



REDES SEM FIO E **REDES MÓVEIS**



No mundo da telefonia pode-se dizer que os quinze últimos anos foram os anos dourados da telefonia celular. O número de assinantes de telefones móveis no mundo inteiro aumentou de 34 milhões em 1993 para quase 5,5 bilhões no final de 2011 e, agora, ultrapassa o número de linhas telefônicas convencionais. As muitas vantagens dos telefones celulares são evidentes para todos — em qualquer lugar, a qualquer hora, acesso desimpedido à rede global de telefonia por meio de um equipamento leve e totalmente portátil. Com o advento de notebooks, palmtops, smartphones e a promessa de acesso desimpedido à Internet global de qualquer lugar, a qualquer hora, será que estamos prestes a assistir a uma explosão semelhante da utilização de dispositivos sem fio para acesso à Internet?

Independentemente do crescimento futuro de equipamentos sem fio para Internet, já ficou claro que redes sem fio e os serviços móveis relacionados que elas possibilitam vieram para ficar. Do ponto de vista de rede, os desafios propostos, em particular nas camadas de enlace e de rede, são tão diferentes dos desafios das redes de computadores cabeadas que é necessário um capítulo inteiro (*este capítulo*) devotado ao estudo de redes sem fio e redes móveis.

Iniciaremos este capítulo com uma discussão sobre usuários móveis, enlaces e redes sem fio e sua relação com as redes maiores (normalmente cabeadas) às quais se conectam. Traçaremos uma distinção entre os desafios propostos pela natureza *sem fio* dos enlaces de comunicação nessas redes e pela *mobilidade* que os enlaces sem fio habilitam. Fazer essa importante distinção — entre sem fio e mobilidade — nos permitirá isolar, identificar e dominar melhor os conceitos fundamentais em cada área. Note que, na realidade, há muitos ambientes de rede nos quais os nós são sem fio, mas não são móveis (por exemplo, redes residenciais sem fio ou redes de escritórios compostas por estações de trabalho estacionárias e monitores de grandes dimensões), e que existem formas limitadas de mobilidade que não requerem enlaces sem fio (por exemplo, um profissional que utiliza um notebook em casa, desliga o equipamento e o leva para seu escritório, onde o liga à rede cabeada da empresa em que trabalha). É claro que muitos dos ambientes sem fio mais interessantes são aqueles em que os usuários são sem fio *e também* móveis — por exemplo, um cenário no qual um usuário móvel (digamos, no banco traseiro de um carro) mantém uma chamada de voz sobre IP (VoIP) e várias conexões TCP ativas enquanto corre pela rodovia a 160 km/h. É nesse ponto, em que o sem fio se cruza com a mobilidade, que encontraremos os desafios técnicos mais interessantes!

Começaremos por ilustrar, primeiro, o cenário no qual consideraremos comunicação e mobilidade sem fio — uma rede na qual usuários sem fio (e possivelmente móveis) estão conectados à infraestrutura da rede maior por um enlace sem fio na borda de rede. Então, consideraremos as características desse enlace sem fio na Seção 6.2. Nessa seção, incluímos uma breve introdução ao acesso múltiplo por divisão de código (*code division*

multiple access — CDMA), um protocolo de acesso ao meio compartilhado que é utilizado com frequência em redes sem fio. Na Seção 6.3, estudaremos com certa profundidade os aspectos da camada de enlace do padrão da LAN sem fio IEEE 802.11 (Wi-Fi); também falaremos um pouco sobre Bluetooth e outras redes pessoais sem fio. Na Seção 6.4 daremos uma visão geral do acesso à Internet por telefone celular, incluindo 3G e as tecnologias celulares emergentes 4G, que fornecem acesso à Internet por voz e em alta velocidade. Na Seção 6.5, voltaremos nossa atenção à mobilidade, focalizando os problemas da localização de um usuário móvel, do roteamento até o usuário móvel e da transferência (*hand-off*) do usuário móvel que passa dinamicamente de um ponto de conexão com a rede para outro. Estudaremos como esses serviços de mobilidade são executados no padrão IP móvel e em GSM nas seções 6.6 e 6.7, respectivamente. Por fim, na Seção 6.8 consideraremos o impacto dos enlaces e da mobilidade sem fio sobre protocolos de camada de transporte e aplicações em rede.

6.1 INTRODUÇÃO

A Figura 6.1 mostra o cenário no qual consideraremos os tópicos de comunicação de dados e mobilidade sem fio. Começaremos mantendo nossa discussão dentro de um contexto geral o suficiente para abranger uma ampla faixa de redes, entre elas LANs sem fio (como a IEEE 802.11) e redes celulares (como uma rede 3G); em outras seções, passaremos então para uma discussão mais detalhada de arquiteturas sem fio específicas. Podemos identificar os seguintes elementos em uma rede sem fio:

- *Hospedeiros sem fio.* Como no caso de redes cabeadas (ou com fio), hospedeiros são os equipamentos de sistemas finais que executam aplicações. Um **hospedeiro sem fio** pode ser um notebook, um palmtop, um smartphone ou um computador de mesa. Os hospedeiros em si podem ser móveis ou não.
- *Enlaces sem fio.* Um hospedeiro se conecta a uma estação-base (definida mais adiante) ou a outro hospedeiro sem fio por meio de um **enlace de comunicação sem fio**. Tecnologias diferentes de enlace sem fio têm taxas de transmissão diversas e podem transmitir a distâncias variadas. A Figura 6.2 mostra duas características fundamentais (área de cobertura e taxa de enlace) dos padrões de enlace sem fio mais populares. (A figura serve apenas para dar uma ideia aproximada dessas características. Por exemplo, alguns desses tipos de redes só estão sendo empregados agora, e algumas taxas de enlace podem aumentar ou diminuir além dos valores mostrados, dependendo da distância, condições do canal e do número de usuários na rede sem fio.) Abordaremos esses padrões mais adiante, na primeira metade deste capítulo; consideraremos também outras características de enlaces sem fio (como suas taxas de erros de bit e as causas desses erros) na Seção 6.2.

Na Figura 6.1, enlaces sem fio conectam hospedeiros localizados na borda da rede com a infraestrutura da rede de maior porte. Não podemos nos esquecer de acrescentar que enlaces sem fio às vezes também são utilizados *dentro* de uma rede para conectar roteadores, comutadores e outros equipamentos de rede. Contudo, neste capítulo, focalizaremos a utilização da comunicação sem fio nas bordas da rede, pois é aqui que estão ocorrendo muitos dos desafios técnicos mais interessantes e a maior parte do crescimento.

- *Estação-base.* A **estação-base** é uma parte fundamental da infraestrutura de rede sem fio. Diferentemente dos hospedeiros e enlaces sem fio, uma estação-base não tem nenhuma contraparte óbvia em uma rede cabeada. Uma estação-base é responsável pelo envio e recebimento de dados (por exemplo, pacotes) de e para um hospedeiro sem fio que está associado a ela. Uma estação-base frequentemente será responsável pela coordenação da transmissão de vários hospedeiros sem fio com os quais está associada. Quando dizemos que um hospedeiro sem fio está “associado” a uma estação-base, isso quer dizer que (1) o hospedeiro está dentro do alcance de comunicação sem fio da estação-base e (2) o hospedeiro usa a estação-base para retransmitir dados entre ele (o hospedeiro) e a rede maior. **Torres celulares** em redes celulares e **pontos de acesso** em LANs sem fio 802.11 são exemplos de estações-base.

Na Figura 6.1, a estação-base está conectada à rede maior (isto é, à Internet, à rede corporativa ou residencial, ou à rede telefônica); portanto, ela funciona como uma retransmissora da camada de enlace entre o hospedeiro sem fio e o resto do mundo com o qual o hospedeiro se comunica.

HISTÓRIA

Acesso público Wi-Fi: em breve, em um poste próximo de você?

Pontos de acesso Wi-Fi — locais públicos onde os usuários podem encontrar acesso sem fio 802.11 — estão se tornando cada vez mais comuns em hotéis, aeroportos e cafés ao redor do mundo. A maioria dos *campi* universitários oferece acesso sem fio espalhado por toda a parte, e é difícil encontrar um hotel que não ofereça acesso à Internet sem fio.

Durante a última década, diversas cidades projetaram, implantaram e operaram redes Wi-Fi municipais. A visão de oferecer acesso Wi-Fi por toda a parte para a comunidade como um serviço público (semelhante aos postes de luz) — ajudando a eliminar a exclusão digital por meio do acesso à Internet para todos os cidadãos e a promover o desenvolvimento econômico — é tentadora. Muitas cidades do mundo inteiro, incluindo Filadélfia, Toronto, Hong Kong, Minneapolis, Londres e Auckland, anunciaram planos de prover esse acesso sem fio dentro de todo o município, ou já fizeram isso de formas variadas. O objetivo na Filadélfia foi “transformar a Filadélfia no maior ponto de acesso Wi-Fi do país e ajudar a melhorar a educação, eliminar a exclusão digital, aprimorar o desenvolvimento da região e reduzir os custos do governo”. O ambicioso programa — um acordo entre a cidade, a

Wireless Philadelphia (entidade sem fins lucrativos) e o ISP Earthlink — construiu uma rede operacional de pontos de acesso 802.11b nos braços de poste de iluminação e semáforos, abrangendo 80% da cidade. Porém, questões financeiras e operacionais fizeram a rede ser vendida a um grupo de investidores privados em 2008, que mais tarde a revenderam para a cidade em 2010. Outros centros, como Minneapolis, Toronto, Hong Kong e Auckland, tiveram sucesso com esforços em menor escala.

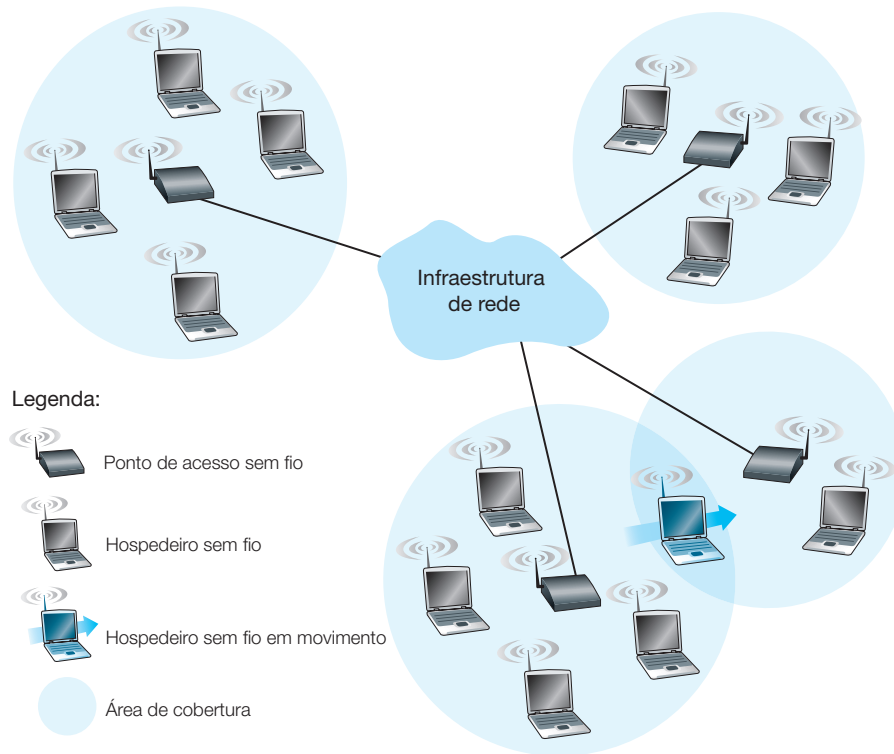
O fato de que redes 802.11 operam no espectro não licenciado (e por isso podem ser realizadas sem a compra dos caríssimos direitos de uso do espectro) parece torná-las financeiramente atraentes. Porém, os pontos de acesso 802.11 (ver Seção 6.3) possuem alcance muito mais curto que as estações-base de celular 3G (ver Seção 6.4), exigindo um maior número de pontos de acesso para cobrir a mesma região geográfica. Por outro lado, as redes de dados por celular, que oferece acesso à Internet, operam no espectro licenciado. As operadoras de celular pagam bilhões de dólares pelos direitos de acesso ao espectro para suas redes, tornando as redes de dados por celular um negócio lucrativo, em vez de um empreendimento municipal.

Quando hospedeiros estão associados com uma estação-base, em geral diz-se que estão operando em **modo de infraestrutura**, já que todos os serviços tradicionais de rede (por exemplo, atribuição de endereço e roteamento) são fornecidos pela rede com a qual estiverem conectados por meio da estação-base. Em **redes *ad hoc***, hospedeiros sem fio não dispõem de qualquer infraestrutura desse tipo com a qual possam se conectar. Na ausência de tal infraestrutura, os próprios hospedeiros devem prover serviços como roteamento, atribuição de endereço, tradução de endereços semelhante ao DNS e outros.

Quando um hospedeiro móvel se desloca para fora da faixa de alcance de uma estação-base e entra na faixa de outra, ele muda seu ponto de conexão com a rede maior (isto é, muda a estação-base com a qual está associado) — um processo denominado **transferência** (*handoff*). Essa mobilidade dá origem a muitas questões desafiadoras. Se um hospedeiro pode se mover, como descobrir sua localização atual na rede de modo que seja possível lhe encaminhar dados? Como é realizado o endereçamento, visto que um hospedeiro pode estar em um dentre muitos locais possíveis? Se o hospedeiro se movimentar *durante* uma conexão TCP ou ligação telefônica, como os dados serão roteados para que a conexão continue sem interrupção? Essas e muitas (mas muitas!) outras questões fazem das redes sem fio e móveis uma área de pesquisa muito interessante sobre redes.

