplicação e, sobretudo, mostra como podem ser obtidas e interpretadas as informações sobre a qualidade de sinal e de conexão em redes IEEE 802.11. A Seção 5.5 descreve os princípios e a implementação de uma estratégia adaptativa que, baseada na qualidade do sinal, aumenta significativamente a eficiência das transmissões sobre redes IEEE 802.11. Por fim, a Seção 5.6 traz as considerações finais deste capítulo.

5.1 Trabalhos relacionados

Ainda é relativamente pequeno o número de trabalhos que abordaram IEEE 802.11 a partir da camada de aplicação. Dentre esses trabalhos, podemos destacar [OCL05], que propôs uma API para monitoração de redes sem fio, e [dSOD⁺01], que apresentou um modelo para exibição adaptativa de conteúdo. A maioria dos trabalhos, porém, concentra-se em camadas subjacentes à camada de aplicação; por exemplo, na otimização da camada MAC [TG05, Gup05, dMdRJ05, dCdR05], dos protocolos de transporte [SdMBC04], ou na caracterização do tráfego [MGCM01, IL02].

Desse modo, ao se observar as pesquisas realizadas anteriormente, é possível identificar uma

lacuna na qual se esquece de que é na camada de aplicação que residem as maiores possibilidades para adaptação [SRC84], pois é nela que se podem aplicar técnicas adaptativas baseadas nos requisitos das aplicações, o que não poderia ser feito nos elementos de rede e protocolos subjacentes, sem prejuízo à generalidade dos mesmos. Contudo, olhando para o passado recente, é possível ver algumas razões para a escassez de pesquisa nessa área. A principal foi, sem dúvida, a limitada capacidade de interação entre a aplicação e as interfaces de rede. Mas essa limitação foi muito reduzida com as novas versões de NDIS (*Network Driver Interface Specification*) e com a incorporação ao núcleo do Linux de ferramentas para redes sem fio (*Wireless Tools*). Outra razão para a lacuna, mas que ainda persiste, é a dificuldade de condução de experimentos práticos sobre redes IEEE 802.11, especialmente daqueles relacionados à mobilidade, pois a alta variabilidade intrínseca às redes sem fio dificulta a execução de experimentos representativos e reproduzíveis [KNE03].

A fim de tentar sanar essa lacuna, este capítulo concentra-se na camada de aplicação, apontando não apenas os principais problemas vinculados à programação para redes IEEE 802.11, mas também algumas técnicas que podem ser úteis à sua solução. Veremos que, apesar de ainda ser difícil conduzir pesquisas empíricas em redes sem fio, é possível implementar estratégias adaptativas capazes de otimizar o desempenho das aplicações executadas sobre redes IEEE 802.11.

5.2 Novos desafios impostos pelas redes IEEE 802.11

As redes IEEE 802.11, devido às suas características e peculiaridades, trazem novos desafios para o desenvolvimento de aplicações adaptativas. Podemos classificar esses desafios em três grupos: os inerentes às redes sem fio, os gerados pelos próprios mecanismos adaptativos presentes na especificação do protocolo IEEE 802.11 e os relativos à diversidade de comportamento entre dispositivos IEEE 802.11. Esta última especificidade pode dar-se tanto pelo uso de dispositivos IEEE 802.11 de diferentes fabricantes, quanto pelo uso de diferentes configurações.

5.2.1 Problemas intrínsecos às redes sem fio

As redes IEEE 802.11 estão sujeitas aos problemas intrínsecos das redes sem fio [Sch00, GC04], dentre os quais estão a maior susceptibilidade a interferências externas, o desvanecimento do sinal devido à mobilidade e à distância ao ponto de acesso, a propagação do sinal por dois ou mais

caminhos (efeito *multipath fading*) e a existência de nós invisíveis (*hidden nodes*).

Uma das razões para a maior susceptibilidade a interferências externas é o fato de outros dispositivos eletrônicos, que vão desde fornos microondas até interfaces Bluetooth [GCR03], usarem a mesma frequência de rádio que IEEE 802.11. Além dos dispositivos eletrônicos, existem também outros fatores capazes de interferir na qualidade de uma conexão sem fio, sendo alguns deles obviamente não gerenciáveis pela camada de aplicação, como a orientação e alturas das antenas, o enfraquecimento do sinal devido a sobreposição de objetos, a umidade do ar etc. Em suma, as redes IEEE 802.11, devido ao grande número de fontes de interferência, são intrinsecamente mais variáveis. Desse modo, a qualidade do sinal nessas redes pode variar substancialmente e em curtos intervalos de tempo, dificultando assim o projeto de algoritmos adaptativos.

Além disso, as redes sem fio estão sujeitas à ocorrência de nós invisíveis — que são unidades móveis visíveis ao ponto de acesso (ou ao nó central em redes *ad hoc*), mas não visíveis para outras unidades móveis. Em outras palavras, em redes sem fio, duas unidades móveis pertencentes à mesma rede podem não “ouvir” uma à outra. Da existência de nós invisíveis decorrem as dificuldades para coordenação do acesso ao meio em redes sem fio; esta é, portanto, uma das razões por que as redes IEEE 802.11 usam mecanismos de *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA), ao invés de *Carrier Sense Multiple Access with Collision Detection* (CSMA/CD). Uma vez que alguns nós podem não ouvir a transmissão de outros, não é possível detectar as colisões; desse modo, resta lançar mão de mecanismos que sejam, ao menos, capazes de evitar essas colisões, que é o caso de CSMA/CA.

Outro problema que pode ser considerado intrínseco às redes sem fio são os *handoffs*. Presentes apenas nas redes sem fio, como visto no Capítulo 3, os *handoffs* podem acarretar períodos de desconexão da ordem de alguns segundos [RS05].

5.2.2 Problemas intrínsecos às redes IEEE 802.11

Os próprios mecanismos adaptativos que fazem parte da especificação de IEEE 802.11 também podem dificultar a interpretação das métricas de adaptação, o que acarreta dificuldades para o projeto de aplicações adaptativas.

O mecanismo de *Automatic Repeat reQuest* (ARQ), por exemplo, é responsável por reenviar um

quadro de transmissão (MSDUs na nomenclatura de IEEE 802.11) caso sua entrega não tenha sido confirmada (*acknowledged*). O número de tentativas de reenvio é definido em função do tamanho do quadro; para os menores do que o *RTS Threshold*, o número padrão de reenvios é sete e, para os maiores, é quatro [Gas02]. Resumindo, devido à atuação do mecanismo de ARQ, um quadro é reenviado pelo menos quatro vezes antes que a sua perda seja notificada ao protocolo de nível superior. Desse modo, ao usar retransmissões para melhorar a confiabilidade da transmissão, IEEE 802.11 faz com que as taxas de perda de pacotes não reflitam fielmente a real condição do canal. Na prática, observamos que, em redes IEEE 802.11, a qualidade da conexão já se encontra bastante deteriorada quando as perdas de pacotes começam a ser percebidas.

Assim como o mecanismo de ARQ, outros mecanismos internos ao protocolo IEEE 802.11 afetam fortemente as métricas de adaptação.¹ Dentre eles podemos citar: *Automatic Rate Control* (ARC), *Power Saving Mode* (PSM) e *Request to Send / Clear to Send* (RTS/CTS). Por exemplo, o mecanismo de RTS/CTS, ao tentar eliminar o problema de *hidden-nodes*, agrega alta variação à latência. PSM, por sua vez, pode gerar períodos de inatividade que podem ser interpretados como congestionamentos pelo protocolo TCP. Por fim, o mecanismo de ARC é o pivô de dois problemas: a má distribuição da capacidade de transmissão em redes IEEE 802.11 [TG05] e a redução da capacidade de vazão total do sistema em função da degradação da qualidade de sinal de apenas uma unidade móvel [HRBSD03, dCdR05].

5.2.3 Problemas relacionados à diversidade das implementações

Outro desafio atualmente imposto pelas redes IEEE 802.11 é a diversidade de comportamento entre dispositivos Wi-Fi, o que advém do uso, tanto de dispositivos IEEE 802.11 de diferentes fabricantes, quanto de diferentes configurações. Por questões de simplicidade, mas também para não invadir o espaço de diferenciação competitiva entre os fabricantes de dispositivos de rede, o padrão IEEE 802.11 deixa em aberto alguns pontos da especificação; uma parte destes, porém, possui forte influência no comportamento das transmissões sobre essas redes.

¹Cabe dizer que não contestamos a importância dos mecanismos adaptativos de IEEE 802.11, pois eles foram e são fundamentais, cada um a seu modo, para a ampla utilização dessas redes. Mas é importante chamar a atenção para as consequências negativas que esses mecanismos trazem para o projeto de aplicações adaptativas para redes IEEE 802.11.

Por exemplo, o padrão IEEE 802.11 não especifica como o mecanismo de ARC deve escolher a modulação (taxa de transmissão) a ser usada em caso de degradação do sinal. Ao não definir quais, nem como essas métricas de adaptação devem ser usadas pelo mecanismo de ARC, o padrão IEEE 802.11 delega aos fabricantes essa definição; sendo assim, cada fabricante define os algoritmos e limiares de adaptação segundo as peculiaridades de seus dispositivos; ou seja, mediante a degradação do sinal, mecanismos ARC de dispositivos de fabricantes diferentes atuam de maneira diversa e, em função disso, experimentam diferentes taxas de perda de pacotes, latência, *jitter* etc.

O padrão IEEE 802.11 também deixa em aberto como as interfaces de rede (*Network Interface Cards*, ou NICs) apresentam a qualidade do sinal. NICs de diferentes fabricantes, quando submetidos a uma mesma condição de conexão, podem relatar diferentes níveis de qualidade de sinal, o que dificulta a interpretação dessa informação e, conseqüentemente, o seu uso como métrica de adaptação. Por exemplo, os NICs IEEE 802.11 dos fabricantes Atheros, Symbol e Cisco usam escalas diferentes para reportar a qualidade do sinal [Wil02]: de 0 a 60, de 0 a 31 e de 0 a 100, respectivamente. Além disso, o valor médio da qualidade do sinal nesses NICs (30, 16 e 50, respectivamente) corresponde a diferentes valores de força do sinal² (-65 dBm, -70 dBm e -58 dBm, respectivamente).

Outro fator que também contribui para a diversidade de comportamentos em IEEE 802.11 é a existência de variantes interoperantes do padrão; por exemplo, NICs IEEE 802.11g podem acessar redes 802.11b e g, mas submetidos a redes com capacidades de vazão diferentes.

Em suma, o uso de diferentes NICs e configurações implica a utilização de conexões com diferentes características. Portanto, a diversidade de comportamentos dificulta o desenvolvimento de aplicações adaptativas, pois a interpretação das métricas de adaptação (tais como taxas de perda de pacotes, latência, *jitter*, qualidade do sinal etc.) pode variar em função dos ambientes utilizados. Por exemplo, a taxa de perda de pacotes varia em função da sensibilidade das antenas utilizadas pelo NIC, o que varia de fabricante para fabricante e, inclusive, de modelo para modelo.

Resta observar que a diversidade de implementação e de comportamento entre interfaces não é um problema necessariamente exclusivo das redes IEEE 802.11: afinal, também existem diversos

²A força do sinal, para conveniência de notação, é normalmente expressa em dB-miliwatts (dBm), que é a representação logarítmica de miliwatts (mW). Por exemplo, a sensibilidade das interfaces IEEE 802.11b vai de 100 mW a 0.0000000002511 mW ou de 20 dBm a -96 dBm, sendo essa última representação bem mais conveniente.

fabricantes para interfaces *Ethernet*. Mas a diversidade de comportamento é mais perceptível nas redes IEEE 802.11, pois a mobilidade implica maior interação entre interfaces de diferentes fabricantes, inclusive com a possibilidade de interação entre interfaces diferentes durante uma única conexão.

Assim, as redes IEEE 802.11 agregam ao desenvolvimento de aplicações novos problemas, advindos da natureza das redes sem fio, das diferenças de implementação e dos próprios mecanismos de IEEE 802.11. Estes problemas são percebidos na camada de aplicação da seguinte maneira:

- maiores taxas de perda de pacotes;
- maior incidência de perda de pacotes em rajadas;
- maior variação na latência;
- maior incidência de entrega de pacotes em rajada;
- variações abruptas na qualidade da conexão;
- períodos de desconexão;
- menor previsibilidade da conexão.

As abstrações atualmente usadas na camada de transporte não são capazes de resolver completamente esses novos problemas trazidos pelas redes IEEE 802.11, tais como os períodos de desconexão em função dos *handoffs*. Desse modo, cabe aos desenvolvedores munir as aplicações com estratégias capazes de resolver ou ao menos amenizar as características das redes sem fio.

5.3 Desenvolvimento para redes IEEE 802.11: práticas recomendadas e técnicas úteis

A seguir, apontamos algumas recomendações e técnicas que devem ser consideradas no desenvolvimento de aplicações para redes IEEE 802.11.

5.3.1 Tamanhos dos pacotes

A primeira recomendação para o desenvolvimento de aplicações para redes IEEE 802.11 é relativa ao tamanho dos pacotes. De maneira geral, a utilização de pacotes pequenos implica menores taxas de erros [SHW03], mas, por outro lado, reduz a eficiência de utilização da rede devido à maior inserção de sobrecargas (*overheads*). Estas advêm, por exemplo, da transmissão de um maior número de cabeçalhos e do maior número de troca de mensagens de confirmação de entrega de quadros (mensagens ACK). Além disso, é importante que os pacotes não sejam maiores do que os limites de fragmentação da camada *Ethernet*, que é de 1.500 octetos. Ou seja, os pacotes UDP e TCP não devem ser maiores do que 1.472 e 1.460 octetos, respectivamente. Ultrapassar esses limites de tamanho implica um número maior de fragmentações e desfragmentações, o que deve ser evitado. A escolha do tamanho ideal dos pacotes também depende fortemente da natureza da aplicação — recomenda-se que a transmissão de fluxos contínuos de vídeo MPEG, quando possível, seja feita usando-se datagramas UDP de aproximadamente 750 octetos [IL02]. Ao usar esse tamanho de pacotes para aplicações de vídeo, chega-se a um compromisso entre a sobrecarga (eficiência de uso da rede), a probabilidade de perda de pacotes e o impacto da perda de pacotes na exibição de vídeo.

5.3.2 Novas métricas de adaptação

Em redes IEEE 802.11, além das métricas de adaptação tradicionais, tais como a taxa de perda de pacotes e a latência, existem novos fatores a serem monitorados e que também podem ser utilizados como métricas de adaptação. Dentre essas novas métricas destacam-se: a qualidade do sinal e o número de retransmissões executadas pela camada MAC. A qualidade do sinal, como veremos adiante, oferece bons indicadores sobre a necessidade de adaptação e pode também indicar a necessidade de execução de *handoffs*. Por sua vez, o número de retransmissões executadas pela camada MAC é uma métrica mais adequada do que a informação de perda de pacotes, pois esta última, como foi mencionado anteriormente, é mascarada pelas retransmissões automáticas. Essas novas métricas de adaptação para redes sem fio, das quais a qualidade do sinal e o número de retransmissões são apenas os exemplos mais importantes, podem ser obtidas pela camada de aplicação através dos métodos que serão descritos na Seção 5.4.

Outro fator importante a ser monitorado é a taxa de conexão. Dispositivos IEEE 802.11 — diferentemente das interfaces Ethernet — podem conectar-se a diferentes taxas de transmissão. A escolha da taxa mais apropriada é feita pelo mecanismo de ARC [Gas02], que pode reduzir significativamente a capacidade de transmissão sem qualquer notificação à aplicação. Em redes IEEE 802.11b, por exemplo, a taxa de conexão pode ser reduzida de 11 Mbps para 5,5 Mbps, o que diminui a vazão efetiva de 5,6 Mbps para 1,6 Mbps, aproximadamente. Portanto, as aplicações devem estar preparadas para reduções abruptas na capacidade de vazão. Mas atualmente, apesar do impacto da atuação do mecanismo de ARC nas aplicações, não há métodos simples para a monitoração da taxa de transmissão. No próximo capítulo propomos um mecanismo para auxiliar a disponibilização das escolhas feitas pelo mecanismo ARC à camada de aplicação.

A importância de monitorar as taxas de transmissão reside também na interpretação confiável das métricas de adaptação, pois as mesmas taxas de erro não podem ser interpretadas da mesma forma quando obtidas usando-se diferentes taxas de transmissão; essa mesma lógica aplica-se também a outras métricas de adaptação, tais como a latência e a vazão.

Também é importante observar que diferentes ambientes de execução (computadores, NICs, pontos de acesso e as respectivas configurações) podem implicar diferentes comportamentos de perdas de pacotes e de latência. A latência, por exemplo, pode variar significativamente após a execução de um *handoff*. Sendo assim, algoritmos adaptativos robustos não devem usar limiares de adaptação fixos (*hard coded*), mas devem valer-se de mecanismos de calibração capazes de determinar dinamicamente os limiares de adaptação.

A carga total na rede seria uma métrica importante para o controle de admissão em redes IEEE 802.11. Porém, esta métrica não está disponível para as aplicações. Nessas redes, a capacidade dos pontos de acesso é compartilhada pelas unidades móveis, de tal sorte que uma unidade móvel pode exceder a capacidade do sistema e, dessa forma, interferir negativamente no tráfego de outras unidades [dCK03]. Na ausência de métodos eficientes para controle de admissão de fluxos, é preciso que as aplicações façam uso moderado dos recursos de rede; por exemplo, adotando estratégias incrementais de utilização dos recursos (como faz o protocolo TCP, apesar de que o controle de congestionamento de TCP é sabidamente não apropriado para redes sem fio [BPSK97]). Isso se aplica especialmente aos sistemas baseados em UDP que, por não possuírem mecanismos integrados

para controle de congestionamentos, podem facilmente exceder a capacidade de transmissão da rede.

5.3.3 Recuperação de erros

Devido às razões expostas anteriormente, as redes IEEE 802.11 estão sujeitas a taxas maiores de perdas de pacotes do que as redes cabeadas tradicionais [SdMBC04], mas o impacto dessas perdas pode ser minimizado pela utilização, na camada de aplicação, de técnicas para recuperação de erros, tais como: *Automatic Repeat reQuest* (ARQ) e *Forward Error Correction* (FEC). Os mecanismos de ARQ³ baseiam-se no reenvio de pacotes [SP04] e os de FEC, por sua vez, transmitem informação redundante nos pacotes, de tal forma que, corrompido um deles, este pode ser reconstruído [HV02a, MGCM01].

Em termos de uso de banda, FEC costuma ser menos eficiente do que ARQ, mas em geral é capaz de recuperar as perdas mais rapidamente do que ARQ. Nas redes IEEE 802.11, entretanto, é comum que a perda de pacotes ocorra em rajada (em seqüência), especialmente quando a conexão já se encontra bastante deteriorada; nessas situações, ARQ é mais adequado do que FEC [VcFRJ01], pois, ao se perderem pacotes em seqüência, normalmente perdem-se também as informações usadas para reconstruí-los. As técnicas para minimização das perdas de pacotes podem ser aplicadas a todos os pacotes da transmissão, ou podem ser aplicadas apenas para evitar a perda dos pacotes considerados mais importantes. Por exemplo, na transmissão de fluxos de vídeo H.264, normalmente feita sobre UDP, existem pacotes do tipo *parameter set* que são essenciais para a decodificação do vídeo. Recomenda-se que a transmissão desses pacotes seja protegida [EY05], quer pela aplicação de FEC, de ARQ, pela simples duplicação de envio desses pacotes, quer pelo uso de canais confiáveis (vide Capítulo 4).

5.3.4 Operação desconectada

Uma das principais diferenças de IEEE 802.11 em relação às outras redes é a maior incidência de períodos de desconexão, que podem ocorrer por várias razões, tais como *handoffs*, interferências, sombreamento etc. Um agravante para a ocorrência das desconexões é o fato de que elas normal-

³A própria camada MAC de IEEE 802.11, ao fazer retransmissões, implementa um mecanismo de ARQ.

mente ocorrem de maneira abrupta, sem que haja tempo hábil para o acionamento de mecanismos capazes de notificá-las ou de amenizar seus efeitos. Tomemos como exemplo uma aplicação de vídeo sob demanda em que uma unidade móvel requisita um vídeo a um servidor na Internet. Nessas aplicações, os fluxos de vídeo são geralmente implementados usando-se UDP e, em geral, os servidores não possuem mecanismos adaptativos para interromper a entrega dos quadros de vídeo quando a unidade móvel estiver experimentando períodos curtos de desconexão. Sem tais mecanismos, os quadros de vídeo enviados pelo servidor continuarão chegando ao ponto de acesso que, por sua vez, tentará transmiti-los para a unidade móvel — mas obviamente sem sucesso. Isso acarreta, antes de tudo, o desperdício da capacidade do ponto de acesso, o que pode afetar negativamente a transmissão para outras unidades móveis por ele servidas. Em suma, as aplicações para redes IEEE 802.11 precisam implementar mecanismos capazes de lidar com períodos de desconexão. Isso pode ser feito através de mecanismos para operação desconectada, ou mesmo através da implementação de mecanismos mais simples — mas muito úteis — como a troca de mensagens *KeepAlive*, que permitem verificar constantemente a conectividade de uma unidade móvel. Uma vantagem adicional de implementar as mensagens *KeepAlive* é que elas podem também carregar informações úteis à adaptação, como por exemplo a qualidade da conexão, servindo desse modo como uma forma de sinalização.

Na Seção 5.5, apresentaremos a implementação de um novo algoritmo que ilustra o uso das mensagens *KeepAlive*, mas antes, vejamos rapidamente como pode ser obtida, entre outras informações, a qualidade do sinal em redes IEEE 802.11.

5.4 Interação entre aplicações e interfaces IEEE 802.11

É importante lembrar que as conexões IEEE 802.11 não precisam ser necessariamente tratadas como caixas-pretas: diversos NICs disponíveis no mercado possibilitam a alteração de suas configurações a partir da camada de aplicação, permitindo, entre outras coisas, a alteração de parâmetros como o limiar de fragmentação e o uso do mecanismo RTS/CTS. Vejamos a seguir como isso pode ser feito para os sistemas Windows e Linux.

5.4.1 Em sistemas Windows: *Network Driver Interface Specification* (NDIS)

Nos sistemas operacionais da família Windows, a forma mais simples para interagir com os NICs a partir da camada de aplicação⁴ é através de NDIS. Para cada tecnologia de rede, tais como Ethernet ou IEEE 802.11, NDIS especifica um conjunto de *Object IDentifiers* (OIDs), cujo formato é `OID_XXX_YYY`, onde XXX é a tecnologia de rede e YYY é o objeto desejado. Por exemplo, `OID_802_3_YYY` representa o objeto YYY das redes Ethernet e `OID_802_11_RSSI` representa a qualidade do sinal (*Received Signal Strength Indication* ou RSSI) de um NIC para redes IEEE 802.11. A lista completa de OIDs pode ser encontrada em <http://msdn.microsoft.com>. Contudo, NDIS define três classes de OIDs: obrigatórios, recomendados e opcionais. Os OIDs obrigatórios são aqueles considerados essenciais à interação com o NIC, como por exemplo `OID_802_11_RSSI`. Os OIDs recomendados são aqueles considerados importantes, mas não essenciais, como por exemplo `OID_802_11_STATISTICS`, que fornece estatísticas sobre a conexão. Por fim, existem também os OIDs opcionais, como a requisição `OID_802_11_RTS_THRESHOLD`, que permite alterar o valor do *RTS Threshold*. A principal implicação da classificação dos OIDs nesses três grupos é que nem todo OID funciona para todo NIC. Por exemplo, em nossos experimentos, dentre quatro NICs testados, a requisição de `OID_802_11_STATISTICS`, cuja implementação não é obrigatória, funcionou em apenas um deles.

A melhor fonte de informação sobre como usar o NDIS são os códigos-fonte da *Wireless Research API* (WRAPI, vide <http://ramp.ucsd.edu/pawn/wrapi>), porém essa biblioteca pode não funcionar corretamente nas versões mais novas do sistema operacional Windows. No endereço <http://gsd.ime.usp.br/software/WirelessMultimedia/ndis> podem ser encontrados os passos necessários para o uso de NDIS e um programa-exemplo, escrito em linguagem C, que chama o objeto `OID_802_11_RSSI`. A partir desse programa exemplo, do código fonte de WRAPI e da documentação de NDIS contendo a funcionalidade de cada OID, é relativamente simples criar suas próprias DLLs para interação com as interfaces de rede.

Para as futuras versões dos sistemas operacionais da Microsoft, por exemplo o Windows Vista, algumas das funções de NDIS serão preferencialmente realizadas usando-se *Native* Wi-Fi. Para

⁴Apesar de ser executado no espaço do núcleo do sistema operacional, NDIS pode ser acessado pela camada de aplicação através de chamadas IOCTL ao *driver* NDISPROT, que se encontra disponível na Web.

mais informações sobre *Native* Wi-Fi, consulte <http://msdn.microsoft.com/library/en-us/nativewifi/nwifi/portal.asp>.

5.4.2 Em sistemas Linux: `/proc/net/wireless`

Nos sistemas operacionais baseados em Linux, a forma mais simples para interagir com os NICs IEEE 802.11 é através das ferramentas para redes sem fio (*Wireless Tools*, vide http://www.hp1.hp.com/personal/Jean_Tourrilhes/Linux/Wireless.html), em particular usando-se as ferramentas `iwconfig` e `iwspy`. Além disso, as *Wireless Tools* também disponibilizam a biblioteca `iwlib.o`, permitindo a configuração da conexão IEEE 802.11 seja feita de dentro das aplicações. As principais informações sobre a qualidade da conexão sem fio, contudo, estão resumidas no pseudo-arquivo⁵ `/proc/net/wireless`, cujo formato é apresentado na Figura 5.1.

Inter-	sta-	Quality				Discarded packets					Missed	WE
face	tus	link	level	noise		nwid	crypt	frag	retry	misc	beacon	16
wlan0:	0000	100	224	0		0	0	0	6	122829	0	
wlan1:	0000	55.	218.	163.		0	0	0	1	0	0	

Figura 5.1: O pseudo-arquivo `/proc/net/wireless`.

A Figura 5.1 mostra as informações referentes a duas interfaces IEEE 802.11, `wlan0` e `wlan1`, e as respectivas estatísticas sobre a qualidade da conexão. Para detalhes sobre o significado de cada informação fornecida no pseudo-arquivo `/proc/net/wireless`, consulte o manual da aplicação `iwconfig`. No contexto deste trabalho, os parâmetros mais relevantes são:

- **link:** qualidade do sinal;
- **level:** força do sinal (*signal strength*);
- **noise:** força do ruído;
- **retry:** número de retransmissões.

Na Figura 5.1, observe que os NICs reportam a qualidade do sinal de maneiras completamente distintas; `wlan0` reporta a qualidade como um indicador relativo (100%) e `wlan1` reporta a relação

⁵Os arquivos localizados no diretório `/proc` são chamados pseudo-arquivos, pois, na verdade, são arquivos virtuais. Vide <http://www.ebcvg.com/articles.php?id=108> para mais informações (última visita em maio de 2006).

entre a força do sinal e do ruído (ou seja, `link=level-noise`). Resumidamente, os valores de `link`, `level` e `noise` obedecem a uma aritmética de oito bits, podendo assumir valores entre 0 e 255; porém, cada fabricante estabelece como os valores de qualidade de sinal são informados. Para saber mais a esse respeito, consulte [Wil02] e os códigos-fonte das ferramentas para redes sem fio de Linux. Na próxima seção, usaremos as informações de qualidade do sinal contidas no pseudo-arquivo `/proc/net/wireless` para implementar um novo algoritmo adaptativo para otimização de transmissões sobre redes IEEE 802.11.

5.5 Adaptação de fluxos contínuos baseada na qualidade do sinal

Vimos nas seções anteriores os principais desafios para o desenvolvimento de aplicações adaptativas para redes IEEE 802.11, as principais técnicas para ajudar nesse desenvolvimento e as formas para obtenção de informações sobre a qualidade das conexões sem fio. Esta seção, por sua vez, apresenta uma nova estratégia adaptativa que, baseada na qualidade do sinal das redes IEEE 802.11, adequou os fluxos de dados à capacidade da rede, aumentando assim a eficiência das transmissões. Mas, antes de apresentar os algoritmos, é necessário entender o comportamento das transmissões sobre redes IEEE 802.11 em condições de mobilidade.

A Figura 5.2 apresenta o comportamento de um fluxo contínuo enviado de uma máquina servidora para a unidade móvel através de um ponto de acesso, na presença de mobilidade. Ou seja, durante o experimento, a unidade móvel é levada de um ponto próximo do ponto de acesso para um ponto mais afastado e trazida de volta. Da Figura 5.2(a) podem-se extrair algumas observações iniciais sobre o comportamento de um fluxo e de seus indicadores em função da mobilidade:

- a qualidade do sinal varia significativamente;
- o ruído apresenta pequenas variações, e;
- a capacidade de recepção é afetada pela mobilidade.