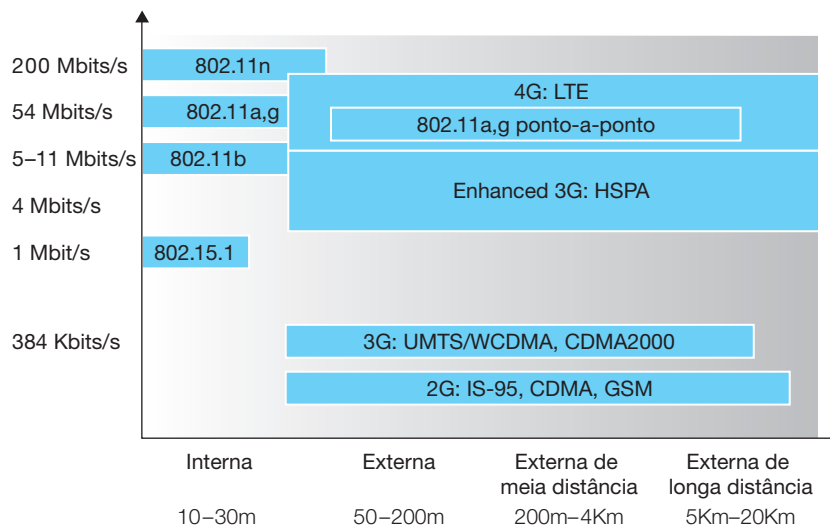
- **Infraestrutura de rede.** É a rede maior com a qual um hospedeiro sem fio pode querer se comunicar.

Após discutir sobre as “partes” da rede sem fio, observamos que essas partes podem ser combinadas de diversas maneiras diferentes para formar diferentes tipos de redes sem fio. Você pode achar uma taxonomia desses tipos de redes sem fio útil ao ler este capítulo, ou ler/aprender mais sobre redes sem fio além deste livro. No nível mais alto,

FIGURA 6.1 ELEMENTOS DE UMA REDE SEM FIO


podemos classificar as redes sem fio de acordo com dois critérios: (i) se um pacote na rede sem fio atravessa exatamente *um salto único sem fio ou múltiplos saltos sem fio*, e (ii) se há *infraestrutura* na rede, como uma estação-base:

- **Salto único, com infraestrutura.** Essas redes têm uma estação-base conectada a uma rede cabeada maior (por exemplo, a Internet). Além disso, toda a comunicação é feita entre a estação-base e um hospedeiro sem fio através de um único salto sem fio. As redes 802.11 que você utiliza na sala de aula, na lanchonete ou na biblioteca; e as redes de dados por celular 3G, que aprenderemos em breve, encaixam-se nesta categoria.

FIGURA 6.2 CARACTERÍSTICAS DE ENLACES DE PADRÕES SELECIONADOS DE REDE SEM FIO


- *Salto único, sem infraestrutura.* Nessas redes, não existe estação-base conectada à rede sem fio. Entretanto, como veremos, um dos nós nessa rede de salto único pode coordenar as transmissões dos outros nós. As redes Bluetooth (que serão estudadas na Seção 6.3.6) e as redes 802.11 no modo *ad hoc* são redes de salto único, sem infraestrutura.
- *Múltiplos saltos, com infraestrutura.* Nessas redes, está presente uma estação-base cabeada para as redes maiores. Entretanto, alguns nós sem fio podem ter que restabelecer sua comunicação através de outros nós sem fio para se comunicarem por meio de uma estação-base. Algumas redes de sensores sem fio e as chamadas redes em malha sem fio se encaixam nesta categoria.
- *Múltiplos saltos, sem infraestrutura.* Não existe estação-base nessas redes, e os nós podem ter de restabelecer mensagens entre diversos outros nós para chegar a um destino. Os nós também podem ser móveis, ocorrendo mudança de conectividade entre eles — uma categoria de redes conhecida como redes móveis *ad hoc* (MANETs). Se os nós móveis forem veículos, essa rede é denominada rede veicular *ad hoc* (VANET). Como você pode imaginar, o desenvolvimento de protocolos para essas redes é desafiador e constitui o assunto de muita pesquisa em andamento.

Neste capítulo, vamos nos limitar às redes de salto único e, depois, principalmente às redes baseadas em infraestrutura.

Agora vamos nos aprofundar um pouco mais nos desafios técnicos que surgem em redes sem fio e móveis. Começaremos considerando, em primeiro lugar, o enlace sem fio individual, deixando nossa discussão sobre mobilidade para outra parte deste capítulo.

6.2 CARACTERÍSTICAS DE ENLACES E REDES SEM FIO

Vamos começar considerando uma rede simples cabeada, por exemplo, uma rede residencial, com hospedeiros interconectados por um comutador Ethernet cabeado (ver Seção 5.4). Se substituíssemos a Ethernet cabeada por uma rede 802.11 sem fio, uma interface de rede sem fio substituiria a interface Ethernet cabeada nos hospedeiros e um ponto de acesso substituiria o comutador Ethernet, mas, na camada de rede ou acima dela, praticamente nenhuma mudança seria necessária. Isso sugere que concentremos nossa atenção na camada de enlace ao procurarmos diferenças importantes entre redes com fio e sem fio. Realmente, podemos encontrar várias diferenças importantes entre um enlace com fio e um enlace sem fio:

- *Redução da força do sinal.* Radiações eletromagnéticas são atenuadas quando atravessam algum tipo de matéria (por exemplo, um sinal de rádio ao atravessar uma parede). O sinal se dispersará mesmo ao ar livre, resultando na redução de sua força (às vezes denominada atenuação de percurso) à medida que aumenta a distância entre emissor e receptor.
- *Interferência de outras fontes.* Várias fontes de rádio transmitindo na mesma banda de frequência sofrerão interferência umas das outras. Por exemplo, telefones sem fio de 2,4 GHz e LANs sem fio 802.11b transmitem na mesma banda de frequência. Assim, o usuário de uma LAN sem fio 802.11b que estiver se comunicando por um telefone sem fio de 2,4 GHz pode esperar que nem a rede nem o telefone funcionem particularmente bem. Além da interferência de fontes transmissoras, o ruído eletromagnético presente no ambiente (por exemplo, um motor ou um equipamento de micro-ondas próximo) pode causar interferência.
- *Propagação multivias.* A propagação multivias (ou multicaminhos) ocorre quando partes da onda eletromagnética se refletem em objetos e no solo e tomam caminhos de comprimentos diferentes entre um emissor e um receptor. Isso resulta no embaralhamento do sinal recebido no destinatário. Objetos que se movimentam entre o emissor e o receptor podem fazer com que a propagação multivias mude ao longo do tempo.

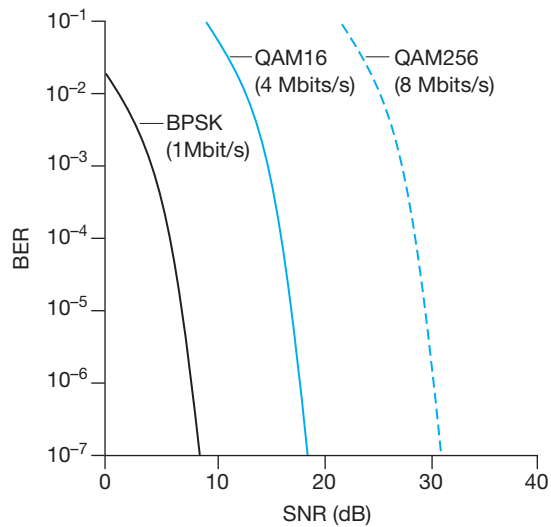
Para obter uma discussão detalhada sobre as características, modelos e medidas do canal sem fio, consulte Anderson [1995].

A discussão anterior sugere que erros de bit serão mais comuns em enlaces sem fio do que em enlaces com fio. Por essa razão, talvez não seja nenhuma surpresa que protocolos de enlace sem fio (como o protocolo 802.11 que examinaremos na seção seguinte) empreguem não só poderosos códigos de detecção de erros por CRC, mas também protocolos de transferência de dados confiável em nível de enlace, que retransmitem quadros corrompidos.

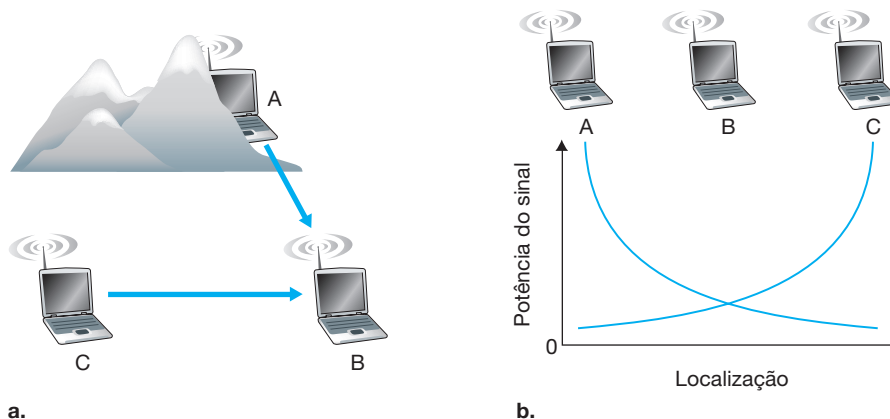
Tendo considerado as falhas que podem ocorrer em um canal sem fio, vamos voltar nossa atenção para o hospedeiro que recebe o sinal sem fio. Esse hospedeiro recebe um sinal eletromagnético que é uma combinação de uma forma degradada do sinal original transmitido pelo remetente (degradada pelos efeitos da atenuação e da propagação multivias, discutidas acima, entre outros) e um ruído de fundo no ambiente. A **relação sinal-ruído** (SNR — *signal-to-noise ratio*) é uma medida relativa da potência do sinal recebido (ou seja, a informação sendo transmitida) e o ruído. A SNR costuma ser calculada em unidades de decibéis (dB), uma unidade de medida que, segundo alguns, é utilizada por engenheiros elétricos principalmente para confundir cientistas da computação. A SNR, medida em dB, é vinte vezes a razão do logaritmo de base 10 da amplitude do sinal recebido à amplitude do ruído. Para nossos fins, precisamos saber apenas que uma SNR maior facilita ainda mais para o destinatário extrair o sinal transmitido de um ruído de fundo.

A Figura 6.3 (adaptada de Holland [2001]) mostra a taxa de erro de bits (BER — *bit error rate*) — em termos simples, a probabilidade de um bit transmitido ser recebido com erro no destinatário — *versus* a SNR para três técnicas de modulação diferentes para codificar informações para a transmissão em um canal sem fio idealizado. A teoria da modulação e da codificação, bem como a extração do sinal e a BER, vai além do escopo deste livro (consulte Schwartz [1980] para obter uma discussão sobre esses assuntos). Não obstante, a Figura 6.3 ilustra diversas características da camada física que são importantes para entender os protocolos de comunicação sem fio da camada superior:

- *Para um determinado esquema de modulação, quanto mais alta for a SNR, mais baixa será a BER.* Visto que um remetente consegue aumentar a SNR elevando sua potência de transmissão, ele pode reduzir a probabilidade de um quadro ser recebido com erro diminuindo tal potência. Observe, entretanto, que há um pequeno ganho prático no aumento da potência além de certo patamar, digamos que para diminuir a BER de 10^{-12} para 10^{-13} . Existem também *desvantagens* associadas com o aumento da potência de transmissão: mais energia deve ser gasta pelo remetente (uma consideração importante para usuários móveis, que utilizam bateria), e as transmissões do remetente têm mais probabilidade de interferir nas transmissões de outro remetente (consulte Figura 6.4(b)).
- *Para determinada SNR, uma técnica de modulação com uma taxa de transmissão de bit maior (com erro ou não) terá uma BER maior.* Por exemplo, na Figura 6.3, com uma SNR de 10 dB, a modulação BPSK com uma taxa de transmissão de 1 Mbit/s possui uma BER menor do que 10^{-7} , enquanto para a modulação QAM16 com uma taxa de transmissão de 4 Mbits/s, a BER é 10^{-1} , longe de ser útil na prática. Entretanto, com uma SNR de 20 dB, a modulação QAM16 possui uma taxa de transmissão de 4 Mbits/s e uma BER de 10^{-7} , enquanto a modulação BPSK possui uma taxa de transmissão de apenas 1 Mbit/s e uma BER tão baixa como estar (literalmente) “fora da parada”. Se é possível suportar uma BER de 10^{-7} , a taxa de transmissão mais alta apresentada pela modulação QAM16 faria desta a técnica de modulação preferida nesta situação. Tais considerações dão origem à característica final, descrita a seguir.
- *A seleção dinâmica da técnica de modulação da camada física pode ser usada para adaptar a técnica de modulação para condições de canal.* A SNR (e, portanto, a BER) pode mudar, como resultado da mobilidade ou em razão das mudanças no ambiente. A modulação adaptativa e a codificação são usadas em sistemas de dados celulares e nas redes de dados Wi-Fi 802.11 e celular 3G, que estudaremos nas Seções 6.3 e 6.4. Isso permite, por exemplo, a seleção de uma técnica de modulação que ofereça a mais alta taxa de transmissão possível sujeita a uma limitação na BER, para as características de determinado canal.