No experimento apresentado na Figura 5.2(a), a taxa de recepção, que é inicialmente de 2 Mbps, é reduzida quando a unidade móvel é afastada do ponto de acesso. Essa redução deve-se à atuação do mecanismo de ARC do ponto de acesso, que reduz a taxa de conexão ao identificar a

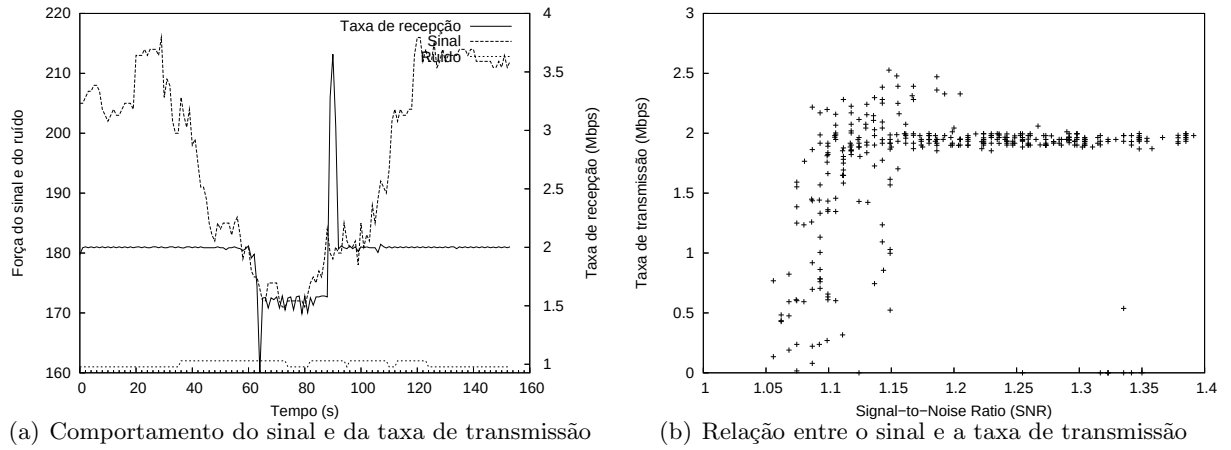


Figura 5.2: Qualidade do sinal e taxa de transmissão na presença de mobilidade.

degradação do sinal em função da mobilidade. Quando o ponto de acesso escolhe transmitir usando uma modulação mais robusta e, portanto, de menor capacidade de vazão, a taxa de recepção de 2 Mbps já não é mais alcançada pela unidade móvel. Usando como indicador da qualidade do sinal a razão entre a força do sinal e do ruído (*Signal-to-Noise Ratio* ou SNR), a Figura 5.2(b) oferece uma visualização da relação entre a qualidade do sinal e a taxa de recepção efetivamente alcançada pela unidade móvel.

O gráfico apresentado na Figura 5.2(b) pode ser dividido em três partes: a primeira, à direita, onde o valor de SNR é alto e a recepção normalmente alcança 2 Mbps. A segunda, ao centro, onde a taxa de recepção é mais instável, mas, em média, ainda alcança 2 Mbps. Outra, por fim, à esquerda, onde o valor de SNR é baixo e a taxa de recepção raramente atinge 2 Mbps. Baseados nesses três patamares de operação, podem ser definidos três limiares de adaptação α , β e γ , onde: α é o menor valor de SNR, de tal modo que a taxa de transmissão desejada é alcançada; β é o menor valor de SNR, de forma que a taxa de transmissão é aceitável; e γ é o menor valor de SNR, para o qual ainda se observa transmissão. Por exemplo, analisando-se a Figura 5.2(b), α , β e γ poderiam ser 1.15, 1.1 e 1.05, respectivamente.

5.5.1 Estratégia adaptativa baseada nos limiares de adaptação α , β e γ

Partindo da observação de que não se atingem taxas elevadas de transmissão quando a qualidade do sinal está degradada, os valores de SNR, α , β e γ podem ser usados em um algoritmo adaptativo

da seguinte maneira:

- **nível 0:** se $\text{SNR} \geq \alpha$, a transmissão continua normalmente, pois a transmissão deve estar transcorrendo normalmente;
- **nível 1:** se $\alpha > \text{SNR} \geq \beta$, o servidor deve transmitir de maneira conservadora. Um exemplo de estratégia conservadora é reduzir o tamanho dos pacotes, mas sem reduzir a taxa de transmissão, pois dessa forma reduz-se a probabilidade de perda de pacotes [SHW03, Sch00];
- **nível 2:** se $\beta > \text{SNR} \geq \gamma$, o servidor deve reduzir a taxa de transmissão, pois, como mostra a Figura 5.2(b), nessas condições a transmissão raramente atinge altas taxas;
- **nível 3:** se $\text{SNR} < \gamma$, a transmissão é interrompida até que SNR volte a ser maior ou igual a γ , pois, para valores de SNR menores do que γ , a rede sem fio normalmente estará inoperante.

A estratégia adaptativa é organizada em níveis para que os clientes possam definir os valores de α , β e γ em função da qualidade de sua recepção. Dessa forma, não é preciso reportar ao servidor a qualidade do sinal e os limiares de adaptação; basta enviar o nível de operação. Em situações de uso real do algoritmo, as estratégias apropriadas para cada nível de operação devem ser implementadas em função da aplicação e da natureza dos dados. Por exemplo, para aplicações de transmissão de fluxos contínuos de vídeo, as estratégias poderiam ser: no nível 1, o envio redundante dos pacotes mais importantes ou o uso de códigos corretores; no nível 2, a redução da qualidade de compressão e, conseqüentemente, da taxa de transmissão ou o descarte de quadros; no nível 3, resta pouco a fazer além de interromper a transmissão, mas pode-se ainda tentar notificar o usuário das condições de conexão ou tentar, quando disponíveis, outras formas de conexão.

Os valores de α , β e γ podem ser encontrados de maneira analítica e com intervenção humana, como fizemos em caráter ilustrativo na Seção 5.5, ou podem ainda ser descobertos automaticamente.

Descoberta automática de α , β e γ

A definição automática de γ é trivial; para fazê-lo, basta armazenar o menor valor de SNR dentre os valores de SNR medidos a cada recepção de pacotes. Os valores de α e β podem ser definidos

a partir da correlação entre a taxa de perda de pacotes e o valor de SNR. Podemos definir α e β como os menores valores de SNR, de tal modo que as taxas de perdas de pacotes sejam, por exemplo, inferiores a 1% e 25%, respectivamente. Para fazer essa correlação, a cada segundo de transmissão, medimos a Taxa de Perdas de Pacotes (TPP) e o valor de SNR e atualizamos α e β segundo o pseudocódigo apresentado na Figura 5.3.

```

se  $[(TPP < 1\%) \text{ e } (SNR < \alpha)]$  então
     $\alpha \leftarrow \alpha - [(\alpha - SNR)/16]$ 
fim se
se  $[(TPP < 25\%) \text{ e } (SNR < \beta)]$  então
     $\beta \leftarrow \beta - [(\beta - SNR)/16]$ 
fim se

```

Figura 5.3: Algoritmo para atualização de α e β .

Os valores de α e β são inicializados com os primeiros valores obtidos de SNR. A fim de aumentar a estabilidade do algoritmo, os valores de TPP e SNR utilizados para atualização de α e β são, na verdade, a média das últimas cinco medições de TPP e SNR. Para evitar que casos excepcionais afetem demasiadamente os valores de α e β , a diferença entre os coeficientes e o SNR é amortizada; em nossa implementação, usamos um fator de amortização 16.

Realizamos experimentos com diferentes valores de amortização. Se o fator de amortização for pequeno demais, por exemplo 2, a convergência de α e β é mais rápida, mas, em contrapartida, os valores de α e β absorvem casos excepcionais na transmissão (casos em que valores excepcionalmente baixos de SNR conseguem — mesmo que por curtos espaços de tempo — taxas altas de transmissão). Se o fator de amortização for grande demais, a convergência dos valores de α e β torna-se lenta.

A amortização também poderia ser implementada usando-se, ao invés de apenas um fator de amortização fixo, funções mais complexas — em que a convergência dos valores de α e β fosse rápida e as exceções fossem evitadas. Em nossos experimentos, avaliamos a utilização de diferentes

funções de amortização, testamos funções lineares e não lineares. No entanto, a estratégia de usar um fator de amortização de valor 16 mostrou-se mais estável e de aplicação mais geral, além de ser mais simples.

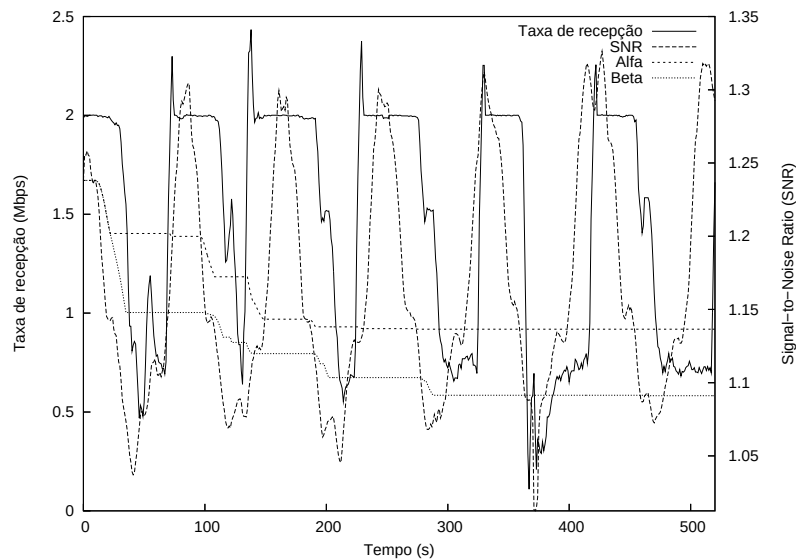


Figura 5.4: Convergência de α e β (TPPs 1% e 25% e fator de amortização 16).

A Figura 5.4 demonstra o processo de busca automática de α e β . Nesta execução, usamos o algoritmo apresentado na Figura 5.3 (com fator de amortização 16) para correlacionar os valores de α e β às taxas de perda de pacotes de 1% e 25%, respectivamente. Na figura, a variação do valor de SNR ilustra a mobilidade presente no experimento, que foi o agente principal das quedas na taxa de recepção. A figura mostra ainda que os valores de α e β convergem para 1,13 e 1,09, respectivamente. Observe que a transmissão sofre quedas bruscas na taxa de recepção e que essas quedas ocorrem — não por acaso — quando o valor de SNR cai ao patamar de α . Repare também que o instante da queda na taxa de transmissão e o instante em que SNR “cruza” α aproximam-se conforme o algoritmo ajusta os valores de α . Em outras palavras, a aproximação de SNR do valor de α é um indício forte de problemas futuros na conexão. Note também que a transmissão normalmente está comprometida quando SNR cai ao patamar de β .

Resta mencionar que, dado que a função usada para atualização de α e β é monotonicamente decrescente, é preciso estabelecer um critério de parada para o algoritmo; caso contrário, a acuidade dos limiares será deteriorada pela — sutil mas contínua — absorção dos casos excepcionais que podem ocorrer ao longo da transmissão. Recomendamos que se use a estabilidade do valor de

α como critério de parada; normalmente pode-se parar o algoritmo quando as atualizações de α forem da ordem de 1% do seu valor absoluto. Usamos apenas α no critério de parada, pois na faixa de operação de β a transmissão é menos estável, o que dificulta a definição de um critério de convergência.

Contudo, apesar da utilidade da determinação automática dos limiares de adaptação, notamos que na maioria das situações práticas os valores para α , β e γ precisam ser determinados apenas uma vez para cada ambiente de execução. Após terem sido determinados esses limiares de adaptação, uma nova busca só se faz necessária caso ocorram mudanças no ambiente, tais como a execução de *handoffs* ou o uso de um NIC diferente.

5.5.2 Implementação

Na implementação da estratégia adaptativa proposta neste trabalho, a unidade móvel monitora os valores de SNR e (a cada segundo) notifica ao servidor o nível de operação (mensagem *Keep Alive*). O servidor, por sua vez, ajusta a taxa de transmissão de acordo com o nível de operação. Para aumentar a estabilidade das adaptações (histerese), o servidor altera o modo de transmissão apenas após receber duas notificações consecutivas contendo o mesmo valor de operação. O mecanismo de troca de mensagens entre o servidor e a unidade móvel é implementado usando-se datagramas UDP. Escolhemos usar UDP e não TCP, como seria natural para mensagens de controle, devido ao comportamento normalmente bloqueante das chamadas TCP.

Entretanto, se $\text{SNR} < \gamma$, como podemos notificar o servidor? Quando o valor de SNR é menor do que γ , a rede provavelmente está inoperante. Sendo assim, é preciso implementar no servidor mecanismos que permitam adaptar os fluxos mesmo durante esses períodos de desconexão. Desse modo, implementamos um mecanismo de *timeout* em que, caso o servidor não receba notificações do cliente durante três segundos, este assume o nível três de operação e interrompe a transmissão, até que uma nova notificação de operação seja recebida.

Em suma, a unidade móvel periodicamente envia datagramas UDP notificando o seu nível de operação e o servidor interrompe a transmissão caso não receba essas notificações. As estratégias de adaptação implementadas foram: no nível de operação 0, o servidor transmite 125 pacotes de

2.000 octetos⁶ por segundo (totalizando 2 Mbps de taxa de transmissão); no nível 1, 250 pacotes de 1.000 octetos por segundo (ainda 2 Mbps); no nível 2, 125 pacotes de 1.000 octetos por segundo (a taxa de transmissão é reduzida para 1 Mbps); e, no nível 3, como dito anteriormente, a transmissão é interrompida.

5.5.3 Ambiente experimental

Os experimentos apresentados neste capítulo foram executados em um ambiente composto por: um computador servidor (Dual-Pentium III, com 2 processadores de 1 GHz e 512 MB de memória; sistema operacional Debian, distribuição Woody), um computador portátil (Pentium III 800 MHz com 128 MB de memória; sistema operacional Debian, distribuição Sarge) e um ponto de acesso IEEE 802.11 (Samsung, modelo SWL-4000 AP). A conectividade entre o computador móvel e o ponto de acesso foi realizada utilizando-se um cartão PCMCIA IEEE 802.11b (Lucent, modelo WaveLAN Silver PC24E-H-FC, *chipset* Orinoco). Também fizemos avaliações utilizando diferentes NICs e pontos de acesso, inclusive avaliamos o algoritmo sobre redes 802.11g, mas, para maior consistência dos resultados, todos os experimentos apresentados neste capítulo foram obtidos usando-se o ambiente experimental acima descrito.

Apesar da dificuldade para realizar experimentos práticos sobre questões de mobilidade, nossos testes foram conduzidos exaustivamente: os resultados apresentados foram obtidos a partir de pelo menos 12 execuções dos algoritmos, sendo que todas as execuções foram consistentes entre si.

5.5.4 Comportamento do algoritmo adaptativo

A Figura 5.5 mostra a comparação entre um fluxo de dados adaptativo e outro não adaptativo. Na execução do algoritmo adaptativo, foram usados valores de $\alpha = 1.13$ e $\beta = 1.09$. A figura mostra que, durante a fase inicial do experimento, aproximadamente até os 30s, ambos os algoritmos alcançam a taxa de recepção de 2 Mbps. Aos 30s, porém, os algoritmos começam a sofrer os efeitos

⁶Como mencionado na Seção 5.3.1, o tamanho dos pacotes, quando possível, não deve superar 1.472 octetos. Entretanto, para gerar fluxos de 2 Mbps, ou mais, tais como os necessários para transmissão de TV Digital, é conveniente utilizar pacotes maiores do que esse valor, caso contrário seria necessário implementar o envio de pacotes em intervalos muito curtos; por essa razão utilizamos pacotes de 2.000 octetos. Além disso, a utilização de valores “redondos” simplifica o entendimento e apresentação dos algoritmos, sem, no entanto, afetar a sua generalidade.

da mobilidade e da queda na qualidade do sinal. Mas, durante essa fase, o algoritmo adaptativo opera uma série de adaptações com o objetivo de reduzir o desperdício da capacidade da rede. Quando a qualidade da conexão volta a ser alta, entretanto, uma questão mostra-se fundamental para a diferença de comportamento entre os algoritmos: o algoritmo adaptativo retoma mais rapidamente as taxas normais de recepção do que o algoritmo não adaptativo.