FIGURA 6.3 TAXA DE ERRO DE BITS, TAXA DE TRANSMISSÃO E SNR

Taxas de erros de bits mais altas e que variam com o tempo não são as únicas diferenças entre um enlace com fio e um enlace sem fio. Lembre-se de que, no caso de enlaces de difusão cabeados, cada nó recebe as transmissões de todos os outros nós. No caso de enlaces sem fio, a situação não é tão simples, conforme mostra a Figura 6.4. Suponha que a estação A esteja transmitindo para a estação B. Suponha também que a estação C esteja transmitindo para a estação B. O denominado **problema do terminal oculto**, obstruções físicas presentes no ambiente (por exemplo, uma montanha ou um prédio), pode impedir que A e C escutem as transmissões um do outro, mesmo que as transmissões de A e C interfiram no destino, B. Isso é mostrado na Figura 6.4(a). Um segundo cenário que resulta em colisões que não são detectadas no receptor é causado pelo **desvanecimento** da força de um sinal à medida que se propaga pelo meio sem fio. A Figura 6.4(b) ilustra o caso em que a localização de A e C é tal que as potências de seus sinais não são suficientes para que eles detectem as transmissões um do outro, mas, mesmo assim, *são* fortes o bastante para interferir uma com a outra na estação B. Como veremos na Seção 6.3, o problema do terminal oculto e o desvanecimento tornam o acesso múltiplo em uma rede sem fio consideravelmente mais complexo do que em uma rede cabeada.

FIGURA 6.4 PROBLEMA DO TERMINAL OCULTO (a) E DO DESVANECIMENTO (b)

6.2.1 CDMA

Lembre-se de que dissemos, no Capítulo 5, que, quando hospedeiros se comunicam por um meio compartilhado, é preciso um protocolo para que os sinais enviados por vários emissores não interfiram nos receptores. No mesmo capítulo descrevemos três classes de protocolos de acesso ao meio: de partição de canal, de acesso aleatório e de revezamento. O acesso múltiplo por divisão de código (*code division multiple access* — CDMA) pertence à família de protocolos de partição de canal. Ele predomina em tecnologias de LAN sem fio e celulares. Por ser tão importante no mundo sem fio, examinaremos o CDMA rapidamente agora, antes de passar para tecnologias específicas de acesso sem fio nas próximas seções.

Com um protocolo CDMA, cada bit que está sendo enviado é codificado pela multiplicação do bit por um sinal (o código) que muda a uma velocidade muito maior (conhecida como **taxa de chipping**) do que a sequência original de bits de dados. A Figura 6.5 mostra um cenário simples e idealizado de codificação/decodificação CDMA. Suponha que a velocidade com que bits de dados originais cheguem ao codificador CDMA defina a unidade de tempo; isto é, cada bit original de dados a ser transmitido requer um intervalo de tempo de um bit. Seja d_i o valor do bit de dados para o i -ésimo intervalo de bit. Por conveniência do cálculo matemático, representamos o bit de dados com valor 0 por -1 . Cada intervalo de bit é ainda subdividido em M mini-intervalos. Na Figura 6.5, $M = 8$, embora, na prática, M seja muito maior. O código CDMA usado pelo remetente consiste em uma sequência de M valores, c_m , $m = 1, \dots, M$, cada um assumindo um valor de $+1$ ou -1 . No exemplo da Figura 6.5, o código CDMA de M bits que está sendo usado pelo remetente é $(1, 1, 1, -1, 1, -1, -1, -1)$.

Para ilustrar como o CDMA funciona, vamos focalizar o i -ésimo bit de dados, d_i . Para o m -ésimo mini-intervalo do tempo de transmissão de bits de d_i , a saída do codificador CDMA, $Z_{i,m}$, é o valor de d_i multiplicado pelo m -ésimo bit do código CDMA escolhido, c_m :

$$Z_{i,m} = d_i \cdot c_m \quad (6.1)$$

Se o mundo fosse simples e não houvesse remetentes interferindo, o receptor receberia os bits codificados, $Z_{i,m}$, e recuperaria os bits de dados originais, d_i , calculando:

$$d_i = \frac{1}{M} \sum_{m=1}^M Z_{i,m} \cdot c_m \quad (6.2)$$

Talvez o leitor queira repassar os detalhes do exemplo da Figura 6.5 para verificar se os bits originais de dados são, de fato, corretamente recuperados no receptor usando a Equação 6.2.

No entanto, o mundo está longe de ser ideal e, como mencionamos antes, o CDMA deve funcionar na presença de remetentes que interferem e que estão codificando e transmitindo seus dados usando um código designado diferente. Mas, como um receptor CDMA pode recuperar bits de dados originais de um remetente quando estes estão sendo embaralhados com bits que estão sendo transmitidos por outros remetentes? O CDMA trabalha na hipótese de que os sinais de bits interferentes sendo transmitidos são aditivos. Isso significa, por exemplo, que, se três remetentes enviam um valor 1 e um quarto envia um valor -1 durante o mesmo mini-intervalo, então o sinal recebido em todos os receptores durante o mini-intervalo é 2 (já que $1 + 1 + 1 - 1 = 2$). Na presença de vários remetentes, s calcula suas transmissões codificadas, $Z_{i,m}^s$, exatamente como na Equação 6.1. O valor recebido no receptor durante o m -ésimo mini-intervalo do i -ésimo intervalo de bit, contudo, é agora a *soma* dos bits transmitidos de todos os N remetentes durante o mini-intervalo:

$$Z_{i,m}^* = \sum_{s=1}^N Z_{i,m}^s$$

Surpreendentemente, se os códigos dos remetentes forem escolhidos com cuidado, cada receptor pode recuperar os dados enviados por um dado remetente a partir do sinal agregado apenas usando o código do remetente, como na Equação 6.2:

$$d_i = \frac{1}{M} \sum_{m=1}^M Z_{i,m}^* \cdot c_m \quad (6.3)$$

A Figura 6.6 ilustra um exemplo de CDMA com dois remetentes. O código CDMA de M bits usado pelo remetente que está acima é (1, 1, 1, -1, 1, -1, -1, -1), ao passo que o código CDMA usado pelo que está embaixo é (1, -1, 1, 1, 1, -1, 1, 1). A Figura 6.6 ilustra um receptor recuperando os bits de dados originais do remetente que está acima. Note que o receptor pode extrair os dados do remetente 1, a despeito da transmissão interferente do remetente 2.

Voltando à analogia do coquetel apresentada no Capítulo 5, um protocolo CDMA é semelhante à situação em que os convidados falam vários idiomas; nessa circunstância, os seres humanos até que são bons para manter conversações no idioma que entendem e, ao mesmo tempo, continuar filtrando (rejeitando) outras conversações. Vemos aqui que o CDMA é um protocolo de partição, pois reparte o espaço de código (e não o tempo ou a frequência) e atribui a cada nó uma parcela dedicada do espaço de código.

Nossa discussão do código CDMA aqui é necessariamente breve; na prática, devem ser abordadas inúmeras questões diferentes. Primeiro, para que receptores CDMA consigam extrair o sinal de um emissor qualquer, os códigos CDMA devem ser escolhidos cuidadosamente. Segundo, nossa discussão considerou que as intensidades dos sinais recebidos de vários emissores são as mesmas; na realidade, isso pode ser difícil de conseguir. Existe muita literatura abordando essas e outras questões relativas ao CDMA; veja Pickholtz [1982]; Viterbi [1995], se quiser mais detalhes.

FIGURA 6.5 UM EXEMPLO SIMPLES DE CDMA: CODIFICAÇÃO NO REMETENTE, DECODIFICAÇÃO NO RECEPTOR

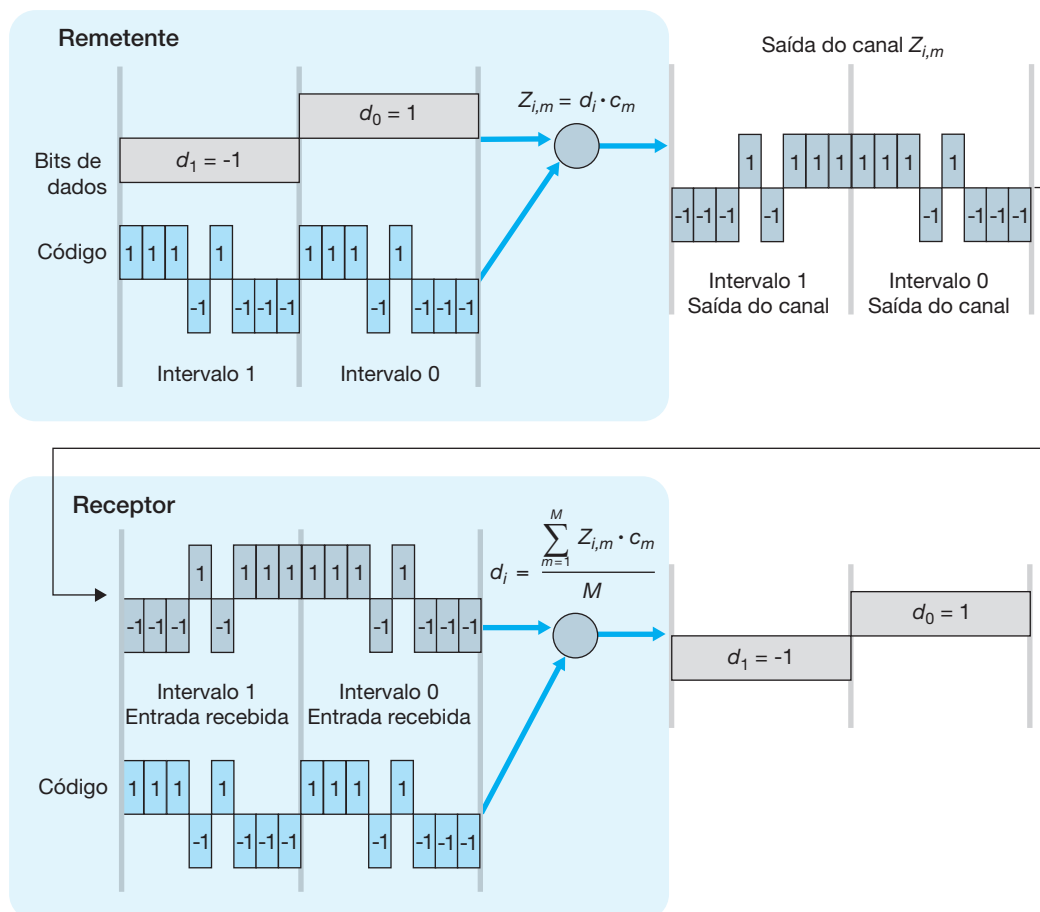
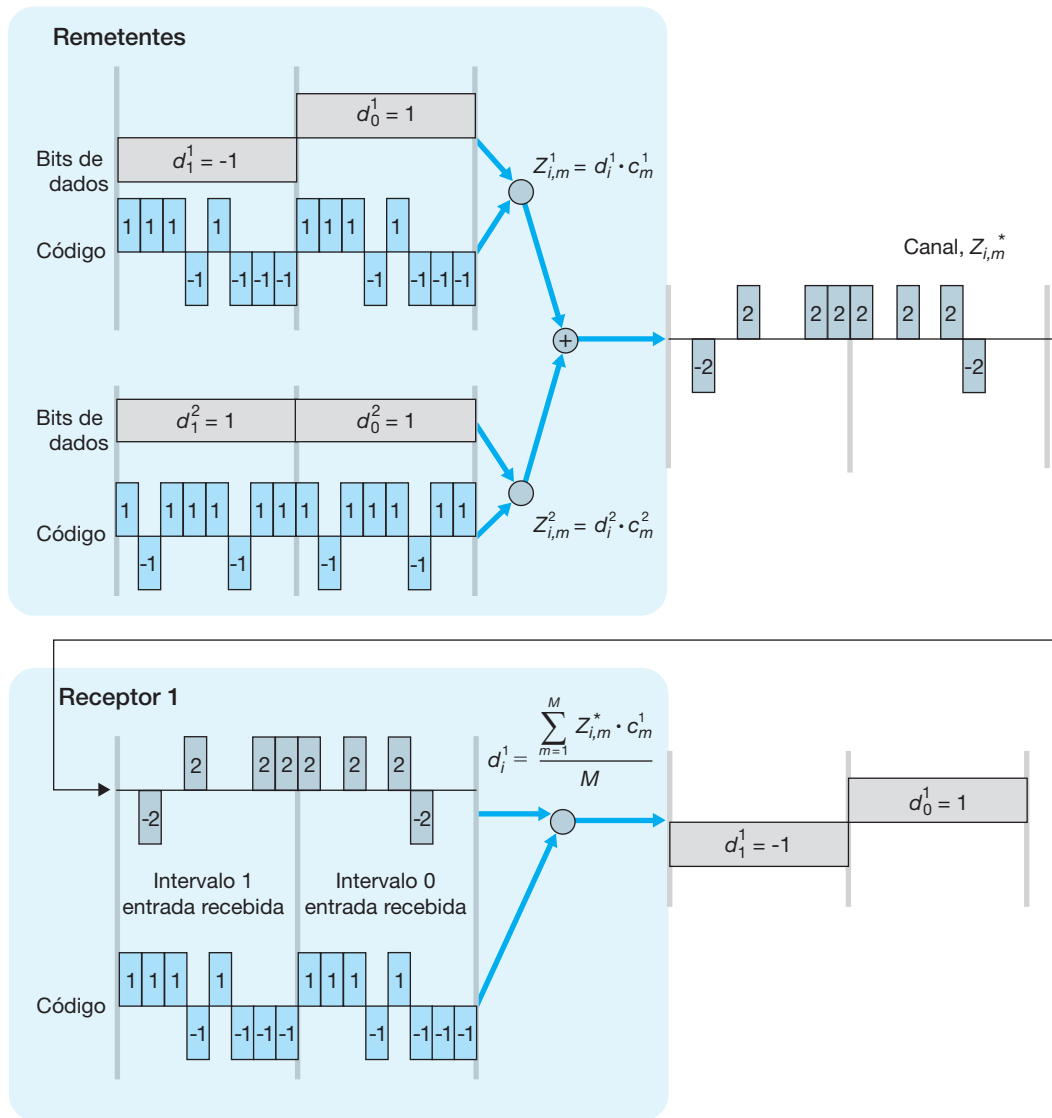


FIGURA 6.6 UM EXEMPLO DE CDMA COM DOIS REMETENTES



6.3 WI-FI: LANS SEM FIO 802.11

Presentes no local de trabalho, em casa, em instituições educacionais, em cafés, aeroportos e esquinas, as LANs sem fio agora são uma das mais importantes tecnologias de rede de acesso na Internet de hoje. Embora muitas tecnologias e padrões para LANs sem fio tenham sido desenvolvidos na década de 1990, uma classe particular de padrões surgiu claramente como a vencedora: a **LAN sem fio IEEE 802.11**, também conhecida como **Wi-Fi**. Nesta seção estudaremos em mais detalhes as LANs sem fio 802.11, examinando a estrutura do quadro 802.11, o protocolo 802.11 de acesso ao meio e a interconexão de LANs 802.11 com LANs Ethernet cabeadas.

Há diversos padrões 802.11 para tecnologia de LAN sem fio, entre eles 802.11b, 802.11a e 802.11g. A Tabela 6.1 apresenta um resumo das principais características desses padrões. 802.11g é, de longe, a tecnologia mais popular. Estão também disponíveis diversos mecanismos de modos duplo (802.11a/g) e triplo (802.11a/b/g).

Os três padrões 802.11 compartilham muitas características. Todos usam o mesmo protocolo de acesso ao meio, CSMA/CA, que discutiremos em breve. Os três também usam a mesma estrutura de quadro para seus quadros de

camada de enlace. Todos os três padrões têm a capacidade de reduzir sua taxa de transmissão para alcançar distâncias maiores. E todos os três padrões permitem “modo de infraestrutura” e “modo *ad hoc*”, como discutiremos em breve. Contudo, conforme mostra a Tabela 6.1, eles apresentam algumas diferenças importantes na camada física.

A LAN sem fio 802.11b tem uma taxa de dados de 11 Mbits/s e opera na faixa de frequência não licenciada de 2,4 a 2,485 GHz, competindo por espectro de frequência com telefones e fornos de micro-ondas de 2,4 GHz. LANs sem fio 802.11a podem funcionar a taxas de bits significativamente mais altas, porém em frequências mais altas. Como operam a uma frequência mais alta, a distância de transmissão dessas LANs é mais curta para determinado nível de potência e elas sofrem mais com a propagação multivias. LANs 802.11g, que operam na mesma faixa de frequência mais baixa das LANs 802.11b e que são compatíveis com a 802.11b (para que se possa atualizar clientes 802.11b de forma incremental), porém com as taxas de transmissão mais altas da 802.11a, devem permitir que os usuários tenham o melhor dos dois mundos.

Um padrão Wi-Fi relativamente novo, 802.11n [IEEE 802.11n, 2012], utiliza antenas de entrada múltipla e saída múltipla (MIMO); ou seja, duas ou mais antenas no lado remetente e duas ou mais antenas no lado destinatário que estão transmitindo/recebendo sinais diferentes [Diggavi, 2004]. Dependendo do esquema de modulação utilizado, é possível alcançar taxas de transmissão de centenas de megabits por segundo com 802.11n.

TABELA 6.1 RESUMO DOS PADRÕES IEEE 802.11

Padrão	Faixa de frequências (EUA)	Taxa de dados
802.11b	2,4–2,485 GHz	até 11 Mbits/s
802.11a	5,1–5,8 GHz	até 54 Mbits/s
802.11g	2,4–2,485 GHz	até 54 Mbits/s

6.3.1 A arquitetura 802.11

A Figura 6.7 ilustra os principais componentes da arquitetura de LAN sem fio 802.11. O bloco de construção fundamental da arquitetura 802.11 é o **conjunto básico de serviço** (*basic service set* — **BSS**). Um BSS contém uma ou mais estações sem fio e uma **estação-base** central, conhecida como um **ponto de acesso** (*access point* — **AP**) na terminologia 802.11. A Figura 6.7 mostra o AP em cada um dos dois BSSs conectando-se a um dispositivo de interconexão (tal como um comutador ou um roteador), que, por sua vez, leva à Internet. Em uma rede residencial típica, há apenas um AP e um roteador (normalmente integrados como uma unidade) que conecta o BSS à Internet.

Como acontece com dispositivos Ethernet, cada estação sem fio 802.11 tem um endereço MAC de 6 bytes que é armazenado no *firmware* do adaptador da estação (isto é, na placa de interface de rede 802.11). Cada AP também tem um endereço MAC para sua interface sem fio. Como na Ethernet, esses endereços MAC são administrados pelo IEEE e são (em teoria) globalmente exclusivos.

Como observamos na Seção 6.1, LANs sem fio que disponibilizam APs em geral são denominadas **LANs sem fio de infraestrutura** e, nesse contexto, “infraestrutura” significa os APs junto com a infraestrutura de Ethernet cabeada que interconecta os APs e um roteador. A Figura 6.8 mostra que estações IEEE 802.11 também podem se agrupar e formar uma rede *ad hoc* — rede sem nenhum controle central e sem nenhuma conexão com o “mundo exterior”. Nesse caso, a rede é formada conforme a necessidade, por dispositivos móveis que, por acaso, estão próximos uns dos outros, têm necessidade de se comunicar e não dispõem de infraestrutura de rede no lugar em que se encontram. Uma rede *ad hoc* pode ser formada quando pessoas que portam notebooks se reúnem (por exemplo, em uma sala de conferências, um trem ou um carro) e querem trocar dados na ausência de um AP centralizado. As redes *ad hoc* estão despertando um interesse extraordinário com a contínua proliferação de equipamentos portáteis que podem se comunicar. Porém, nesta seção, concentraremos nossa atenção em LANs sem fio com infraestrutura.

FIGURA 6.7 A ARQUITETURA DE LAN IEEE 802.11

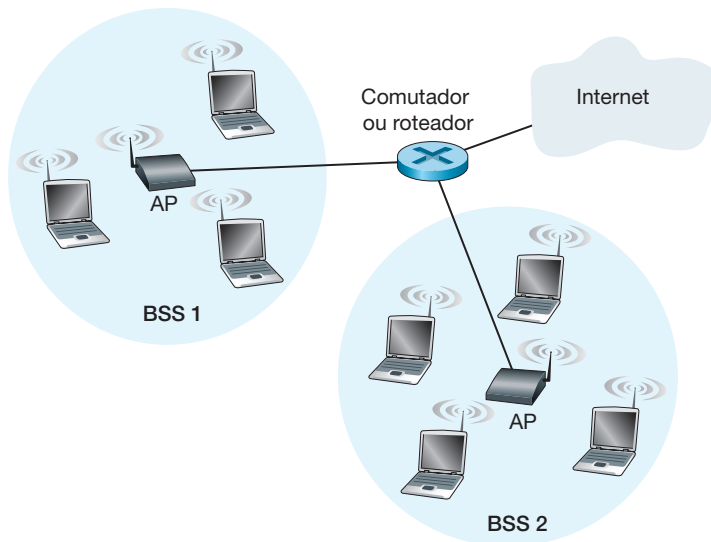
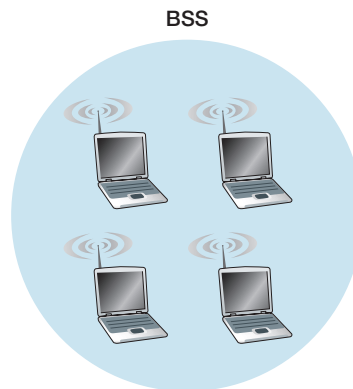


FIGURA 6.8 UMA REDE AD HOC IEEE 802.11



Canais e associação

Em 802.11, cada estação sem fio precisa se associar com um AP antes de poder enviar ou receber dados da camada de rede. Embora todos os padrões 802.11 usem associação, discutiremos esse tópico especificamente no contexto da IEEE 802.11b/g.

Ao instalar um AP, um administrador de rede designa ao ponto de acesso um **Identificador de Conjunto de Serviços** (*Service Set Identifier* — SSID) composto de uma ou duas palavras. (O comando “veja redes disponíveis” no Microsoft Windows XP, por exemplo, apresenta uma lista que mostra o SSID de todos os APs ordenado por faixa.) O administrador também deve designar um número de canal ao AP. Para entender números de canal, lembre-se de que as redes 802.11b operam na faixa de frequência de 2,4 GHz a 2,485 GHz. Dentro dessa faixa de 85 MHz, o padrão 802.11 define 11 canais que se sobrepõem em parte. Não há sobreposição entre quaisquer dois canais se, e somente se, eles estiverem separados por quatro ou mais canais. Em particular, o conjunto dos canais 1, 6 e 11 é o único de três canais não sobrepostos. Isso significa que um administrador poderia criar uma LAN sem fio com uma taxa máxima de transmissão agregada de 33 Mbits/s instalando três APs 802.11b na mesma localização física, designando os canais 1, 6 e 11 aos APs e interconectando cada um desses APs com um comutador.

Agora que já entendemos o básico sobre canais 802.11, vamos descrever uma situação interessante (e que não é completamente fora do comum) — uma selva de Wi-Fis. Uma **selva de Wi-Fis** (Wi-Fi *jungle*) é qualquer localização física na qual uma estação sem fio recebe um sinal suficientemente forte de dois ou mais APs. Por exemplo, em muitos cafés da cidade de Nova York, uma estação sem fio pode captar um sinal de diversos APs próximos. Um deles pode ser o AP gerenciado pelo café, enquanto os outros podem estar localizados em apartamentos vizinhos. Cada ponto de acesso provavelmente estaria localizado em uma sub-rede IP diferente e teria sido designado independentemente a um canal.

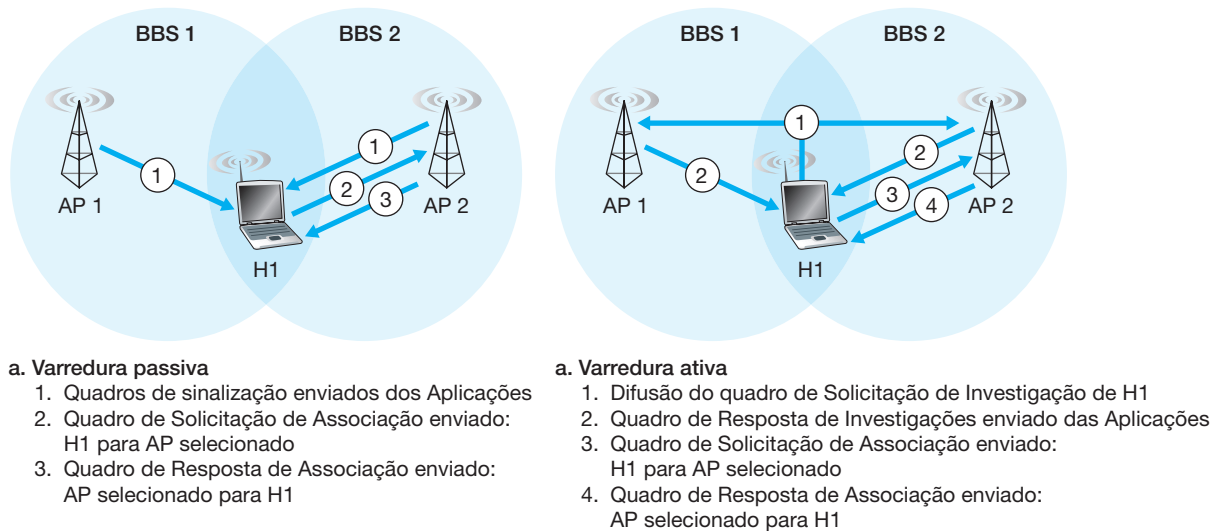
Agora suponha que você entre nessa selva de Wi-Fis com seu computador portátil, em busca de acesso à Internet sem fio e de um cafezinho. Suponha que há cinco APs na selva de Wi-Fis. Para conseguir acesso à Internet, sua estação sem fio terá de se juntar a exatamente uma das sub-redes e, portanto, precisará se **associar** com exatamente um dos APs. Associar significa que a estação sem fio cria um fio virtual entre ela mesma e o AP. De modo específico, só o AP associado enviará quadros de dados (isto é, quadros contendo dados, tal como um datagrama) a sua estação sem fio e esta enviará quadros de dados à Internet apenas por meio do AP associado. Mas como sua estação sem fio se associa com um determinado AP? E, o que é mais fundamental, como sua estação sem fio sabe quais APs estão dentro da selva, se é que há algum?

O padrão 802.11b requer que um AP envie periodicamente **quadros de sinalização**, cada qual incluindo o SSID e o endereço MAC do AP. Sua estação sem fio, sabendo que os APs estão enviando quadros de sinalização, faz uma varredura dos 11 canais em busca de quadros de sinalização de quaisquer APs que possam estar por lá (alguns dos quais talvez estejam transmitindo no mesmo canal — afinal, estamos na selva!). Ao tomar conhecimento dos APs disponíveis por meio dos quadros de sinalização, você (ou seu hospedeiro sem fio) seleciona um desses pontos de acesso para se associar.