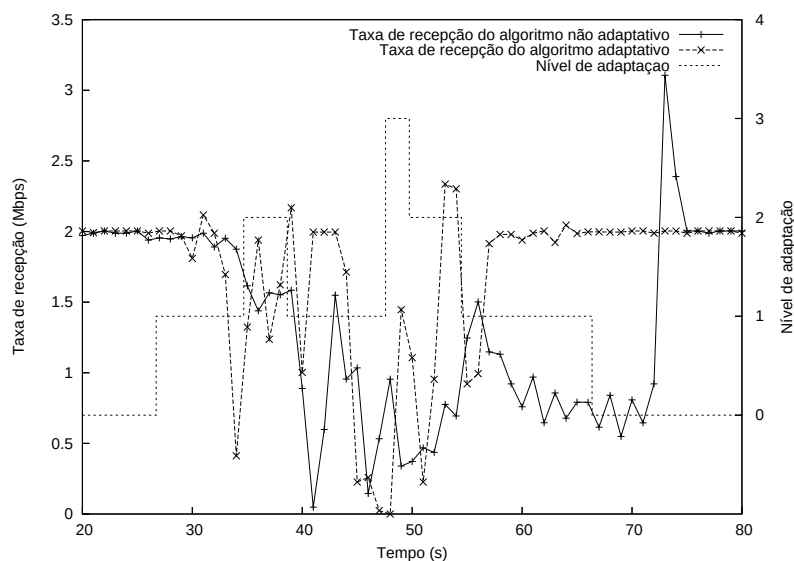


Figura 5.5: Comparação entre o algoritmo adaptativo e o não adaptativo.

O algoritmo adaptativo recupera-se mais rapidamente do que o não adaptativo por duas razões. Primeiro, porque o algoritmo não adaptativo mantém a taxa de transmissão em 2 Mbps e, desse modo, sobrecarrega a capacidade de transmissão do ponto de acesso. Segundo, e mais importante, as adaptações executadas pelo algoritmo adaptativo reduzem as perdas de pacotes, evitando assim a necessidade de atuação do mecanismo ARC e, conseqüentemente, a redução da taxa de transmissão. É importante lembrar que as métricas usadas pelo mecanismo de ARC para adaptação da taxa de transmissão baseiam-se principalmente na qualidade da transmissão e da recepção em termos de perda de quadros e pacotes, número de mensagens ACK perdidas, vazão etc. [HTL⁺05]; e não na qualidade do sinal. A qualidade do sinal é apenas uma métrica auxiliar para o mecanismo de ARC, porque é difícil estabelecer uma relação apenas entre a qualidade do sinal e a respectiva taxa ótima de transmissão; afinal, isso depende também de outros fatores, como a sensibilidade dos NICs usados, o tamanho dos pacotes e o volume de dados trafegados. Assim, ao adaptar os fluxos, mesmo não interferindo diretamente na força do sinal, reduzimos a necessidade de atuação

<i>Algoritmo</i>	<i>Transmitido</i>	<i>Recebido</i>	<i>Taxa Média</i>	<i>Eficiência</i>
Não adaptativo	1.848 Mbits	1.390 Mbits	1,505 Mbps	0,75
Adaptativo	1.659 Mbits	1.604 Mbits	1,736 Mbps	0,96

Tabela 5.1: Comparação entre o algoritmo adaptativo e o não adaptativo.

dos mecanismos de ARC.

Além disso, a fim de evitar excessivas alterações na taxa de transmissão, que poderiam levar à instabilidade do sistema, o mecanismo de ARC aguarda certo período de tempo antes de retomar as taxas mais altas de transmissão [HTL⁺05] e, por essa razão, o fluxo não adaptativo continua sofrendo os efeitos da mobilidade, mesmo já estando próximo ao ponto de acesso.

Análise dos resultados

A Tabela 5.1 resume os resultados obtidos pelos algoritmos adaptativo e não adaptativo. O cliente do algoritmo adaptativo recebeu efetivamente mais dados do que o não adaptativo, mesmo tendo o servidor enviado mais dados para o algoritmo não adaptativo do que para o adaptativo. Em números, o algoritmo adaptativo alcançou 96% de eficiência⁷ contra apenas 75% do algoritmo não adaptativo. Além disso, atingiu taxas médias de recepção 15% maiores do que o algoritmo não adaptativo (1,736 Mbps contra 1,505 Mbps). Ademais, poderíamos traçar cenários em que essa diferença seria ainda mais acentuada, como por exemplo experimentos sujeitos a longos períodos de desconexão.

5.6 Conclusões

Neste capítulo apresentamos os principais problemas relacionados ao desenvolvimento de aplicações adaptativas para redes IEEE 802.11, assim como as técnicas que podem ser usadas para resolver esses problemas.

A contribuição central deste trabalho, contudo, foi a proposta de uma nova estratégia adap-

⁷Eficiência é definida como a quantidade de dados transmitida dividida pela quantidade de dados recebida. Desse modo, a eficiência é tão melhor quanto mais próxima for de 1.

tativa que, ao adequar as transmissões à qualidade do sinal, reduz o desperdício de recursos da rede e, conseqüentemente, aumenta a eficiência das transmissões. Destacamos que o algoritmo adaptativo proposto obteve 96% de eficiência nas transmissões sujeitas à mobilidade, enquanto transmissões não adaptativas atingiram apenas 75% de eficiência. Além disso, a estratégia adaptativa proposta pode ser aplicada em diferentes ambientes de redes sem fio, sendo que grande parte dessa generalidade deve-se ao algoritmo para a descoberta automática dos limiares de adaptação α , β e γ .

Do ponto de vista do desenvolvimento de aplicações, as redes IEEE 802.11, até o presente momento, foram erroneamente tratadas como redes *Ethernet*; isso deve-se principalmente ao fato de que não existiam formas padronizadas de interação entre as aplicações e as interfaces para essas redes. Entretanto, esta pesquisa demonstrou, por meio de implementações concretas, que as peculiaridades das redes IEEE 802.11 podem — e por vezes devem — ser tratadas pelos desenvolvedores na camada de aplicação. Alguns servidores de fluxos já implementam estratégias adaptativas; porém, estas estratégias são inadequadas às redes sem fio. Ademais, este trabalho também foi elaborado com o objetivo de servir como referência para o desenvolvimento de aplicações cada vez mais adequadas às novas realidades trazidas pelas redes sem fio e, em especial, pelas redes IEEE 802.11. Desse modo, esperamos incentivar a pesquisa e o desenvolvimento de técnicas adaptativas para redes IEEE 802.11, cujo aprimoramento nos parece um tema urgente.

Capítulo 6

Pares de pacotes: uma nova métrica de adaptação

*Firmissima est inter pares amicitia.*¹

Curtius Rufus

As aplicações adaptativas são capazes de adequar-se às condições de execução; para isso, monitoram certas métricas de adaptação e, em função delas, executam as devidas adaptações. No entanto, foi visto nos capítulos anteriores que o desenvolvimento de aplicações adaptativas para redes IEEE 802.11 é mais difícil do que para as redes tradicionais. Isso ocorre porque algumas características das redes IEEE 802.11 reduzem a acuidade das métricas de adaptação, tornando sua interpretação mais complexa. Além disso, em redes IEEE 802.11 existem novas métricas de adaptação a serem monitoradas, como por exemplo a qualidade do sinal.

Dentre os fatores que reduzem a acuidade das métricas de adaptação também podemos citar os próprios mecanismos da camada MAC de IEEE 802.11. O mecanismo de retransmissão automática de quadros (*Automatic Repeat reQuest* ou ARQ) oferece o exemplo mais comum de como as métricas de adaptação podem ser negativamente afetadas pelos mecanismos internos de IEEE 802.11. O mecanismo ARQ é responsável por reenviar um quadro caso a entrega deste quadro não

¹Extremamente sólida é a amizade entre os pares.

tenha sido confirmada (*acknowledged*); o número de tentativas de reenvio é definido em função do tamanho do quadro: para os quadros menores do que o *RTS Threshold*, o número padrão de reenvios é sete e, para os quadros maiores, é quatro [Gas02]. Porém, ao fazer uso de retransmissões para melhorar a confiabilidade da comunicação, IEEE 802.11 afeta a acuidade das métricas de adaptação baseadas na informação de perda de pacotes, dificultando assim o uso dessas métricas no desenvolvimento de estratégias adaptativas. Na prática, devido às retransmissões, as perdas de pacotes em redes IEEE 802.11 são perceptíveis apenas quando a qualidade do canal já está bastante deteriorada, ou seja, quando pode ser tarde demais para a realização de adaptações.

O mecanismo para controle automático da taxa de transmissão (*Automatic Rate Control* ou ARC) também tem grande impacto sobre a acuidade das métricas de adaptação [HTL⁺05] e é responsável por escolher a modulação adequada às condições do canal, com o objetivo de explorar o compromisso entre a taxa de erros e a capacidade de transmissão. Seu princípio de atuação é: quanto pior a situação do canal, mais robusta deve ser a modulação utilizada (menor a taxa de transmissão); quanto melhor o canal, menos robusta pode ser a modulação utilizada (maior a taxa de transmissão). O padrão IEEE 802.11g define o uso de até 14 diferentes modulações, que resultam nas taxas de transmissão: 54, 48, 36, 33, 24, 22, 18, 12, 11, 9, 6, 5.5, 2 e 1 Mbps. Porém, o padrão deixa em aberto o algoritmo e os critérios usados pelo mecanismo de ARC para escolher dentre essas possíveis modulações². Como consequência dessa não-definição, os mecanismos de ARC são implementados pelos fabricantes de interfaces de rede IEEE 802.11 de maneira proprietária (“fechada”); isso não seria um problema não fosse o impacto direto do mecanismo de ARC no comportamento da conexão, nas métricas de adaptação e, inclusive, na rede sem fio como um todo.

Do ponto de vista do desenvolvimento de aplicações, uma das principais consequências da atuação do mecanismo de ARC são as variações relativamente bruscas na capacidade de vazão do sistema. Como o uso de diferentes modulações é uma estratégia presente apenas nas redes sem fio — e não nas redes tradicionais —, as aplicações atuais não estão preparadas para essas reduções abruptas na capacidade de vazão. Nem mesmo as aplicações adaptativas levam em consideração as alterações na taxa de transmissão, pois em geral esta informação não está disponível para a camada

²Na página 12 de [IEE03] afirma-se literalmente que a definição do algoritmo usado pelo mecanismo de ARC está além do escopo da padronização.

de aplicação. Grosso modo, o impacto do mecanismo de ARC nas aplicações tem sido negligenciado; na verdade, mais do que isso, as aplicações para redes IEEE 802.11 têm sido tratadas apenas como aplicações para redes *Ethernet*.

Dessa forma, não considerar o impacto do mecanismo de ARC pode fazer com que as métricas de adaptação sejam erroneamente interpretadas. Afinal, as mesmas taxas de perdas de pacotes, ou mesmo a qualidade do sinal, não podem ser interpretadas da mesma maneira, quando obtidas a diferentes taxas de transmissão; o mesmo argumento se aplica também para a interpretação de outras métricas, tais como as baseadas em perda de quadros, atraso, *jitter* etc. Portanto, a taxa de transmissão é uma informação fundamental para o projeto de algoritmos adaptativos para redes IEEE 802.11.

Dada essa necessidade, propomos neste capítulo uma técnica para monitorar o impacto nas taxas de transmissão das alterações feitas pelo mecanismo de ARC; a técnica baseia-se no mecanismo de pares de pacotes e tem como objetivo ser uma métrica de adaptação complementar — quiçá fundamental — para o projeto de aplicações adaptativas para redes IEEE 802.11. Com esse objetivo, a Seção 6.1 descreve o mecanismo de ARC e suas implicações para o projeto de aplicações. A Seção 6.2 apresenta o mecanismo de pares de pacotes e os trabalhos relacionados. A Seção 6.3 descreve o ambiente experimental utilizado. A Seção 6.4 demonstra os resultados obtidos e a utilização do mecanismo de pares de pacotes para a monitoração de alterações na taxa de transmissão. A Seção 6.5 discute algumas limitações à aplicação do mecanismo. Por fim, a Seção 6.6 apresenta nossas conclusões e experiências obtidas na utilização do mecanismo.

6.1 Controle automático da taxa de transmissão

O mecanismo para controle automático da taxa de transmissão (*Automatic Rate Control* ou ARC) é um dos maiores responsáveis pelo sucesso das redes IEEE 802.11, pois permite tanto altas taxas de transmissão, quanto maiores áreas de cobertura. A função do mecanismo é analisar a condição do canal de comunicação e escolher a modulação mais adequada a essa condição.

O uso de modulações mais robustas permite o tráfego de dados mesmo sobre canais deteriorados, ou seja, canais sujeitos a taxas de erro elevadas. Conseqüentemente, essas modulações também permitem o maior alcance das transmissões. Porém, a robustez da modulação é obtida

em detrimento da taxa de transmissão. A Figura 6.1 ilustra a relação entre vazão e o alcance de transmissão de três modulações diferentes³. Em suma, as modulações mais robustas atingem menores taxas de transmissão, mas podem alcançar distâncias maiores, e as modulações menos robustas atingem taxas de transmissão maiores, mas alcançam distâncias menores.

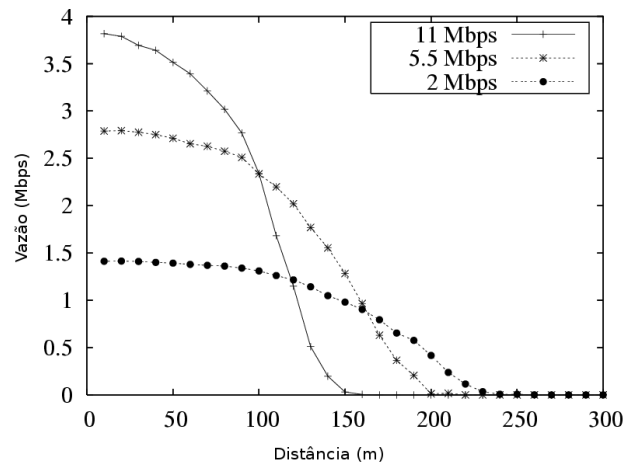


Figura 6.1: Vazão TCP em uma rede IEEE 802.11b (com o mecanismo RTS/CTS habilitado).

O mecanismo de ARC é, portanto, um mecanismo adaptativo capaz de transparentemente adaptar a modulação de transmissão às condições do canal, o que é desejável e por vezes necessário. Essa adaptação, porém, dificulta o desenvolvimento de aplicações adaptativas para redes IEEE 802.11, pois interfere na interpretação das métricas de adaptação e gera distorções indesejáveis no compartilhamento da capacidade da rede sem fio [HRBSD03].