O padrão 802.11 não especifica um algoritmo para selecionar com quais dos APs disponíveis se associar; esse algoritmo é de responsabilidade dos projetistas do *firmware* e do software 802.11 em seu hospedeiro sem fio. Em geral, o hospedeiro escolhe o AP cujo quadro de sinalização é recebido com a intensidade de sinal mais alta. Embora uma intensidade alta do sinal seja algo bom (veja, por exemplo, a Figura 6.3), esta não é a única característica do AP que determinará o desempenho que um hospedeiro recebe. Em particular, é possível que o AP selecionado tenha um sinal forte, mas pode ser sobrecarregado com outros hospedeiros associados (que precisarão compartilhar a largura de banda sem fio naquele AP), enquanto um AP não carregado não é selecionado em razão de um sinal levemente mais fraco. Diversas formas alternativas de escolher os APs foram propostas recentemente [Vasudevan, 2005; Nicholson, 2006; Sudaresan, 2006]. Para obter uma discussão interessante e prática de como a intensidade do sinal é medida, consulte Bardwell [2004].

O processo de varrer canais e ouvir quadros de sinalização é conhecido como **varredura passiva** (veja a Figura 6.9(a)). Um hospedeiro sem fio pode também realizar uma **varredura ativa**, transmitindo um quadro de investigação que será recebido por todos os APs dentro de uma faixa do hospedeiro sem fio, como mostrado na Figura 6.9(b). Os APs respondem ao quadro de requisição de investigação com um quadro de resposta de investigação. O hospedeiro sem fio pode, então, escolher o AP com o qual irá se associar dentre os APs que estão respondendo.

Após selecionar o AP ao qual se associará, o hospedeiro sem fio envia um quadro de solicitação de associação ao AP, e este responde com um quadro de resposta de associação. Observe que essa segunda apresentação de solicitação/resposta é necessária com a varredura ativa, visto que um AP de resposta ao quadro de solicitação de investigação inicial não sabe quais dos (possivelmente muitos) APs de resposta o hospedeiro escolherá para se associar, do mesmo modo que um cliente DHCP pode escolher entre servidores múltiplos DHCP (veja a Figura 4.21). Uma vez associado ao AP, o hospedeiro desejará entrar na sub-rede (no sentido do endereçamento IP da Seção 4.4.2) à qual pertence o AP. Assim, o hospedeiro normalmente enviará uma mensagem de descoberta DHCP (veja a Figura 4.21) à sub-rede por meio de um AP a fim de obter um endereço IP na sub-rede. Logo que o endereço é obtido, o resto do mundo, então, vê esse hospedeiro apenas como outro hospedeiro com um endereço IP naquela sub-rede.

FIGURA 6.9 VARREDURA PASSIVA E ATIVA PARA PONTOS DE ACESSO

Para criar uma associação com um determinado AP, a estação sem fio talvez tenha de se autenticar perante o AP. LANs sem fio 802.11 dispõem de várias alternativas para autenticação e acesso. Uma abordagem, usada por muitas empresas, é permitir o acesso a uma rede sem fio com base no endereço MAC de uma estação. Uma segunda abordagem, usada por muitos cafés Internet, emprega nomes de usuários e senhas. Em ambos os casos, o AP em geral se comunica com um servidor de autenticação usando um protocolo como o RADIUS [RFC 2865] ou o DIAMETER [RFC 3588]. Separar o servidor de autenticação do AP permite que um servidor de autenticação atenda a muitos APs, centralizando as decisões de autenticação e acesso (quase sempre delicadas) em um único servidor e mantendo baixos os custos e a complexidade do AP. Veremos, na Seção 8.8, que o novo protocolo IEEE 802.11i, que define aspectos de segurança da família de protocolos 802.11, adota exatamente essa técnica.

6.3.2 O protocolo MAC 802.11

Uma vez associada com um AP, uma estação sem fio pode começar a enviar e receber quadros de dados de e para o ponto de acesso. Porém, como várias estações podem querer transmitir quadros de dados ao mesmo tempo sobre o mesmo canal, é preciso um protocolo de acesso múltiplo para coordenar as transmissões. Aqui, **estação** significa uma estação sem fio ou um AP. Como discutimos no Capítulo 5 e na Seção 6.2.1, em termos gerais, há três classes de protocolos de acesso múltiplo: partição de canal (incluindo CDMA), acesso aleatório e revezamento. Inspirados pelo enorme sucesso da Ethernet e seu protocolo de acesso aleatório, os projetistas do 802.11 escolheram um protocolo de acesso aleatório para as LANs sem fio 802.11. Esse protocolo de acesso aleatório é denominado **CSMA com prevenção de colisão** ou, mais sucintamente, **CSMA/CA**. Do mesmo modo que o CSMA/CD da Ethernet, o “CSMA” de CSMA/CA quer dizer “acesso múltiplo por detecção de portadora”, o que significa que cada estação sonda o canal antes de transmitir e abstém-se de transmitir quando percebe que o canal está ocupado. Embora tanto a Ethernet quanto o 802.11 usem acesso aleatório por detecção de portadora, os dois protocolos MAC apresentam diferenças importantes. Primeiro, em vez de usar detecção de colisão, o 802.11 usa técnicas de prevenção de colisão. Segundo, por causa das taxas relativamente altas de erros de bits em canais sem fio, o 802.11 (ao contrário da Ethernet) usa um esquema de reconhecimento/retransmissão (ARQ) de camada de enlace. Mais adiante descreveremos os esquemas usados pelo 802.11 para prevenção de colisão e reconhecimento na camada de enlace.

Lembre-se de que, nas seções 5.3.2 e 5.4.2, dissemos que, com o algoritmo de detecção de colisão, uma estação Ethernet ouve o canal à medida que transmite. Se, enquanto estiver transmitindo, a estação detectar que alguma outra estação também está, ela abortará sua transmissão e tentará novamente após uma pequena unidade de tempo aleatória. Ao contrário do protocolo Ethernet 802.3, o protocolo MAC 802.11 *não* implementa detecção de colisão. Isso se deve a duas razões importantes:

- A capacidade de detectar colisões exige as capacidades de enviar (o próprio sinal da estação) e de receber (para determinar se alguma outra estação está transmitindo) ao mesmo tempo. Como a potência do sinal recebido em geral é muito pequena em comparação com a potência do sinal transmitido no adaptador 802.11, é caro construir um hardware que possa detectar colisões.
- Mais importante, mesmo que o adaptador pudesse transmitir e ouvir ao mesmo tempo (e, presumivelmente, abortar transmissões quando percebesse um canal ocupado), ainda assim ele não seria capaz de detectar todas as colisões, devido ao problema do terminal escondido e do desvanecimento, como discutimos na Seção 6.2.

Como LANs 802.11 sem fio não usam detecção de colisão, uma vez que uma estação comece a transmitir um quadro, *ela o transmite integralmente*; isto é, tão logo uma estação inicie, não há volta. Como é de se esperar, transmitir quadros inteiros (em particular os longos) quando existe grande possibilidade de colisão pode degradar significativamente o desempenho de um protocolo de acesso múltiplo. Para reduzir a probabilidade de colisões, o 802.11 emprega diversas técnicas de prevenção de colisão, que discutiremos em breve.

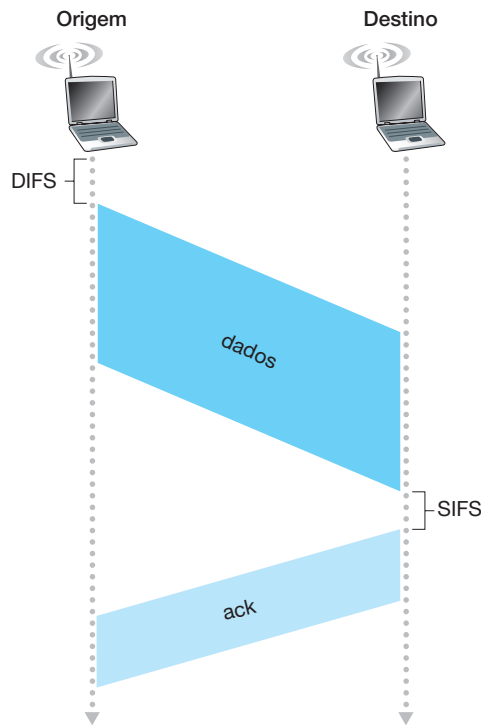
Antes de considerar prevenção de colisão, contudo, primeiro devemos examinar o esquema de **reconhecimento na camada de enlace** do 802.11. Lembre-se de que dissemos, na Seção 6.2, que, quando uma estação em uma LAN sem fio envia um quadro, este talvez não chegue intacto à estação de destino, por diversos motivos. Para lidar com essa probabilidade não desprezível de falha, o protocolo MAC 802.11 usa reconhecimentos de camada de enlace. Como ilustrado na Figura 6.10, quando a estação de destino recebe um quadro que passou na verificação de CRC, ela espera um curto período de tempo, conhecido como **Espaçamento Curto Interquadros** (*Short Inter-Frame Spacing* — SIFS), e então devolve um quadro de reconhecimento. Se a estação transmissora não receber um reconhecimento em dado período de tempo, ela admitirá que ocorreu um erro e retransmitirá o quadro usando de novo o protocolo CSMA/CA para acessar o canal. Se a estação transmissora não receber um reconhecimento após certo número fixo de retransmissões, desistirá e descartará o quadro.

Agora que já discutimos como o 802.11 usa reconhecimentos da camada de enlace, estamos prontos para descrever o protocolo CSMA/CA 802.11. Suponha que uma estação (pode ser uma estação sem fio ou um AP) tenha um quadro para transmitir.

1. Se inicialmente a estação perceber que o canal está ocioso, ela transmitirá seu quadro após um curto período de tempo conhecido como **Espaçamento Interquadros Distribuído** (*Distributed Inter-Frame Space* — DIFS); ver Figura 6.10.
2. Caso contrário, a estação escolherá um valor aleatório de recuo usando o recuo exponencial binário (conforme encontramos na Seção 5.3.2) e fará a contagem regressiva a partir desse valor quando perceber que o canal está ocioso. Se a estação perceber que o canal está ocupado, o valor do contador permanecerá congelado.
3. Quando o contador chegar a zero (note que isso pode ocorrer somente quando a estação percebe que o canal está ocioso), a estação transmitirá o quadro inteiro e então ficará esperando um reconhecimento.
4. Se receber um reconhecimento, a estação transmissora saberá que o quadro foi corretamente recebido na estação de destino. Se a estação tiver outro quadro para transmitir, iniciará o protocolo CSMA/CA na etapa 2. Se não receber um reconhecimento, a estação entrará de novo na fase de recuo na etapa 2 e escolherá um valor aleatório em um intervalo maior.

Lembre-se de que, no protocolo de acesso múltiplo CSMA/CD (Seção 5.5.2), uma estação começa a transmitir tão logo percebe que o canal está ocioso. Com o CSMA/CA, entretanto, a estação priva-se de transmitir

FIGURA 6.10 802.11 USA RECONHECIMENTOS DA CAMADA DE ENLACE

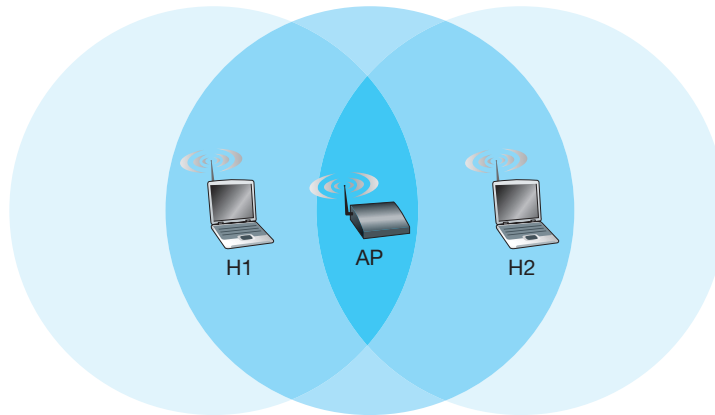


enquanto realiza a contagem regressiva, mesmo quando percebe que o canal está ocioso. Por que o CSMA/CD e o CDMA/CA adotam essas abordagens diferentes aqui?

Para responder a essa pergunta, vamos considerar um cenário com duas estações em que cada uma tem um quadro a transmitir, mas nenhuma transmite imediatamente porque percebe que uma terceira estação já está transmitindo. Com o CSMA/CD da Ethernet, cada uma das duas estações transmitiria tão logo detectasse que a terceira estação terminou de transmitir. Isso causaria uma colisão, o que não é um problema sério em CSMA/CD, já que ambas as estações abortariam suas transmissões e assim evitariam a transmissão inútil do restante dos seus quadros. Entretanto, com 802.11 a situação é bem diferente. Como o 802.11 não detecta uma colisão nem aborta transmissão, um quadro que sofra uma colisão será transmitido integralmente. Assim, a meta do 802.11 é evitar colisões sempre que possível. Com esse protocolo, se duas estações perceberem que o canal está ocupado, ambas entrarão imediatamente em *backoff* aleatório e, esperamos, escolherão valores diferentes de *backoff*. Se esses valores forem, de fato, diferentes, assim que o canal ficar ocioso, uma das duas começará a transmitir antes da outra e (se as duas não estiverem ocultas uma da outra) a “estação perdedora” ouvirá o sinal da “estação vencedora”, interromperá seu contador e não transmitirá até que a estação vencedora tenha concluído sua transmissão. Desse modo é evitada uma colisão dispendiosa. É claro que ainda podem ocorrer colisões com 802.11 nesse cenário: as duas estações podem estar ocultas uma da outra ou podem escolher valores de *backoff* aleatório próximos o bastante para que a transmissão da estação que se inicia primeiro tenha ainda de atingir a segunda. Lembre-se de que já vimos esse problema antes em nossa discussão sobre algoritmos de acesso aleatório no contexto da Figura 5.12.

Tratando de terminais ocultos: RTS e CTS

O protocolo 802.11 MAC também inclui um esquema de reserva inteligente (mas opcional) que ajuda a evitar colisões mesmo na presença de terminais ocultos. Vamos estudar esse esquema no contexto da Figura 6.11, que mostra duas estações sem fio e um ponto de acesso. Ambas as estações estão dentro da faixa do AP (cuja área de cobertura é representada por um círculo sombreado) e ambas se associaram com o AP. Contudo, pelo desvanecimento, as faixas

FIGURA 6.11 EXEMPLO DE TERMINAL OCULTO: H1 ESTÁ OCULTO DE H2, E VICE-VERSA

de sinal de estações sem fio estão limitadas ao interior dos círculos sombreados mostrados na Figura 6.11. Assim, cada uma das estações está oculta da outra, embora nenhuma esteja oculta do AP.

Agora vamos considerar por que terminais ocultos podem ser problemáticos. Suponha que a estação H1 esteja transmitindo um quadro e, a meio caminho da transmissão, a estação H2 queira enviar um quadro para o AP. O H2, que não está ouvindo a transmissão de H1, primeiro esperará um intervalo DIFS para, então, transmitir o quadro, resultando em uma colisão. Por conseguinte, o canal será desperdiçado durante todo o período da transmissão de H1, bem como durante a transmissão de H2.

Para evitar esse problema, o protocolo IEEE 802.11 permite que uma estação utilize um quadro de controle **RTS** (*Request to Send* — solicitação de envio) curto e um quadro de controle **CTS** (*Clear to Send* — pronto para envio) curto para *reservar* acesso ao canal. Quando um remetente quer enviar um quadro DATA, ele pode enviar primeiro um quadro RTS ao AP indicando o tempo total requerido para transmitir o quadro DATA e o quadro de reconhecimento (ACK). Quando o AP recebe o quadro RTS, responde fazendo a transmissão por difusão de um quadro CTS. Esse quadro CTS tem duas finalidades: dá ao remetente uma permissão explícita para enviar e também instrui as outras estações a não enviar durante o tempo reservado.

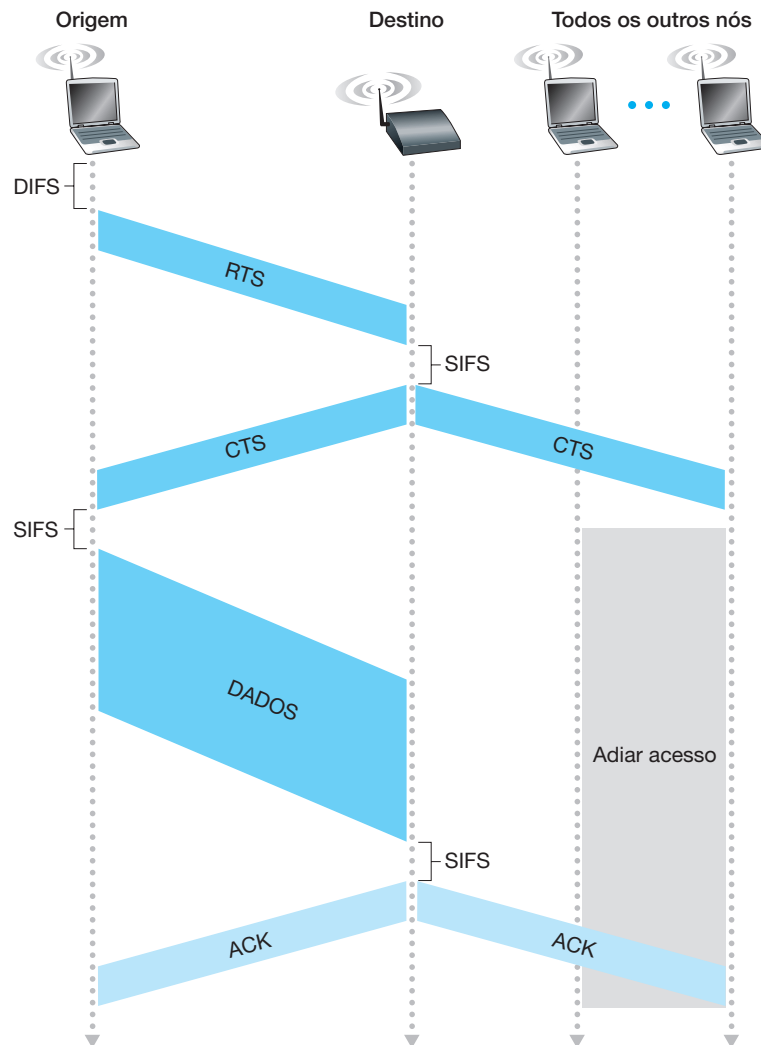
Assim, na Figura 6.12, antes de transmitir um quadro DATA, H1 primeiro faz uma transmissão por difusão de um quadro RTS, que é ouvida por todas as estações que estiverem dentro do seu círculo de alcance, incluindo o AP. O AP então responde com um quadro CTS, que é ouvido por todas as estações dentro de sua faixa de alcance, incluindo H1 e H2. Como ouviu o CTS, a estação H2 deixa de transmitir durante o tempo especificado no quadro CTS. Os quadros RTS, CTS, DATA e ACK são mostrados na Figura 6.12.

A utilização dos quadros RTS e CTS pode melhorar o desempenho de dois modos importantes:

- O problema da estação oculta é atenuado, visto que um quadro DATA longo é transmitido apenas após o canal ter sido reservado.
- Como os quadros RTS e CTS são curtos, uma colisão que envolva um quadro RTS ou CTS terá apenas a duração dos quadros RTS ou CTS curtos. Desde que os quadros RTS e CTS sejam corretamente transmitidos, os quadros DATA e ACK subsequentes deverão ser transmitidos sem colisões.

Aconselhamos o leitor a verificar o applet 802.11 no site deste livro. Esse applet interativo ilustra o protocolo CSMA/CA, incluindo a sequência de troca RTS/CTS.

Embora a troca RTS/CTS ajude a reduzir colisões, também introduz atraso e consome recursos do canal. Por essa razão, a troca RTS/CTS é utilizada (quando utilizada) apenas para reservar o canal para a transmissão de um quadro DATA longo. Na prática, cada estação sem fio pode estabelecer um patamar RTS tal que a sequência RTS/CTS seja utilizada somente quando o quadro for mais longo do que o patamar. Para muitas estações sem fio, o valor *default* do patamar RTS é maior do que o comprimento máximo do quadro, de modo que a sequência RTS/CTS é omitida para todos os quadros DATA enviados.

FIGURA 6.12 PREVENÇÃO DE COLISÃO USANDO OS QUADROS RTS E CTS

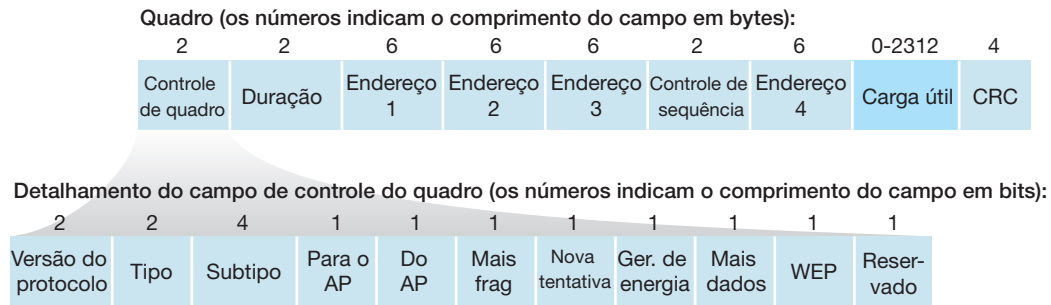
Usando o 802.11 como enlace ponto a ponto

Até aqui nossa discussão focalizou a utilização do 802.11 em um cenário de múltiplo acesso. Devemos mencionar que, se dois nós tiverem, cada um, uma antena direcional, eles poderão dirigir suas antenas um para o outro e executar o protocolo 802.11 sobre o que é, essencialmente, um enlace ponto-a-ponto. Dado o baixo custo comercial do hardware 802.11, a utilização de antenas direcionais e uma maior potência de transmissão permitem que o 802.11 seja utilizado como um meio barato de prover conexões sem fio ponto-a-ponto por dezenas de quilômetros. Raman [2007] descreve uma dessas redes sem fio multissaltos que funciona nas planícies rurais do rio Ganges, na Índia, e que contém enlaces 802.11 ponto a ponto.

6.3.3 O quadro IEEE 802.11

Embora o quadro 802.11 tenha muitas semelhanças com um quadro Ethernet, ele também contém vários campos que são específicos para sua utilização para enlaces sem fio. O quadro 802.11 é mostrado na Figura 6.13.

FIGURA 6.13 O QUADRO 802.11



Os números acima de cada campo no quadro representam os comprimentos dos campos em *bytes*; os números acima de cada subcampo no campo de controle do quadro representam os comprimentos dos subcampos em *bits*. Agora vamos examinar os campos no quadro, bem como alguns dos subcampos mais importantes no campo de controle do quadro.

Campos de carga útil e de CRC

No coração do quadro está a carga útil, que consiste, tipicamente, em um datagrama IP ou em um pacote ARP. Embora o comprimento permitido do campo seja 2.312 bytes, em geral ele é menor do que 1.500 bytes, contendo um datagrama IP ou um pacote ARP. Como um quadro Ethernet, um quadro 802.11 inclui uma verificação de redundância cíclica (CRC), de modo que o receptor possa detectar erros de bits no quadro recebido. Como já vimos, erros de bits são muito mais comuns em LANs sem fio do que em LANs cabeadas, portanto, aqui, CRC é ainda mais útil.

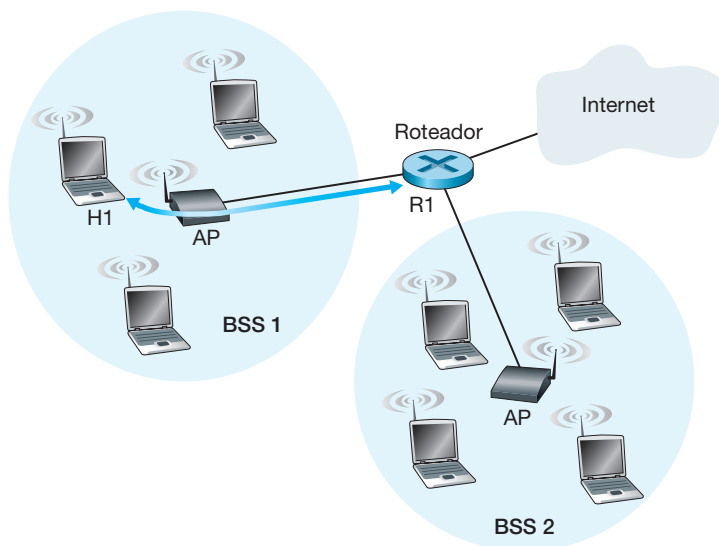
Campos de endereço

Talvez a diferença mais marcante no quadro 802.11 é que ele tem *quatro* campos de endereço e cada um pode conter um endereço MAC de 6 bytes. Mas por que quatro campos de endereço? Um campo de origem MAC e um campo de destino MAC não são suficientes como são na Ethernet? Acontece que aqueles três campos de endereço são necessários para finalidades de interconexão em rede — especificamente, para mover o datagrama de camada de enlace de uma estação sem fio, passando por um AP, até uma interface de roteador. O quarto campo de endereço é usado quando APs encaminham quadros uns aos outros em modo *ad hoc*. Visto que estamos considerando apenas redes de infraestrutura, vamos concentrar nossa atenção nos três primeiros campos de endereço. O padrão 802.11 define esses campos da seguinte forma:

- Endereço 2 é o endereço MAC da estação que transmite o quadro. Assim, se uma estação sem fio transmitir o quadro, o endereço MAC daquela estação será inserido no campo de endereço 2. De modo semelhante, se um AP transmitir o quadro, o endereço MAC do AP será inserido no campo de endereço 2.
- Endereço 1 é o endereço MAC da estação sem fio que deve receber o quadro. Assim, se uma estação móvel sem fio transmitir o quadro, o endereço 1 conterá o endereço MAC do AP de destino. De modo semelhante, se um AP transmitir o quadro, o endereço 1 conterá o endereço MAC da estação sem fio de destino.
- Para entender o endereço 3, lembre-se de que o BSS (que consiste no AP e estações sem fio) faz parte de uma sub-rede, e que esta se conecta com outras sub-redes, por meio de alguma interface de roteador. O endereço 3 contém o endereço MAC dessa interface de roteador.

Para compreender melhor a finalidade do endereço 3, vamos examinar um exemplo de interconexão em rede no contexto da Figura 6.14. Nessa figura há dois APs, cada um responsável por certo número de estações sem fio. Cada AP tem uma conexão direta com um roteador que, por sua vez, se liga com a Internet global.

FIGURA 6.14 A UTILIZAÇÃO DE CAMPOS DE ENDEREÇO EM QUADROS 802.11: MOVENDO UM QUADRO ENTRE H1 E R1



Devemos ter sempre em mente que um AP é um dispositivo da camada de enlace e, portanto, não “fala” IP nem entende endereços IP. Agora, considere mover um datagrama da interface de roteador R1 até a estação sem fio H1. O roteador não está ciente de que há um AP entre ele e H1; do ponto de vista do roteador, H1 é apenas um hospedeiro em uma das sub-redes às quais ele (o roteador) está conectado.