6.1.1 Problemas acarretados pelo uso de ARC

Uma vez que os algoritmos e as métricas de adaptação usados pelo mecanismo ARC não são definidos pelo padrão IEEE 802.11, os fabricantes definem os critérios de adaptação segundo as peculiaridades de seus dispositivos IEEE 802.11. Ou seja, cada interface de rede pode responder de maneira diferente às condições do canal de comunicação, acionando diferentemente os mecanismos de ARC, mesmo quando sujeitos às mesmas condições de sinal e conexão.

O projeto de aplicações adaptativas para redes IEEE 802.11, portanto, deve levar em consi-

³A Figura 6.1 foi gentilmente cedida por Godfrey Tan, ex-doutorando no grupo de Redes e Sistemas Móveis do MIT (<http://nms.lcs.mit.edu>).

deração a diversidade de comportamento nas interfaces de rede. Ou, no mínimo, é preciso que a avaliação das métricas de adaptação considere as taxas de transmissão exercidas, pois, como mencionamos, métricas iguais não podem ser interpretadas da mesma maneira quando obtidas a diferentes taxas de transmissão. Sem a informação de taxa de transmissão, as métricas de adaptação ficam “míopes” e podem não refletir a real condição de uma conexão.

Além disso, apesar da importância da informação de taxa de transmissão para a interpretação das métricas de adaptação, não existem mecanismos para notificar a camada de aplicação das alterações na taxa de transmissão. Atualmente, a forma “natural” de se fazer essa notificação seria através da API de NDIS (*Network Driver Interface Specification*); mas a API não disponibiliza a informação de taxa de transmissão em uso, disponibiliza apenas o conjunto de modulações (taxas de transmissão) que a interface de rede é capaz de usar. Devido à necessidade de melhores métricas de adaptação, foi criado um comitê para a elaboração do padrão IEEE 802.11k (*Radio Resource Measurement*), que deve prover mecanismos mais confiáveis para a avaliação da qualidade do sinal e da conexão sem fio. Mas esse novo padrão não auxiliará os dispositivos atuais e, por ainda estar em desenvolvimento, levará vários anos para chegar aos usuários finais.

Além das consequências para a interpretação das métricas de adaptação, o uso de diferentes modulações pode acarretar algumas anomalias indesejáveis, como a reportada em [HRBSD03]. Em uma rede IEEE 802.11, unidades móveis que transmitem a taxas menores, por exemplo a 1 Mbps, afetam negativamente a capacidade de transmissão das unidades que operam a 11 Mbps. Isso ocorre porque o protocolo MAC de IEEE 802.11 foi projetado para que todas as unidades móveis tenham a mesma probabilidade de acesso ao meio, independente da taxa de transmissão, do tamanho dos pacotes ou da qualidade da conexão. Sendo assim, duas unidades A e B, transmitindo respectivamente a 11 e 1 Mbps, têm a mesma probabilidade de acesso ao meio e, conseqüentemente, a unidade A tem que “esperar” pela transmissão da unidade B. O mecanismo de ARC é o pivô dessa injustiça no compartilhamento da capacidade de transmissão entre unidades móveis [TG05].

Antes de prosseguir e descrever o método de pares de pacotes, cabe questionar: por que realizar adaptações no nível das aplicações, se o mecanismo ARC já o faz? É verdade que o mecanismo ARC possui melhores condições para decidir sobre a adaptação do canal sem fio do que a aplicação,

pois tem pleno acesso à camada MAC e PHY, mas, por outro lado, o mecanismo ARC não possui informações sobre a natureza quer da aplicação e de seus requisitos, quer das possíveis estratégias adaptativas, o que torna desejável a implementação de mecanismos adaptativos na camada de aplicação. Além disso, deve-se observar que as adaptações realizadas pelo mecanismo ARC também são feitas sob certo grau de miopia, pois, para definir a taxa de *transmissão*, o mecanismo de ARC baseia-se principalmente nas estatísticas locais de *recepção* [HTL⁺05], isto é, a estratégia de transmissão é definida em função da recepção, o que pode levar a erros, dado que as transmissões em redes sem fio não são necessariamente simétricas [KNE03, KNG⁺04].

6.2 O mecanismo de pares de pacotes e trabalhos relacionados

O mecanismo de pares de pacotes usado nesse trabalho foi inspirado num mecanismo mais geral, o mecanismo de trem de pacotes, ilustrado na Figura 6.2, que é composto por um emissor e um receptor, tal que ao emissor cabe enviar pacotes a intervalos constantes de tempo e ao receptor, por sua vez, cabe analisar a recepção desses pacotes. De maneira geral, a recepção sofre influência das condições da rede. Em redes sujeitas a condições mais severas, como por exemplo congestionamentos, observam-se maiores variações no intervalo de tempo entre cada recepção de pacotes, assim como a eventual perda de pacotes. Por outro lado, em redes ociosas, as recepções são mais estáveis.

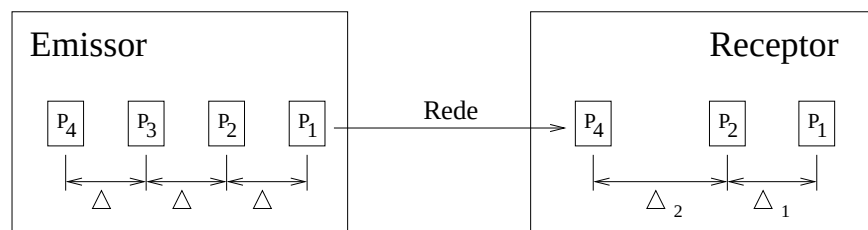


Figura 6.2: Mecanismo de trem de pacotes.

Os mecanismos baseados em trem de pacotes têm sido usados em algoritmos para a inferência das condições de redes desde o final da década de 80. O conceito de análise da recepção de pacotes foi inicialmente apresentado por Van Jacobson [Jac88]. Outro pioneiro foi Keshav [Kes91a], que considerava o intervalo entre pacotes no controle de fluxos; essa abordagem também pode ser

encontrada em [Kes91b] e [Kes94]. Em [Bol93], pequenos datagramas UDP (*probing packets*) eram usados para caracterizar o comportamento fim-a-fim de perdas e atrasos na Internet; esse foi um dos primeiros trabalhos a observar a importância de tais mecanismos para as então emergentes aplicações de áudio e vídeo. O mecanismo também foi utilizado para estimar a capacidade máxima de redes cabeadas [RFdAL03]. Mais recentemente, [DRM04] concluiu um estudo abrangente sobre as técnicas de avaliação de redes baseadas em medidas de dispersão, ou seja, baseadas nos intervalos entre pacotes. Em suma, os trabalhos anteriores constataram que o método de trem de pacotes possui as seguintes características:

- maior estabilidade e precisão do método, quanto maiores forem os pacotes utilizados;
- maior confiabilidade das inferências, quanto maior o número de pacotes analisados;
- adequação do método para inferência da capacidade máxima da rede, mas não para a inferência da utilização da rede.

As razões para as duas primeiras constatações são relativamente intuitivas; afinal, quanto maior a amostra, maior a sua precisão. A razão da terceira constatação é que, para inferir a utilização da rede, é preciso que os pacotes pertencentes ao trem de pacotes sejam atrasados pelo tráfego presente na rede, mas isso ocorre com pouca frequência, pois em geral os pacotes do trem “cabem” nos períodos de transmissão.

Uma vez que não encontramos, na literatura, a exploração do mecanismo em redes IEEE 802.11, recentemente avaliamos o comportamento do mecanismo de trem de pacotes nas redes IEEE 802.11 [dCK04]. A nossa avaliação, em suma, corroborou os resultados anteriormente obtidos para as redes cabeadas, exceto nas situações que envolviam mobilidade, pois esta não é uma característica das redes Ethernet.

Este trabalho, por sua vez, explora a aplicação de uma variação do método de trem de pacotes, o método de pares de pacotes, sobre as redes IEEE 802.11, o qual consiste em enviar pares de pacotes do emissor para o receptor a intervalos regulares; a diferença deste para o mecanismo de trem de pacotes é que os pacotes pertencentes a um par são inseridos na rede sem intervalo entre eles, ou seja, um imediatamente após o outro, como ilustra a Figura 6.3. Para simplicidade do texto, a

diferença de tempo entre o recebimento dos pacotes de um par será chamada de IEP (Intervalo Entre Pacotes). É na avaliação dos IEPs que se baseiam nossos algoritmos para monitoração das taxas de transmissão e para inferência da capacidade da rede.

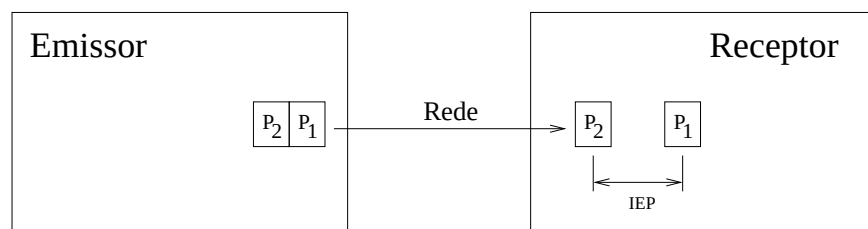


Figura 6.3: Mecanismo de pares de pacotes.

A implementação do mecanismo de pares de pacotes requer algumas decisões de projeto (protocolo de transporte, tamanho dos pacotes, intervalos entre pares de pacotes etc.) que dependem da aplicação. Neste trabalho, usamos como referência para a implementação do mecanismo de pares de pacotes as aplicações de transmissão de fluxos de vídeo (*video streaming*), desse modo, o tamanho dos pacotes e o intervalo entre eles são compatíveis com os utilizados nestas aplicações. Além da crescente importância dessas aplicações nos dias atuais, elas foram usadas como referência porque se adequam bem ao método de pares de pacotes. As aplicações de vídeo, de maneira geral, adotam protocolos construídos sobre o protocolo UDP para a transmissão dos quadros de vídeo [SCFJ03, IL02]. A implementação de pares de pacotes também deve preferir UDP a TCP, pois protocolos de transporte confiáveis valem-se de estratégias de retransmissão e de controle que afetariam a medição dos IEPs. Ademais, protocolos confiáveis omitem a perda de pacotes, o que é uma informação relevante na análise do par de pacotes.

Além disso, o tamanho dos pacotes normalmente usados nas aplicações de vídeo e no mecanismo de pares de pacotes também é compatível. Em [IL02], recomenda-se a transmissão de fluxos contínuos de vídeo MPEG em pacotes de aproximadamente 750 octetos. Em [RFdAL03], constatou-se que pacotes de 500 a 1.000 octetos são adequados aos mecanismos de trem de pacotes. Estes valores também estão abaixo dos limites de fragmentação estipulados para as redes IEEE 802.11 e Ethernet (respectivamente, 2.346 e 1.500 octetos), evitando assim a fragmentação de quadros de vídeo, ou pacotes UDP, em dois ou mais quadros de transmissão da camada MAC. O tamanho dos pacotes usados em nossas avaliações é de 1.000 octetos.

6.3 Ambiente experimental

Para a realização dos experimentos, implementamos o mecanismo de pares de pacotes em linguagem de programação Java, versão 1.5. O código-fonte do protótipo desenvolvido está disponível em <http://gsd.ime.usp.br/software/WirelessMultimedia/sourcecode>. Também usamos as ferramentas *Netperf* e *Etherreal* para validar os resultados obtidos.

O ambiente experimental utilizado é composto de um computador servidor (Dual-Pentium III, com 2 processadores de 1 GHz e 512 MB de memória, sistema operacional Debian-Linux, distribuição Woody) e de dois computadores portáteis da marca Compaq (Pentium III 800 MHz com 128 MB de memória e Pentium III 1 GHz com 256 MB de memória, ambos usando Windows XP, SP2). Para a conectividade sem fio entre os computadores foi usado um ponto de acesso da marca Samsung, modelo SWL-4000 AP. O servidor era conectado à mesma sub-rede do ponto de acesso por cabeamento Fast Ethernet de 100 Mbps. Os computadores portáteis eram conectados ao ponto de acesso através de interfaces IEEE 802.11 de diferentes fabricantes e modelos, tais como: WMP55AG, AIR-PCM340, WUSB54GP Ver. 4, WUSB54AG, WUSB54G Ver. 4, WPC54G Ver. 1.2, SWL-2100N e PC24E-H-FC. Para garantir a consistência dos resultados, todos os experimentos apresentados neste capítulo foram obtidos usando-se esta última interface, um cartão PCMCIA IEEE 802.11b, da Lucent, modelo WaveLAN Silver PC24E-H-FC, *chipset* Orinoco.

6.4 Comportamento dos IEPs mediante mobilidade e suas aplicações práticas

O comportamento dos IEPs em redes IEEE 802.11 é semelhante ao observado nas redes Ethernet tradicionais, exceto nas situações de mobilidade [dCK04]. Esta seção descreve o comportamento dos IEPs e apresenta duas aplicações do mecanismo na monitoração de fluxos em redes IEEE 802.11. Demonstraremos que o mecanismo de pares de pacote permite monitorar, no nível da aplicação, as alterações na taxa de transmissão executadas pelo mecanismo ARC, isto é, permite monitorar qual modulação está sendo utilizada. Além disso, o mecanismo permite inferir a vazão máxima da rede.

A Figura 6.4 apresenta um experimento em que os pares de pacotes são enviados do servidor para

a unidade móvel enquanto a unidade móvel é deslocada dentro da área de cobertura do ponto de acesso. Apresentam-se duas informações: os IEPs e a qualidade do sinal, sendo esta última medida no instante de recebimento do par de pacotes, conforme os métodos descritos em [dCK06a, dCK05]. A qualidade do sinal é apresentada a fim de denotar o deslocamento da unidade móvel em relação ao ponto de acesso — quanto maior a força do sinal, menor a distância para o ponto de acesso e vice-versa. Desse experimento derivam duas observações importantes: que o valor dos IEPs se concentra em patamares (o primeiro patamar entre 1 e 2ms e o segundo em torno de 5ms) e que existe forte correlação entre os patamares e a qualidade do sinal.