- O roteador, que conhece o endereço IP de H1 (pelo endereço de destino do datagrama), utiliza ARP para determinar o endereço MAC de H1, exatamente como aconteceria em uma LAN Ethernet comum. Após obter o endereço MAC de H1, a interface do roteador R1 encapsula o datagrama em um quadro Ethernet. O campo de endereço de origem desse quadro contém o endereço MAC de R1 e o campo de endereço de destino contém o endereço MAC de H1.
- Quando o quadro Ethernet chega ao AP, este converte o quadro Ethernet 802.3 para um quadro 802.11 antes de transmiti-lo para o canal sem fio. O AP preenche o endereço 1 e o endereço 2 com o endereço MAC de H1 e seu próprio endereço MAC, respectivamente, como descrito. Para o endereço 3, o AP insere o endereço MAC de R1. Dessa maneira, H1 pode determinar (a partir do endereço 3) o endereço MAC da interface de roteador que enviou o datagrama para a sub-rede.

Agora considere o que acontece quando a estação sem fio H1 responde movendo um datagrama de H1 para R1.

- H1 cria um quadro 802.11, preenchendo os campos de endereço 1 e 2 com o endereço MAC do AP e com o endereço MAC de H1, respectivamente, como descrito. Para o endereço 3, H1 insere o endereço MAC de R1.
- Quando recebe o quadro 802.11, o AP o converte para um quadro Ethernet. O campo de endereço de origem para esse quadro é o endereço MAC de H1 e o campo de endereço de destino é o endereço MAC de R1. Assim, o endereço 3 permite que o AP determine o endereço de destino MAC apropriado ao construir o quadro Ethernet.

Em resumo, o endereço 3 desempenha um papel crucial na interconexão do BSS com uma LAN cabeada.

Campos de número de sequência, duração e controle de quadro

Lembre-se de que, em 802.11, sempre que uma estação recebe corretamente um quadro de outra, devolve um reconhecimento. Como reconhecimentos podem ser perdidos, a estação emissora pode enviar várias cópias

de determinado quadro. Como vimos em nossa discussão sobre o protocolo rdt2.1 (Seção 3.4.1), a utilização de números de sequência permite que o receptor distinga entre um quadro recém-transmitido e a retransmissão de um quadro anterior. Assim, o campo de número de sequência no quadro 802.11 cumpre aqui, na camada de enlace, exatamente a mesma finalidade que cumpria na camada de transporte do Capítulo 3.

Não se esqueça de que o protocolo 802.11 permite que uma estação transmissora reserve o canal durante um período que inclui o tempo para transmitir seu quadro de dados e o tempo para transmitir um reconhecimento. Esse valor de duração é incluído no campo de duração do quadro (tanto para quadros de dados quanto para os quadros RTS e CTS).

Como mostrado na Figura 6.13, o campo de controle de quadro inclui muitos subcampos. Diremos apenas umas poucas palavras sobre alguns dos mais importantes; se o leitor quiser uma discussão mais completa, aconselhamos que consulte a especificação 802.11 [Held, 2001; Crow, 1997; IEEE 802.11, 1999]. Os campos *tipo* e *subtipo* são usados para distinguir os quadros de associação, RTS, CTS, ACK e de dados. Os campos *de* e *para* são usados para definir os significados dos diferentes campos de endereço. (Esses significados mudam dependendo da utilização dos modos *ad hoc* ou de infraestrutura e, no caso do modo de infraestrutura, mudam dependendo de o emissor do quadro ser uma estação sem fio ou um AP.) Finalmente, o campo WEP (*Wireless Equivalent Privacy*) indica se está sendo ou não utilizada criptografia. (A WEP é discutida no Capítulo 8.)

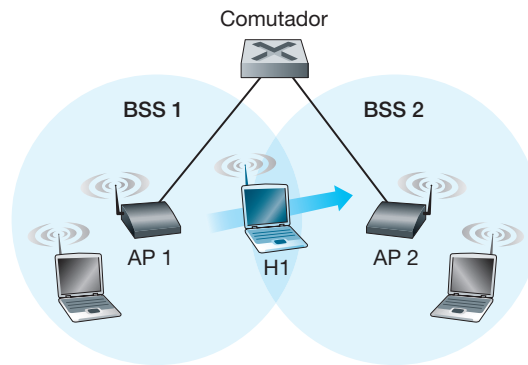
6.3.4 Mobilidade na mesma sub-rede IP

Para ampliar a faixa física de uma LAN sem fio, empresas e universidades frequentemente distribuirão vários BSSs dentro da mesma sub-rede IP. Isso, claro, levanta a questão da mobilidade entre os BSSs — como estações sem fio passam imperceptivelmente de um BSS para outro enquanto mantêm sessões TCP em curso? Como veremos nesta subseção, a mobilidade pode ser manipulada de uma maneira relativamente direta quando os BSSs são parte de uma sub-rede. Quando estações se movimentam entre sub-redes, são necessários protocolos de gerenciamento de mobilidade mais sofisticados, tais como os que estudaremos nas seções 6.5 e 6.6.

Agora vamos examinar um exemplo específico de mobilidade entre BSSs na mesma sub-rede. A Figura 6.15 mostra dois BSSs interconectados e um hospedeiro, H1, que se move entre BSS1 e BSS2. Como nesse exemplo o dispositivo de interconexão entre os dois BSSs *não* é um roteador, todas as estações nos dois BSSs, incluindo os APs, pertencem à mesma sub-rede IP. Assim, quando H1 se move de BSS1 para BSS2, ele pode manter seu endereço IP e todas as suas conexões TCP em curso. Se o dispositivo de interconexão fosse um roteador, então H1 teria de obter um novo endereço IP na sub-rede na qual estava percorrendo. Essa mudança de endereço interromperia (e, em consequência, finalizaria) qualquer conexão TCP em curso em H1. Na Seção 6.6, veremos como um protocolo de mobilidade da camada de rede, como o IP móvel, pode ser usado para evitar esse problema.

Mas o que acontece especificamente quando H1 passa de BSS1 para BSS2? À medida que se afasta de AP1, o H1 detecta um enfraquecimento do sinal de AP1 e começa a fazer uma varredura em busca de um sinal mais forte. H1 recebe quadros de sinalização de AP2 (que em muitos ambientes empresariais e universitários terão o mesmo SSID do AP1). Então, H1 se desassocia de AP1 e se associa com AP2, mantendo, ao mesmo tempo, seu endereço IP e suas sessões TCP em curso.

Isso resolve o problema de transferência do ponto de vista do hospedeiro e do AP. Mas e quanto ao comutador na Figura 6.15? Como é possível saber que o hospedeiro se locomoveu de um AP a outro? O leitor talvez se lembre de que, no Capítulo 5, dissemos que comutadores são “autodidatas” e constroem automaticamente suas tabelas de repasse. Essa característica de autoaprendizagem funciona bem para movimentações ocasionais (por exemplo, quando um profissional é transferido de um departamento para outro); contudo, comutadores não são projetados para suportar usuários com alto grau de mobilidade, que querem manter conexões TCP enquanto se movimentam entre BSSs. Para avaliar este problema aqui, lembre-se de que, antes da movimentação, o comutador tem um registro em sua tabela de repasse que vincula o endereço MAC de H1 com a interface de saída do comutador por meio da qual o H1 pode ser alcançado. Se H1 estiver inicialmente em BSS1, então um datagrama

FIGURA 6.15 MOBILIDADE NA MESMA SUB-REDE

destinado a H1 será direcionado a ele via AP1. Contudo, tão logo H1 se associe com BSS2, seus quadros deverão ser direcionados para AP2. Uma solução (na verdade um tanto forçada) é o AP2 enviar ao comutador um quadro Ethernet de difusão com o endereço de origem de H1 logo após a nova associação. Quando o comutador recebe o quadro, atualiza sua tabela de repasse, permitindo que H1 seja alcançado via AP2. O grupo de padrões 802.11f está desenvolvendo um protocolo entre APs para cuidar dessas e outras questões associadas.

6.3.5 Recursos avançados em 802.11

Finalizaremos nossa abordagem sobre 802.11 com uma breve discussão sobre capacidades avançadas encontradas nas redes 802.11. Como veremos, essas capacidades *não* são completamente especificadas no padrão 802.11, mas, em vez disso, são habilitadas por mecanismos especificados no padrão. Isso permite que fornecedores diferentes implementem essas capacidades usando suas próprias abordagens, aparentemente trazendo vantagens sobre a concorrência.

Adaptação da taxa 802.11

Vimos antes, na Figura 6.3, que as diferentes técnicas de modulação (com as diferentes taxas de transmissão que elas fornecem) são adequadas para diferentes cenários de SNR. Considere, por exemplo, um usuário 802.11 móvel que está, inicialmente, a 20 metros de distância da estação-base, com uma alta relação sinal-ruído. Dada a alta SNR, o usuário pode se comunicar com essa estação usando uma técnica de modulação da camada física que oferece altas taxas de transmissão enquanto mantém uma BER baixa. Esse usuário é um felizarado! Suponha agora que o usuário se torne móvel, se distanciando da estação-base, e que a SNR diminui à medida que a distância da estação-base aumenta. Neste caso, se a técnica de modulação usada no protocolo 802.11 que está operando entre a estação-base e o usuário não mudar, a BER será inaceitavelmente alta à medida que a SNR diminui e, por conseguinte, nenhum quadro transmitido será recebido corretamente.

Por essa razão, algumas execuções de 802.11 possuem uma capacidade de adaptação de taxa que seleciona, de maneira adaptável, a técnica de modulação da camada física sobreposta a ser usada com base em características atuais ou recentes do canal. Se um nó enviar dois quadros seguidos sem receber confirmação (uma indicação implícita de erros de bit no canal), a taxa de transmissão cai para a próxima taxa mais baixa. Se dez quadros seguidos forem confirmados, ou se um temporizador (que registra o tempo desde o último *fallback*) expirar, a taxa de transmissão aumenta para a próxima taxa mais alta. Esse mecanismo de adaptação da taxa compartilha a mesma filosofia de “investigação” (refletida por recebimentos de ACK); quando algo “ruim” acontece, a taxa de transmissão é reduzida. A adaptação da taxa 802.11 e o controle de congestionamento TCP são, desse modo, semelhantes à criança: está sempre exigindo mais e mais de seus pais (digamos, mais doces, para uma criança, e chegar mais

tarde em casa, para adolescentes) até eles por fim dizerem “Chega!” e a criança desistir (para tentar novamente após a situação melhorar!). Diversos métodos também foram propostos para aperfeiçoar esse esquema básico de ajuste automático de taxa [Kamerman, 1997; Holland, 2001; Lacage, 2004].

Gerenciamento de energia

A energia é uma fonte preciosa em aparelhos móveis e, assim, o padrão 802.11 provê capacidades de gerenciamento de energia, permitindo que os nós 802.11 minimizem o tempo de suas funções de percepção, transmissão e recebimento, e outros circuitos necessários para “funcionar”. O gerenciamento de energia 802.11 opera da seguinte maneira. Um nó é capaz de alternar claramente entre os estados “dormir” e “acordar” (como um aluno com sono em sala de aula!). Um nó indica ao ponto de acesso que entrará no modo de dormir ajustando o bit de gerenciamento de energia no cabeçalho de um quadro 802.11 para 1. Um temporizador localizado no nó é, então, ajustado para acordar o nó antes de um AP ser programado para enviar seu quadro de sinalização (lembre-se de que, em geral, um AP envia um quadro de sinalização a cada 100 ms). Uma vez que o AP descobre, pelo bit de transmissão de energia, que o nó vai dormir, ele (o AP) sabe que não deve enviar nenhum quadro àquele nó, e armazenará qualquer quadro destinado ao hospedeiro que está dormindo para transmissão posterior.

Um nó acordará logo após o AP enviar um quadro de sinalização, e logo entrará no modo ativo (ao contrário do aluno sonolento, esse despertar leva apenas 250 μ s [Kamerman, 1997]!). Os quadros de sinalização enviados pelo AP contêm uma relação de nós cujos quadros foram mantidos em buffer no AP. Se não houver quadros mantidos em buffer para o nó, ele pode voltar a dormir. Caso contrário, o nó pode solicitar explicitamente o envio dos quadros armazenados, enviando uma mensagem de *polling* ao AP. Com um tempo de 100 ms entre as sinalizações, um despertar de 250 μ s e um tempo semelhantemente pequeno para receber um quadro de sinalização e verificar que não haja quadros em buffer, um nó que não possui quadros para enviar ou receber pode dormir 99% do tempo, resultando em uma economia de energia significativa.

6.3.6 Redes pessoais: Bluetooth e Zigbee

Como ilustrado na Figura 6.2, o padrão Wi-Fi IEEE 802.11 é voltado para a comunicação entre aparelhos separados por até 100 m (exceto quando 802.11 é usado em configuração ponto a ponto com antena direcional). Dois outros protocolos IEEE 802 — Bluetooth e Zigbee (definidos nos padrões IEEE 802.15.1 e IEEE 802.15.4 [IEEE 802.15, 2012]) e o WiMAX (definido no padrão IEEE 802.16 [IEEE 802.16d, 2004; IEEE 802.16e, 2005]) — são padrões para se comunicar a distâncias mais curtas e mais longas, nessa ordem. Examinaremos o WiMAX rapidamente quando discutirmos sobre redes de dados celulares, na Seção 6.4; aqui, vamos nos concentrar em redes para distâncias mais curtas.