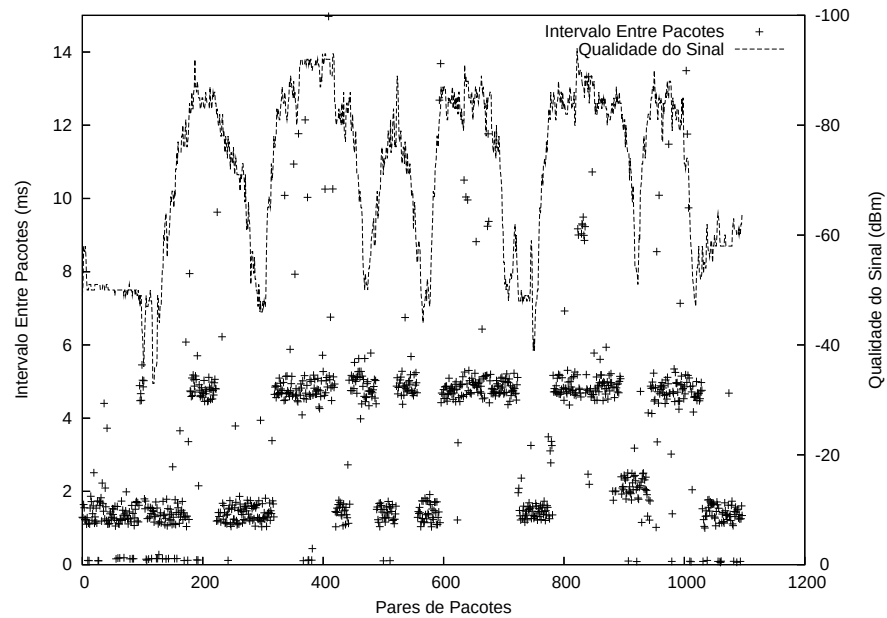


Figura 6.4: Comportamento dos IEPs mediante mobilidade (pacotes de 1000 octetos).

A distribuição dos IEPs em patamares ocorre devido à atuação do mecanismo de ARC, que diminui a taxa de transmissão quando a qualidade da transmissão se deteriora. O patamar inferior corresponde à taxa de transmissão de 11 Mbps e o segundo patamar, à taxa de 5,5 Mbps.

A correlação entre os patamares e a qualidade do sinal, porém, é indireta. Segundo [HTL⁺05], as métricas usadas pelo mecanismo de ARC para decidir alterar a taxa de transmissão baseiam-se em estatísticas sobre a qualidade da transmissão e da recepção, como a taxa de perda de quadros (*Frame Error Rate*), a taxa de erros (*Bit Error Rate*), taxa de mensagens ACK perdidas, vazão etc.; e não na qualidade do sinal, esta é apenas uma métrica de adaptação auxiliar, pois, devido à sua alta variabilidade, é difícil estabelecer uma relação entre o valor da qualidade do sinal e a

respectiva taxa ótima de transmissão. Além disso, diferentes interfaces IEEE 802.11 podem oferecer diferentes valores para a qualidade do sinal, pois as interfaces não oferecem valores calibrados para a qualidade do sinal (*Signal Strength Indication* ou SSI). Observe ainda que, para decidir qual modulação usar, o transmissor precisaria saber a qualidade de sinal medida no receptor. A troca de informações entre receptores e transmissores sobre a qualidade do sinal faz parte da especificação IEEE 802.11h (*Dynamic channel selection and transmission power control*), mas essa especificação aplica-se apenas às redes IEEE 802.11a e, ao que parece, não foi amplamente adotada por implicar adicionais trocas de mensagens de sinalização.

6.4.1 Aplicação 1: monitoração das alterações na taxa de transmissão

Em função dos IEPs podem-se inferir as alterações na taxa de transmissão efetuadas pelos mecanismos de ARC; para isso é preciso determinar os patamares e a qual deles o IEP observado pertence. A Figura 6.5 apresenta a distribuição dos IEPs referentes ao experimento da Figura 6.4. No gráfico pode-se notar que os patamares são suficientemente bem definidos para se projetar um algoritmo capaz de, em função dos IEPs, inferir a taxa de transmissão. Na verdade, mais importante do que inferir a taxa de transmissão, é definir se ela foi reduzida, ou seja, se os IEPs pertencem ou não ao primeiro patamar de recepção. Se os IEPs não pertencem ao primeiro patamar, significa que o mecanismo de ARC identificou problemas na conexão e estabeleceu um modo de transmissão mais robusto.

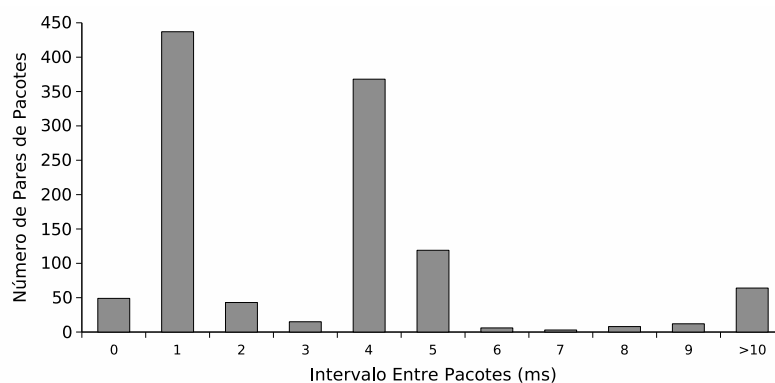


Figura 6.5: Distribuição dos IEPs.

No experimento apresentado na Figura 6.4, o valor médio dos IEPs “pertinentes” ao primeiro

patamar, aqueles abaixo de 4ms, é de 1,43ms (desvio padrão⁴ de 0,61). A partir dessas medidas, que determinam o primeiro patamar, podemos definir uma heurística para encontrar um valor de corte α , tal que um IEP é dito pertencente ao primeiro patamar se o seu valor for menor ou igual a α .

Definimos α como sendo igual ao valor do IEP médio mais duas vezes o valor do desvio padrão; no nosso exemplo, $\alpha = 2,65\text{ms}$. Desse modo, a heurística infere a alteração da taxa de transmissão comparando o valor dos IEPs ao valor de α : se $IEP > \alpha$, então consideramos que a taxa de transmissão foi reduzida pelo mecanismo de ARC; senão, consideramos que não houve alteração. Para garantir a estabilidade das inferências (histerese), levamos em conta três comparações consecutivas, ou seja, consideramos que a taxa de transmissão foi reduzida quando são observados três IEPs consecutivos maiores do que α e consideramos que a taxa de transmissão foi restabelecida quando são observados três IEPs consecutivos menores ou iguais a α . A Figura 6.6 mostra os resultados obtidos pela nossa estratégia; observe que os cortes foram precisos.

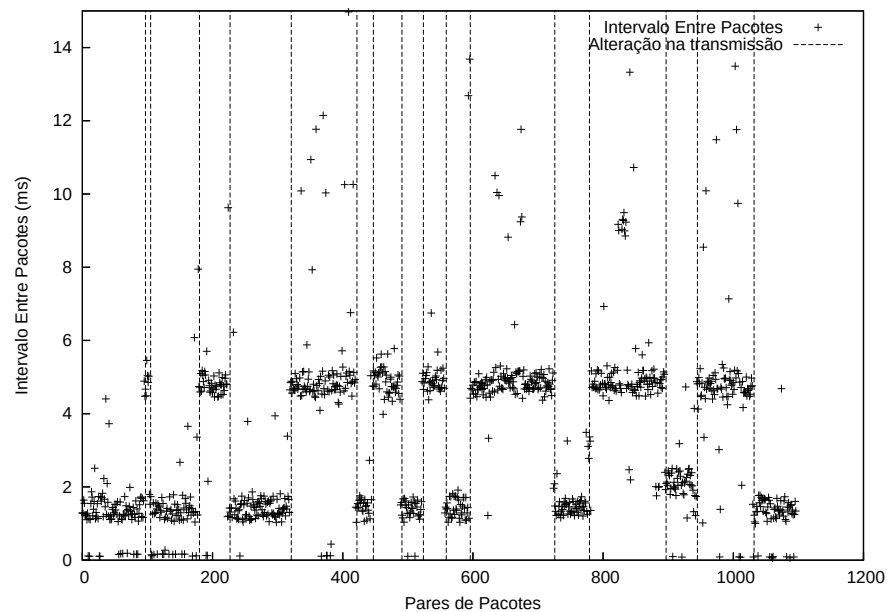


Figura 6.6: Resultados da estratégia para determinação das alterações na taxa de transmissão.

Resta dizer que, em nossos experimentos, não notamos a interferência do mecanismo de pares

⁴O histograma de distribuição apresentado na Figura 6.5 mostra que a distribuição dos IEPs não é uma distribuição normal. Apesar disso, usamos a medida de desvio padrão devido ao fato de que as distribuições para cada patamar assemelham-se a uma distribuição normal.

de pacotes na qualidade das transmissões. Ao comparar fluxos com e sem pares de pacotes, não observamos quaisquer alterações nas taxas de perdas de pacotes, nem nas características dos fluxos, exceto, é claro, a entrega consecutiva dos pacotes pertencentes ao par de pacotes.

6.4.2 Aplicação 2: inferência da capacidade máxima da rede

Além de monitorar as ações dos mecanismos de ARC, o valor médio dos IEPs pode ainda ser usado para calcular uma aproximação da capacidade máxima de vazão da rede. De maneira simplificada, a carga máxima da rede, em Mbps, pode ser estimada através da seguinte fórmula:

$$Carga = (8 * (TamanhoDoPacote)) / (IEP * 1000),$$

onde *TamanhoDoPacote* é o tamanho total dos pacotes, incluindo cabeçalhos, em número de octetos e IEP é o valor médio do IEP em milissegundos.

Aplicando-se a fórmula ao nosso exemplo, em que *TamanhoDoPacote* = 1.000 e *IEP* = 1,43, a estimativa para a vazão máxima comportada pelo ambiente é de 5,59 Mbps. Segundo medições realizadas usando a ferramenta *Netperf*, a capacidade do ambiente 802.11 utilizado é de 5,6 Mbps; ou seja, a estimativa baseada na medição dos IEPs gerou uma aproximação muito boa para a capacidade real do ambiente. Obviamente, existem outros fatores que poderiam ser considerados e incluídos na fórmula acima para aumentar a precisão da estimativa, como as sobrecargas de transmissão dos cabeçalhos das camadas inferiores e a variabilidade das amostras, mas este modelo, apesar de simples, gerou uma ótima estimativa.

A mesma estratégia pode ser usada para estimar a capacidade da rede para o segundo patamar. Em nosso exemplo, o IEP médio do segundo patamar (IEPs maiores do que 2,65ms) é de 4,91ms. Aplicando-se a fórmula definida acima, obtém-se uma capacidade máxima de 1,63 Mbps para o segundo patamar. Essa informação é importante para o projeto de aplicações para ambientes de rede IEEE 802.11, pois nestas redes, diferentemente das redes *Ethernet*, os mecanismos de ARC podem reduzir a taxa de transmissão sem qualquer notificação aos protocolos de nível superior ou à aplicação.

As aplicações para redes IEEE 802.11 devem ser capazes de adaptar-se a reduções bruscas nos níveis de vazão. Uma aplicação de *video streaming* executada em um ambiente IEEE 802.11b pode,

por exemplo, transmitir o vídeo a aproximadamente 5,6 Mbps de taxa real de transmissão, mas essa capacidade de vazão pode ser subitamente reduzida, pela ação do mecanismo de ARC, para aproximadamente 1,63 Mbps.

Se as aplicações executadas sobre redes IEEE 802.11 não forem capazes de se adequar rapidamente às reduções na vazão, não apenas suas funcionalidades podem ser afetadas, mas também a rede como um todo. Nas redes IEEE 802.11, diferentemente de algumas outras tecnologias sem fio, exceder a capacidade de vazão afeta negativa e diretamente todos os usuários do ambiente sem fio [dCK03], pois a capacidade de transmissão é compartilhada entre todos os usuários. Portanto, é importante que as aplicações para redes IEEE 802.11 sejam capazes de monitorar a vazão real a que estão submetidas — e o mecanismo de pares de pacotes é uma alternativa para a implementação desse requisito.

6.5 Limitações do mecanismo de pares de pacotes

A implementação do mecanismo de pares de pacotes requer certos cuidados e possui algumas limitações. O principal deles é o de enviar os pacotes do par um imediatamente após o outro, ou seja, sem atividades adicionais entre essas tarefas; o mesmo procedimento se aplica à recepção dos pacotes.

Outra precaução refere-se à resolução do relógio do sistema; se essa resolução é baixa, por exemplo da ordem de milissegundos, a acuidade da medição dos IEPs fica muito reduzida. Nas aplicações escritas na linguagem Java, por exemplo, ao invés do método-padrão de marcação de tempo `System.currentTimeMillis()`, cuja resolução é da ordem de milissegundos, deve-se usar o método `System.nanoTime()`, cuja resolução é maior e, desse modo, permite que sejam feitas avaliações mais precisas dos IEPs.

A principal limitação para o uso do mecanismo de pares de pacotes é o tamanho dos pacotes. Eles devem ser suficientemente grandes para gerar variações significativas no valor dos IEPs, mas não devem ultrapassar os limites de fragmentação das camadas de rede. Na camada MAC de IEEE 802.11, o tamanho máximo de um quadro de transmissão é 4.096 octetos. Alguns fabricantes, porém, definem o limiar de fragmentação (*Fragmentation Threshold*) em 1.472 octetos (o valor-padrão, segundo o protocolo, é de 2.346 octetos); de modo tal que o limite de fragmentação das

redes Ethernet, que é de 1.500 octetos, também seja atendido [Gas02]. A diferença de 28 octetos entre 1.472 e 1.500 octetos é para conter os cabeçalhos dos protocolos IP e UDP que, juntos, somam 28 octetos. Em suma, o tamanho dos pacotes não deve ultrapassar 1.472 octetos, pois isso implicaria fragmentações, desfragmentações e mensagens ACK adicionais que reduziriam a acuidade dos IEPs.

Certas aplicações, entretanto, usam pacotes bem menores do que 1.472 octetos, ou mesmo do que os 1.000 octetos que usamos em nossos experimentos. A aplicação Skype, por exemplo, usa pacotes UDP de 67 octetos [BS04] para a transmissão de voz (VoIP). Ao usar pacotes pequenos, o tempo de transmissão dos dados (a transmissão de fato) pode ser inferior à sobrecarga inerente à camada MAC de IEEE 802.11, dificultando assim a formação de patamares claros na distribuição dos IEPs. Essa sobrecarga inclui os tempos de espera SIFS e DIFS, o tempo de *backoff*, a transmissão dos cabeçalhos e o tempo de transmissão da mensagem de ACK [Gas02]. Em redes IEEE 802.11b operando a 11 Mbps, a sobrecarga total alcança em média 0,84ms [GK03], podendo chegar a 1,6ms quando operando a taxas menores de transmissão [AAMM01]. Por outro lado, o tempo para transmissão de um pacote UDP com carga útil de 67 octetos é de aproximadamente 0,043ms a 11 Mbps e de 0,085ms a 5,5 Mbps. Portanto, a sobrecarga pode ser facilmente o fator dominante no tempo total da transmissão dos pares de pacotes e, conseqüentemente, do valor dos IEPs.

A Figura 6.7 mostra um experimento em que foram usados pacotes UDP de 67 octetos. A variação nos IEPs é menor, não há a formação clara de múltiplos patamares. O valor médio do IEP é de aproximadamente 0,8ms, mas é difícil definir a ocorrência do segundo patamar. Isso ocorre também porque, para pacotes pequenos, as taxas de erros são menores e, desse modo, a atuação do mecanismo de ARC é menos necessária — lembre-se que a atuação do ARC baseia-se principalmente na perda de pacotes, quadros e mensagens, e não na qualidade do sinal. Sem atuação do ARC, não há alteração das taxas de transmissão, nem formação de patamares.

Ainda sobre o tamanho dos pacotes, resta dizer que o tempo para transmissão dos dados começa a dominar o tempo total da transmissão, sobrepujando as sobrecargas, quando o tamanho dos pacotes atinge cerca de 1.000 octetos, consumindo cerca de 0,715ms a 11 Mbps, o que é necessário para que haja uma clara formação de patamares em caso de alteração da taxa de transmissão.

Outra limitação para o uso do mecanismo de pares de pacotes decorre da diferença de com-

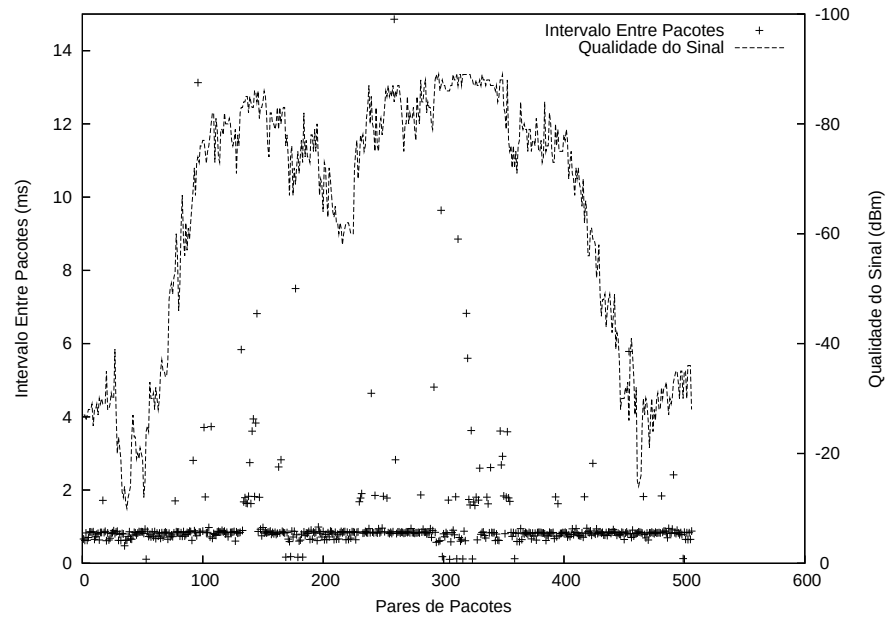


Figura 6.7: Comportamento dos IEPs mediante mobilidade (pacotes de 67 octetos).

portamento entre interfaces IEEE 802.11 de diferentes fabricantes e da compatibilidade entre elas. O mecanismo de pares de pacotes pode não ter exatamente o mesmo comportamento quando utilizadas interfaces de diferentes modelos e fabricantes; as diferenças podem ser sutis, como as decorrentes da diferença entre os algoritmos propostos por cada fabricante para o mecanismo de ARC, mas também podem ser severas a ponto de inviabilizar o uso do mecanismo de pares de pacotes; por exemplo, em nossos experimentos não conseguimos observar a atuação do mecanismo de ARC ao usar a interface Linksys, modelo WPC54G, e o ponto de acesso Samsung, modelo SWL-4000 AP. Aparentemente, apenas conexões de 11 Mbps podem ser estabelecidas entre esses dois produtos. Contudo, acreditamos que, com a evolução dos dispositivos, esses problemas serão mitigados, pois a cada versão os dispositivos IEEE 802.11 têm melhorado sensivelmente, tanto em compatibilidade, quanto em desempenho.

6.6 Conclusões

Neste capítulo, avaliamos empiricamente o comportamento do mecanismo de pares de pacotes em redes IEEE 802.11b e, principalmente, mostramos que ele pode ser usado para monitorar, no nível da aplicação, as alterações na taxa de transmissão feitas pelo mecanismo ARC; a estratégia

apresentada foi capaz de determinar com precisão essas alterações. Além disso, mostramos que o mecanismo de pares de pacotes pode ser usado para inferir a capacidade de vazão das redes IEEE 802.11.

Apesar da importância das informações sobre capacidade de vazão e taxa efetiva de conexão para o projeto de aplicações, não temos conhecimento de trabalhos que disponibilizem essas informações para as camadas de aplicação. Somadas às métricas de adaptação tradicionais, as informações fornecidas pelos pares de pacotes podem resultar em estratégias adaptativas mais adequadas para as redes IEEE 802.11. Além disso, para certas aplicações, como as de transmissão de vídeo, o mecanismo pode ser implementado com pouco ou nenhum prejuízo aos requisitos da aplicação; inclusive sem a adição de novas mensagens de controle, ou seja, os próprios pacotes da aplicação podem ser utilizados para compor os pares de pacotes.

Portanto, o mecanismo de pares de pacotes pode, a baixo custo computacional, melhorar a interpretação das métricas de adaptação, contribuindo assim para o projeto de aplicações cada vez mais adequadas aos novos desafios impostos pelas redes sem fio. Em nossos trabalhos futuros, continuaremos buscando novas métricas e estratégias de adaptação para as redes IEEE 802.11.

Capítulo 7

Conclusões

*There is no chance, no destiny, no fact,
that can circumvent or hinder or control
the firm resolve of a determined soul.*

Ella Wheeler Wilcox

Esta tese, através de uma abordagem predominantemente empírica, analisou o comportamento das transmissões de voz e vídeo sobre redes IEEE 802.11. Dentre os principais problemas identificados figuram:

- longos períodos de desconexão em função dos *handoffs*;
- maior incidência de tráfegos em rajadas;
- baixo isolamento entre unidades móveis, ou seja, a transmissão para uma unidade móvel pode afetar negativamente a transmissão para todas as unidades móveis atendidas por um ponto de acesso;
- menor acuidade das métricas de adaptação;
- substancial variação de comportamento das transmissões mediante o uso de diferentes NICs e configurações;

- comportamento “liga/desliga” em caso de falhas.

Para as aplicações de VoWiFi, as principais restrições são os acréscimos à latência causados pelos *handoffs* e pela incidência de tráfegos em rajadas. Por sua vez, as principais restrições para as aplicações de fluxos de vídeo são as perdas de pacotes causadas pelos *handoffs* e, além disso, o baixo isolamento, pois nessas aplicações, devido ao maior uso de banda passante, as implicações do baixo isolamento para a rede como um todo são bem maiores do que nas aplicações de voz.

Quanto à metodologia de trabalho, destacamos que a adoção de uma abordagem empírica foi essencial para a observação de detalhes que não poderiam ser detectados por meio analítico ou pelo uso de simulações, como, por exemplo, a diferença de comportamento entre interfaces de rede de diferentes fabricantes.

7.1 Contribuições

As principais contribuições desta tese foram:

- caracterização dos principais problemas para a transmissão de voz e vídeo (fluxos H.264/AVC) sobre redes IEEE 802.11 (Capítulos 2, 3 e 4);
- avaliação dos cenários atualmente viáveis para o uso de VoIP sobre redes IEEE 802.11 (Capítulo 3);
- levantamento das funcionalidades que um ponto de acesso IEEE 802.11 deve implementar para a distribuição eficiente de fluxos de vídeo H.264/AVC (Capítulo 4);
- proposta de um novo algoritmo para a utilização da qualidade de sinal como uma métrica de adaptação em aplicações de fluxos de vídeo (Capítulo 5);
- demonstração da viabilidade de uso do mecanismo de pares de pacotes para — no nível da aplicação — monitorar a taxa efetiva de transmissão exercida entre o ponto de acesso e a unidade móvel (Capítulo 6).

7.2 Cenários e trabalhos futuros

As redes IEEE 802.11, em breve, deixarão de ser uma novidade tecnológica para tornarem-se “mais um” meio de acesso à Internet — se não o mais utilizado. Neste cenário, as aplicações multimídia certamente também terão o seu lugar. Assim, as redes sem fio devem tornar-se ubíquas e as aplicações de voz e vídeo deverão ser executadas transparentemente sobre elas; para que isso ocorra, porém, a partir do ponto em que chegamos nesta pesquisa, identificamos que os seguintes trabalhos ainda precisam ser realizados:

1. desenvolvimento de interfaces e mecanismos adaptativos para, durante o transcorrer das sessões de VoWiFi, migrar gentilmente em função das métricas de adaptação do modo conversacional para o modo *Push-To-Talk* (P2T);
2. identificação exata das razões para ocorrência dos tráfegos em rajadas;
3. definição de uma arquitetura padronizada, modular e configurável para os MANEs, tal que esses elementos de rede possam ser configurados a fim de conter apenas os módulos (*plug-ins*) que lhes forem necessários;
4. validação em outras tecnologias de rede sem fio do algoritmo proposto no Capítulo 5 para definição automática dos limiares de adaptação. A validação desse algoritmo em redes IEEE 802.11n pode ser particularmente interessante, dado que essas redes possuem múltiplas antenas e, portanto, múltiplos níveis de sinal;
5. pesquisa de novas métricas de adaptação, especialmente de métricas capazes de determinar o uso total da rede, pois esta é uma informação ainda não disponível em redes IEEE 802.11. Uma das possíveis abordagens para obtenção dessas métricas seria a contagem, a partir da camada MAC, dos *slots* livres de transmissão. Quanto maior o número de *slots* livres, maior a ociosidade da rede, exceto pela existência de *hidden nodes*.

Além desses itens, que são as extensões naturais deste trabalho, também consideramos inevitável o projeto de um novo protocolo de transporte que, em complemento aos protocolos UDP e TCP, sejam mais adequado para a manipulação transparente de múltiplas conexões e das peculiaridades

das redes sem fio. Esta nova abstração poderia também ser encapsulada em uma camada de *middleware*.

7.3 Perspectivas

Para além dos problemas levantados, das nossas contribuições e das extensões propostas, dentre as mensagens principais desta tese está a importância de considerar as redes sem fio no processo de desenvolvimento de aplicações. Mostramos, por meio de avaliações práticas, que é possível melhorar efetivamente a eficiência dos ambientes IEEE 802.11 observando suas características e limitações. Assim, esperamos que esta pesquisa seja útil aos interessados em desenvolver aplicações para redes sem fio. É verdade que, anteriormente, o desenvolvimento para redes sem fio era uma tarefa complexa, principalmente porque não existiam métodos padronizados para o controle das conexões. Entretanto, atualmente, existem APIs e ferramentas para isso.

Em outras palavras, não há mais razão para continuarmos usando as redes IEEE 802.11 da mesma maneira que usamos as redes Ethernet; assim como não podemos mais considerar o uso de redes sem fio algo casual — uma exceção. A tendência é clara: as redes sem fio, quer IEEE 802.11, quer padrões que ainda estão por vir, devem estar cada vez mais presentes em nosso cotidiano; devemos, portanto, sempre considerá-las.

Referências Bibliográficas

- [AAMM01] M.G. Arranz, R. Agüero, L. Muñoz, and P. Mähönen. Behavior of UDP-based applications over IEEE 802.11 wireless networks. In *12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, volume 2, pages F72–F77, San Diego, USA, September 2001.
- [ACKL98] Oguz Angin, Andrew T. Campbell, Michael E. Kounavis, and Raymond R.-F. Liao. The mobeware toolkit: programmable support for adaptive mobile networking. *IEEE Personal Communications Magazine*, 5(4):32–43, August 1998. Special Issue on Adaptive Mobile Systems.
- [BAPW04] Paramvir Bahl, Atul Adya, Jitendra Padhye, and Alec Walman. Reconsidering wireless systems with multiple radios. *ACM SIGCOMM Computer Communication Review*, 34(5):39–46, October 2004.
- [BBC⁺05] Pierpaolo Baccichet, Daniele Bagni, Antonio Chimienti, Luca Pezzoni, and Fabrizio S. Rovati. Frame concealment for H.264/AVC decoders. *IEEE Transactions on Consumer Electronics*, 51(1):227–233, February 2005.
- [BCD⁺97] G. S. Blair, G. Coulson, N. Davies, P. Robin, and T. Fitzpatrick. Adaptive middleware for mobile multimedia applications. In *7th International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV)*, Kobe, Japan, May 1997.

- [Bej02] Yigal Bejerano. Efficient integration of multi-hop wireless and wired networks with QoS constraints. In *8th International Conference on Mobile Computing and Networking (MOBICOM)*, Atlanta, USA, September 2002.
- [Blu03] Bluetooth SIG. Bluetooth specifications, 2003. <http://www.bluetooth.com>.
- [Bol93] Jean-Chrysostome Bolot. Characterizing end-to-end packet delay and loss in the internet. *Journal of High Speed Networks*, 2(3):305–323, December 1993.
- [BPSK97] Hari Balakrishnan, Venkata N. Padmanabhan, Srinivasan Seshan, and Randy H. Katz. A comparison of mechanisms for improving TCP performance over wireless links. *IEEE/ACM Transactions on Networking (TON)*, 5(6):756–769, 1997.
- [BS04] Salman A. Baset and Henning Schulzrinne. An analysis of the Skype peer-to-peer Internet telephony protocol. Technical Report CUCS-039-04, Department of Computer Science, University of Columbia, New York, USA, September 2004.
- [CBB04] Ranveer Chandra, Victor Bahl, and Pradeep Bahl. MultiNet: connecting to multiple IEEE 802.11 networks using a single wireless card. In *23rd Conference on Computer Communications (INFOCOM)*, volume 2, pages 882–893, Hong Kong, China, March 2004.
- [Cis02] Cisco Systems. Understanding delay in packet voice networks. White Paper, September 2002. Available at <http://www.cisco.com/warp/public/788/voip/delay-details.pdf>. Last visit in July 2005.
- [CM03] Carlos Miguel Tavares Calafate and Manuel Pérez Malumbres. Evaluation of the H.264 codec. Technical Report DISCA/60-2003, Polytechnic University of Valencia, Valencia, Spain, 2003.
- [CMM06] Carlos T. Calafate, Manuel P. Malumbres, and Pietro Manzoni. Performance issues of H.264 compressed video streams over IEEE 802.11b based MANETs. *To appear in International Journal of Wireless and Mobile Computing (IJWMC)*, 2006. Special issue on Wireless Ad Hoc Networking.

- [dCdR05] Celso Barbosa de Carvalho and José Ferreira de Rezende. Definindo critérios no compartilhamento de banda de redes 802.11b sob degradação de performance. In *23rd Brazilian Symposium on Computer Networks (SBRC)*, Fortaleza-CE, Brazil, May 2005.
- [dCK03] Arlindo Flávio da Conceição and Fabio Kon. Adaptação de fluxos contínuos UDP sobre redes IEEE 802.11b. In *5o Workshop de comunicação sem fio e computação móvel (WCSF)*, pages 91–101, São Lourenço-MG, Brazil, October 2003.
- [dCK04] Arlindo Flávio da Conceição and Fabio Kon. O uso do mecanismo de pares de pacotes sobre redes IEEE 802.11b. Technical Report RT-MAC-2004-08, Departamento de Ciência da Computação, IME-USP, São Paulo, Brazil, October 2004.
- [dCK05] Arlindo Flávio da Conceição and Fabio Kon. Adaptive streaming based on IEEE 802.11 signal quality. In *23rd Brazilian Symposium on Computer Networks (SBRC)*, pages 1147–1150, Fortaleza-CE, Brazil, May 2005. Short paper.
- [dCK06a] Arlindo F. da Conceição and Fabio Kon. Desenvolvimento de aplicações adaptativas para redes IEEE 802.11. In *24th Brazilian Symposium on Computer Networks (SBRC)*, pages 329–344, Curitiba-PR, Brazil, May 2006.
- [dCK06b] Arlindo F. da Conceição and Fabio Kon. O uso de pares de pacotes para monitoração da taxa de transmissão e da capacidade de vazão em redes IEEE 802.11. In *24th Brazilian Symposium on Computer Networks (SBRC)*, pages 345–358, Curitiba-PR, Brazil, May 2006.
- [dCLFK06] Arlindo F. da Conceição, Jin Li, Dinei A. Florêncio, and Fabio Kon. Is IEEE 802.11 ready for VoIP? In *8th International Workshop on Multimedia Signal Processing (IEEE MMSP)*, Victoria, Canada, October 2006.
- [dMdRJ05] Guilherme M. de Moura, José F. de Rezende, and Guilherme D. G. Jaime. Mecanismos de fragmentação adaptativa em redes 802.11. In *23rd Brazilian Symposium on Computer Networks (SBRC)*, Fortaleza-CE, Brazil, May 2005.

- [DRM04] Constantinos Dovrolis, Parameswaran Ramanathan, and David Moore. Packet-dispersion techniques and a capacity-estimation methodology. *IEEE/ACM Transactions on Networking (TON)*, 12(6):963–977, December 2004.
- [dSOD⁺01] Daniela A. S. dos Santos, M. Oliveira, R. Duarte, A. A. Loureiro, G. R. Mateus, and B. Ribeiro-Neto. Exibição adaptativa de vídeo MPEG para ambientes móveis. In *3rd Workshop de Comunicação sem Fio e Computação Móvel (WCSF)*, Recife-PE, Brasil, October 2001.
- [EY05] Minoru Etoh and Takeshi Yoshimura. Advances in wireless video delivery. *Proceedings of the IEEE*, 93(1):111–122, January 2005.
- [Fli03] Rob Flickenger. *Building Wireless Community Networks*. O’Reilly Media, 2003.
- [FMS01] Scott Fluhrer, Itsik Mantin, and Adi Shamir. Weaknesses in the key scheduling algorithm of RC4. In *8th Annual Workshop on Selected Areas in Cryptography*, Toronto, Canada, August 2001.
- [Gas02] Matthew S. Gast. *802.11 Wireless Networks. The definitive guide*. O’Reilly Media, 2002.
- [GC04] Gregor Gaertner and Vinny Cahill. Understanding link quality in 802.11 mobile ad hoc networks. *IEEE Internet Computing*, 8(1):55–60, January-february 2004.
- [GCR03] N. Golmie, N. Chevrollier, and O. Rebala. Bluetooth and WLAN coexistence: challenges and solutions. *IEEE Wireless Communications Magazine*, 10(6):22–29, December 2003.
- [GK03] Sachin Garg and Martin Kappes. An experimental study of throughput for UDP and VoIP traffic in IEEE 802.11b networks. In *IEEE Wireless Communications and Network Conference (WCNC)*, pages 1748–1753, New Orleans, USA, March 2003.
- [Gup05] Rajarshi Gupta. *Quality of Service in Ad-Hoc Networks*. PhD thesis, Department of Electrical Engineering and Computer Science, University of California, Berkeley, USA, May 2005.

- [HJ98] M. Handley and V. Jacobson. RFC 2327 - SDP: Session description protocol, April 1998.
- [HM05] Changhua He and John C. Mitchell. Security analysis and improvements for IEEE 802.11i. In *The 12th Annual Network and Distributed System Security Symposium (NDSS)*, pages 90–110, San Diego, USA, February 2005.
- [HRBSD03] Martin Heusse, Franck Rousseau, Gilles Berger-Sabbatel, and Andrzej Duda. Performance anomaly of 802.11b. In *22nd Conference on Computer Communications (INFOCOM)*, volume 2, pages 836–843, San Francisco, USA, March 2003.
- [HTL⁺05] Ivaylo Haratcherev, Jacco Taal, Koen Langendoen, Reginald Lagendijk, and Henk Sips. Automatic IEEE 802.11 rate control for streaming applications. *Wireless Communications and Mobile Computing*, 5:421–427, June 2005. Special issue on Radio Link and Transport Protocol Engineering for Future-Generation Wireless Mobile Data Networks.
- [HV02a] Abramo Hefez and Maria Lúcia T. Villela. *Códigos Corretores de Erros*. Série de Computação e Matemática. IMPA, Rio de Janeiro, Brazil, 2002.
- [HV02b] Dzung Tien Hoang and Jeffrey Scott Vitter. *Efficient algorithms for MPEG video compression*. Wiley, 2002.
- [IEE99] IEEE Working Group for WLAN Standards. IEEE standard for wireless LAN medium access control (MAC) and physical layer (PHY), 1999. <http://grouper.ieee.org/groups/802/11>.
- [IEE03] IEEE Working Group for WLAN Standards. IEEE 802.11g: Further Higher Data Rate Extension in the 2.4 GHz Band, June 2003.
- [IEE04] IEEE Working Group on Broadband Wireless Access Standards. IEEE 802.16 WirelessMAN standard for wireless metropolitan area networks, June 2004. <http://grouper.ieee.org/groups/802/16>.

- [IL02] Praveen Ikkurthy and Miguel A. Labrador. Characterization of MPEG-4 traffic over IEEE 802.11b wireless LANs. In *Annual IEEE Conference on Local Computer Networks (LCN)*, volume 27, pages 421–430, Tampa, USA, November 2002.
- [Int03] International Telecommunications Union. International mobile telecommunications standard 2000 (IMT-2000), 2003. <http://www.itu.int/home/imt.html>.
- [ITU90] ITU-T Recommendation G.764. Voice packetization - packetized voice protocols, December 1990.
- [ITU03] ITU-T Recommendation G.114. One-way transmission time, May 2003.
- [Jac88] Van Jacobson. Congestion avoidance and control. In *Proceedings of the Conference of the Special Interest Group on Data Communication (ACM SIGCOMM)*, Stanford, USA, August 1988.
- [JCG04] J. H. James, Bing Chen, and Laurie Garrison. Implementing VoIP: a voice transmission performance progress report. *IEEE Communication Magazine*, 42(7):36–41, July 2004.
- [Joi03] Joint Video Team. ISO/IEC international standard 14496-10, 2003.
- [Kat94] Randy H. Katz. Adaptation and mobility in wireless information systems. *IEEE Personal Communications Magazine*, 1(1):6–17, 1994.
- [KCN01] Fabio Kon, R. H. Campbell, and Klara Nahrstedt. Using dynamic configuration to manage a scalable multimedia distribution system. *Computer Communications*, 24:105–123, 2001.
- [Kes91a] Srinivasan Keshav. *Congestion Control in Computer Networks*. PhD thesis, Department of Electrical Engineering and Computer Science, University of California, Berkeley, USA, September 1991. Published as UC Berkeley TR-654.
- [Kes91b] Srinivasan Keshav. A control-theoretic approach to flow control. In *Proceedings of the Conference of the Special Interest Group on Data Communication (ACM SIGCOMM)*, pages 3–15, Zürich, Switzerland, September 1991.

- [Kes94] Srinivasan Keshav. Packet-pair flow control. <http://www.cs.cornell.edu/skeshav/doc/94/2-17.ps>, 1994.
- [KNE03] David Kotz, Calvin Newport, and Chip Elliott. The mistaken axioms of wireless-network research. Technical Report TR2003-467, Dartmouth College Computer Science, Hanover, USA, July 2003.
- [KNG⁺04] David Kotz, Calvin Newport, Robert S. Gray, Jason Liu, Uougu Yuan, and Chip Elliott. Experimental evaluation of wireless simulation assumptions. In *7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, pages 78–82, Venezia, Italy, October 2004.
- [LDP⁺04] L-A. Larzon, M. Degermark, S. Pink, L-E. Jonsson (Ed.), and G. Fairhurst (Ed.). RFC 3828 - The Lightweight User Datagram Protocol (UDP-Lite), July 2004.
- [LFG03] Yi J. Liang, Nikolaus Färber, and Bernd Girod. Adaptive playout scheduling and loss concealment for voice communication over IP networks. *IEEE Transactions on Multimedia*, 5(4):532–543, December 2003.
- [LSE03] A. Luthra, G. J. Sullivan, and T. Wiegand (Eds.). Special issue on H.264/AVC. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7), July 2003.
- [MCP03] Liam Murphy, Nicola Cranley, and Philip Perry. User-perceived quality-aware adaptive delivery of MPEG-4 content. In *13rd International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, Monterey, USA, June 2003.
- [Men02] Eyal Menin. *The Streaming Media Handbook*. Prentice Hall Computer Books, 2002.
- [MGCM01] Luis Muñoz, Marta Garcia, Johnny Choque, and Ramón Agüero Petri Mähönen. Optimizing internet flows over IEEE 802.11b wireless local area networks: A performance-enhancing proxy based on forward error correction. *IEEE Communication Magazine*, 39(12):60–67, December 2001.

- [MGFK04] Kamesh Medepalli, Praveen Gopalakrishnan, David Famolari, and Toshikazu Kodama. Voice capacity of IEEE 802.11b, 802.11a and 802.11g wireless LANs. In *IEEE Global Telecommunications Conference (GLOBECOM)*, Dallas, USA, 2004.
- [Mil92] David L. Mills. RFC 1305 - network time protocol (version 3): Specification, implementation and analysis, March 1992.
- [Mil95] David L. Mills. Improved algorithms for synchronizing computer network clocks. *IEEE/ACM Transactions on Networking (TON)*, 3(3):245–254, June 1995.
- [Mov03] Moving Picture Experts Group. The MPEG home page, 2003. <http://mpeg.telecomitalialab.com>.
- [NTT03] NTT DoCoMo. Freedom of multimedia access (FOMA), 2003. <http://foma.nttdocomo.co.jp/english>.
- [OCL05] Ricardo A. Rabelo Oliveira, Rainer P. Couto, and Antonio A. F. Loureiro. Uma API para monitoração e análise de redes 802.11 e Bluetooth. In *23rd Brazilian Symposium on Computer Networks (SBRC)*, pages 563–566, Fortaleza-CE, Brazil, May 2005.
- [PRR⁺03] Saar Pilosof, Ramachandran Ramjee, Danny Raz, Yuval Shavitt, and Prasun Sinha. Understanding TCP fairness over wireless LAN. In *22nd Conference on Computer Communications (INFOCOM)*, volume 2, pages 863–872, San Francisco, USA, March 2003.
- [PTVF02] William H. Press (Ed.), Saul A. Teukolsky (Ed.), William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C++: the Art of Scientific Computing*. Cambridge University Press, 2 edition, February 2002.
- [Qua03] Qualcomm. CDMA2000 1xEV-DO, 2003. <http://www.qualcomm.com/cdma/1xEV/>.
- [Rey03] Janice Reynolds. *Going Wi-Fi: a practical guide to planning and building an 802.11 network*. CPM Books, 2003.

- [RFdAL03] Valter Roesler, Peter Max Finzsch, Maiko de Andrade, and José Valdeni Lima. Análise do mecanismo de pares de pacotes visando estimar a banda da rede via UDP. In *21st Brazilian Symposium on Computer Networks (SBRC)*, Natal-RN, Brazil, May 2003.
- [RS05] Ishwar Ramani and Stefan Savage. SyncScan: Practical fast handoff for 802.11 infrastructure networks. In *24th Conference on Computer Communications (INFOCOM)*, Miami, USA, March 2005.
- [RSC⁺02] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler. RFC 3261 - SIP: Session initiation protocol, June 2002.
- [SCFJ03] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. RTP: A Transport Protocol for Real-Time Applications. RFC 3550, July 2003.
- [Sch00] J. Schiller. *Mobile Communications*. Addison Wesley, 2000.
- [SdMBC04] André A. Santana and Tereza C. de M. B. Carvalho. Proposta para otimização de desempenho do protocolo TCP em redes sem fio 802.11. In *22nd Brazilian Symposium on Computer Networks (SBRC)*, pages 563–566, Gramado-RS, Brazil, May 2004. Short paper.
- [SHW03] Thomas Stockhammer, Miska M. Hannuksela, and Thomas Wiegand. H.264/AVC video in wireless environments. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7):657–673, July 2003.
- [SP04] Rishi Sinha and Christos Papadopoulos. An adaptive multiple retransmission technique for continuous media streams. In *14th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, pages 16–21, Kinsale, Ireland, May 2004.
- [SRC84] J. H. Saltzer, D. P. Reed, and D. D. Clark. End-to-end arguments in system design. *ACM Transactions on Computer Systems (TOCS)*, 2(4):277–288, November 1984.
- [SRL98] H. Schulzrinne, A. Rao, and R. Lanphier. RFC 2326 - real time streaming protocol (RTSP), April 1998.

- [Sta05] IEEE Working Group for WLAN Standards. IEEE 802.11e: MAC enhancements (QoS), July 2005.
- [TG05] Godfrey Tan and John Guttag. The 802.11 MAC protocol leads to inefficient equilibria. In *24th Conference on Computer Communications (INFOCOM)*, Miami, USA, March 2005.
- [TXE⁺05] Shu Tao, Kuai Xu, Antonio Estepa, Teng Fei, Lixin Gao, Roch Guerin, Jim Kurose, Don Towsley, and Zhi-Li Zhang. Improving VoIP quality through path switching. In *24th Conference on Computer Communications (INFOCOM)*, Miami, USA, March 2005.
- [VcFRJ01] Bobby Vandalore, Wu chi Feng, and Sonia Fahmy Raj Jain. A survey of application layer techniques for adaptive streaming of multimedia. *Real-Time Imaging*, 7(3):221–235, 2001.
- [VK04] Héctor Velayos and Gunnar Karlsson. Techniques to reduce IEEE 802.11b handoff time. In *IEEE International Conference on Communications (ICC)*, Paris, France, June 2004.
- [Wen03] Stephan Wenger. H.264/AVC video over IP. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7):645–656, July 2003.
- [WHS⁺05] S. Wenger, M. M. Hannuksela, T. Stockhammer, M. Westerlund, and D. Singer. RFC 3984 - RTP payload format for H.264 video, February 2005.
- [WHV⁺02] Ye-Kui Wang, Miska M. Hannuksela, Viktor Varsa, Ari Hourunranta, and Moncef Gabbouj. The error concealment feature in the H.26L test model. In *IEEE International Conference on Image Processing (ICIP)*, volume II, pages 729–732, Rochester, USA, September 2002.
- [Wil02] WildPackets. Converting signal strength percentage to dBm values. Available in <http://www.wildpackets.com/elements/whitepapers/Converting-Signal-Strength.pdf>, 2002.

- [WLL05] Wei Wang, Soung Chang Liew, and Victor O. K. Li. Solutions to performance problems in VoIP over a 802.11 wireless LAN. *Transaction on Vehicular Technology*, 54(1):366–384, January 2005.