Bluetooth

Uma rede IEEE 802.15.1 opera sobre uma curta faixa, a baixa potência, e a um custo baixo. Esta é, basicamente, uma tecnologia de “substituição de cabos” de baixa velocidade, curto alcance e baixa potência para interconectar notebooks, dispositivos periféricos, telefones celulares e smartphones, enquanto 802.11 é uma tecnologia de “acesso” de velocidade mais alta, alcance médio e potência mais alta. Por essa razão, as redes 802.15.1, às vezes, são denominadas redes pessoais sem fio (WPAN — *Wireless Personal Area Networks*). As camadas de enlace e física do 802.15.1 são baseadas na especificação do **Bluetooth** anterior para redes pessoais [Held, 2001; Bisdikian, 2001]. Redes 802.15.1 operam na faixa de rádio não licenciada de 2,4 GHz em modo TDM, com intervalos de tempo (*time slots*) de 625 μ s. Durante cada intervalo de tempo, um emissor transmite por um entre 79 canais, sendo que, de intervalo para intervalo, o canal muda de uma maneira conhecida, porém pseudoaleatória. Essa forma de saltar de canal em canal, conhecida como **espectro espalhado com salto de frequência** (FHSS

— *Frequency-Hopping Spread Spectrum*), espalha transmissões sobre o espectro de frequência pelo tempo. O 802.15.1 pode oferecer velocidades de dados de até 4 Mbits/s.

Redes 802.15.1 são redes *ad hoc*: não é preciso que haja uma infraestrutura de rede (por exemplo, um ponto de acesso) para interconectar dispositivos 802.15.1. Assim, esses dispositivos devem se organizar por si sós. Dispositivos 802.15.1 são primeiro organizados em uma **picorrede** (*piconet*: pequena rede) de até oito dispositivos ativos, como ilustra a Figura 6.16. Um desses dispositivos é designado como o mestre, e os outros agem como escravos. Na verdade, o nó mestre comanda a picorrede como um rei — seu relógio determina o tempo na picorrede, ele pode transmitir em cada intervalo de tempo de número ímpar e um escravo pode transmitir somente após o mestre ter se comunicado com ele no intervalo de tempo anterior e, mesmo assim, só pode transmitir para o mestre. Além dos dispositivos escravos, a rede também pode conter até 255 dispositivos estacionados. Esses dispositivos não podem se comunicar até que o nó mestre tenha mudado seus estados de estacionado para ativo.

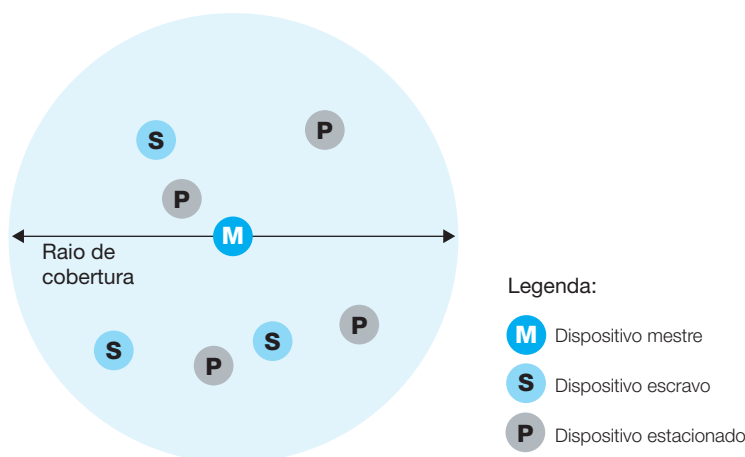
Se o leitor quiser obter mais informações sobre WPANs 802.15.1, pode consultar as referências do Bluetooth [Held, 2001; Bisdikian, 2001] ou o site oficial do IEEE 802.15 [IEEE 802.15, 2012].

Zigbee

Uma segunda rede pessoal padronizada pelo IEEE é o padrão 802.14.5 [IEEE 802.15, 2012], conhecido como Zigbee. Enquanto as redes Bluetooth oferecem uma taxa de dados de “substituição de cabo” de mais de um Mbit por segundo, Zigbee é voltada para aplicações de menos potência, menor taxa de dados e menor ciclo de trabalho do que Bluetooth. Embora sejamos tentados a pensar que “maior e mais rápido é melhor”, nem todas as aplicações de rede precisam de muita largura de banda e dos custos mais altos decorrentes disso (custos tanto econômicos quanto de energia). Por exemplo, sensores domésticos de temperatura e iluminação, dispositivos de segurança e interruptores de parede são todos dispositivos muito simples, de baixa potência, baixo ciclo de trabalho e baixo custo. Portanto, Zigbee é ideal para esses dispositivos. Zigbee define taxas de canal de 20, 40, 100 e 250 Kbits/s, dependendo da frequência do canal.

Os nós em uma rede Zigbee podem ser de dois tipos. Os chamados “dispositivos de função reduzida” operam como escravos controlados por um único “dispositivo de função completa”, assim como dispositivos Bluetooth escravos. Um dispositivo de função completa pode operar como um dispositivo mestre no Bluetooth controlando vários dispositivos escravos, e múltiplos dispositivos de função completa podem, além disso, ser configurados em uma rede de malha, na qual direcionam quadros entre si. Zigbee compartilha muitos mecanismos de protocolo que já encontramos em outros protocolos da camada de enlace: quadros de sinalização e confirmações da camada de enlace (semelhantes ao 802.11), protocolos de acesso aleatório de detecção de portadora com

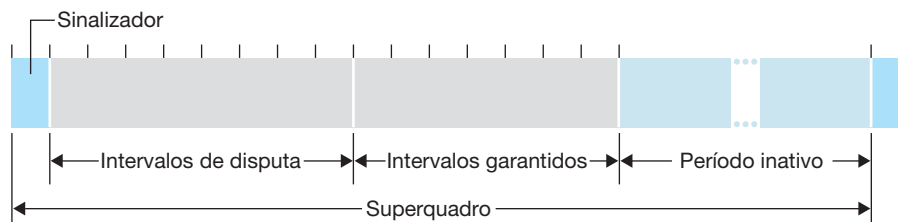
FIGURA 6.16 UMA PICORREDE BLUETOOTH



recuo exponencial binário (semelhante ao 802.11 e Ethernet) e alocação fixa, garantida, de intervalos de tempo (semelhante ao DOCSIS).

Redes Zigbee podem ser configuradas de muitas maneiras diferentes. Vamos considerar o caso simples de um único dispositivo de função completa controlando vários dispositivos de função reduzida com intervalos de tempo e usando quadros de sinalização. A Figura 6.17 mostra um caso em que a rede Zigbee divide o tempo em superquadros recorrentes, cada qual começando com um quadro de sinalização. Cada quadro de sinalização divide o superquadro em um período ativo (durante o qual os dispositivos podem transmitir) e um período inativo (durante o qual todos os dispositivos, inclusive o controlador, podem dormir e, portanto, economizar energia). O período ativo consiste em 16 intervalos de tempo, alguns usados pelos dispositivos em uma forma de acesso aleatório CSMA/CA, e alguns são alocados pelo controlador para dispositivos específicos, oferecendo assim o acesso garantido ao canal para esses dispositivos. Outros detalhes sobre redes Zigbee poderão ser encontrados em Baronti [2007], IEEE 802.15.4 [2012].

FIGURA 6.17 ESTRUTURA DO SUPERQUADRO ZIGBEE 802.14.4



6.4 ACESSO CELULAR À INTERNET

Na seção anterior vimos como um hospedeiro da Internet pode acessá-la quando estiver dentro de um *hotspot* Wi-Fi, isto é, quando estiver nas vizinhanças de um ponto de acesso 802.11. Mas a área de cobertura da maioria dos *hotspots* Wi-Fi é pequena, entre 10 e 100 m de diâmetro. O que fazer, então, quando precisamos desesperadamente de um acesso sem fio à Internet e não podemos acessar um *hotspot* Wi-Fi?

Dado que a telefonia celular agora está presente por toda parte, em muitas áreas no mundo inteiro, uma estratégia natural é estender redes celulares de modo que suportem não apenas a telefonia de voz, mas também o acesso sem fio à Internet. Idealmente, esse acesso teria uma velocidade razoavelmente alta e ofereceria mobilidade imperceptível, permitindo que usuários mantivessem suas sessões TCP enquanto viajassem, por exemplo, em um ônibus ou em um trem. Com taxas de bits altas o bastante entre a origem e o usuário e vice-versa, o usuário poderia até mesmo manter sessões de videoconferência enquanto estivesse em trânsito. Esse cenário não é assim tão implausível. Em 2012, muitas operadoras de telefonia celular ofereciam a seus assinantes um serviço de acesso celular à Internet por menos de 50 dólares mensais com taxas de bits entre a operadora e o usuário e vice-versa na faixa de algumas centenas de kilobits por segundo. Taxas de dados de vários megabits por segundo estão se tornando acessíveis à medida que os serviços de dados de banda larga, como aqueles que veremos aqui, são cada vez mais empregados.

Nesta seção, damos uma breve visão geral das tecnologias de acesso celular à Internet já existentes e emergentes. Mais uma vez, aqui focalizaremos o primeiro salto sem fio, bem como a rede que conecta o primeiro salto sem fio à rede telefônica maior e/ou à Internet; na Seção 6.7 consideraremos como as chamadas são roteadas para um usuário que está se movimentando entre estações-base. Nossa breve discussão oferecerá, necessariamente, apenas uma descrição simplificada e de alto nível das tecnologias celulares. É claro que a questão das comunicações celulares modernas é de grande amplitude e profundidade, e muitas universidades oferecem diversos cursos sobre o assunto. O leitor que desejar um conhecimento mais profundo da questão deve consultar Goodman [1997]; Kaaranen [2001]; Lin [2001]; Korhonen [2003]; Schiller [2003]; Scourias [2012]; Turner [2012]; Akyildiz [2010], bem como as referências particularmente excelentes e completas em Mouly [1992].

6.4.1 Visão geral da arquitetura de rede celular

Em nossa discussão sobre a arquitetura da rede celular nesta seção, adotaremos a terminologia dos padrões do *Sistema Global para Comunicações Móveis* (GSM — *Global System for Mobile Communications*). (Para os amantes de história, a sigla GSM, na origem, veio de *Groupe Spécial Mobile*, até a forma anglicizada ser adotada, mantendo as letras iniciais originais.) Nos anos 1980, os europeus reconheceram a necessidade de um sistema pan-europeu de telefonia celular digital que substituiria os diversos sistemas analógicos de telefonia celular incompatíveis, levando ao padrão GSM [Mouly, 1992]. Os europeus implantaram a tecnologia GSM com um grande sucesso no início dos anos 1990 e, desde então, o GSM cresceu e se tornou um gigante do mundo da telefonia celular, com mais de 80% de assinantes no mundo utilizando essa tecnologia.

Quando as pessoas falam sobre tecnologia celular, em geral a classificam como pertencendo a uma das diversas “gerações”. As primeiras gerações foram, em essência, projetadas para o tráfego de voz. Os sistemas de primeira geração (1G) eram sistemas FDMA analógicos, desenvolvidos especialmente para a comunicação apenas por voz. Esses sistemas 1G quase não existem mais, tendo sido substituídos pelos sistemas digitais 2G. Os sistemas originais 2G também foram projetados para voz, sendo estendidos mais tarde (2,5G) para suportar dados (por exemplo, a Internet), bem como o serviço de voz. Os sistemas 3G, que estão sendo executados hoje, também suportam voz e dados, mas com uma ênfase cada vez maior nas capacidades de dados e enlaces de acesso via rádio com maior velocidade.

HISTÓRIA

Celular móvel 3G versus LANs sem fio

Muitas operadoras de telefonia celular móvel estão disponibilizando sistemas celulares móveis 3G, com taxas de dados de 2 Mbits/s em ambiente fechado e 384 kbits/s, ou mais, ao ar livre. Esses sistemas 3G estão sendo disponibilizados em faixas de radiofrequência licenciadas, e o preço que algumas operadoras pagam aos governos pela licença é uma quantia considerável. Os sistemas 3G permitirão que os usuários acessem a Internet a partir de locais externos remotos enquanto estão em trânsito, de modo semelhante ao acesso por telefone celular existente hoje. Por exemplo, a tecnologia 3G permite que um usuário acesse informações sobre mapas de estradas enquanto dirige seu carro, ou informações sobre a programação de cinema enquanto estiver tomando sol na praia. Não obstante, pode-se questionar a extensão à qual os sistemas 3G serão usados, dados o seu custo e o fato de que os usuários em geral possuem acesso simultâneo a LANs sem fio e 3G:

- A infraestrutura emergente de LAN sem fio logo estará presente em quase toda parte. As LANs sem fio IEEE 802.11, que operam a taxas de 54 Mbits/s,

estão sendo amplamente disponibilizadas. Quase todos os computadores portáteis e smartphones já vêm da fábrica equipados com capacidades de LAN 802.11. Além disso, as novidades em equipamentos de Internet — como câmeras sem fio e molduras para fotografias — também terão capacidades de LAN sem fio de baixa potência.

- Estações-base de LANs sem fio também podem manipular equipamentos de telefonia móvel. Muitos telefones já podem se conectar com a rede de telefonia celular ou a uma rede IP de forma nativa ou usando o serviço de voz sobre IP semelhante ao Skype, ultrapassando, assim, os serviços de dados 3G e de voz por celular da operadora.

É claro que muitos outros especialistas acham que a tecnologia 3G não só será o maior sucesso, mas também revolucionará drasticamente o modo como trabalhamos e vivemos. É claro que ambas, Wi-Fi e 3G, podem se tornar tecnologias sem fio dominantes, com equipamentos sem fio em trânsito que selecionam automaticamente a tecnologia de acesso que oferece o melhor serviço em seu local físico atual.

Arquitetura de rede celular, 2G: conexões por voz à rede telefônica

O termo *celular* refere-se ao fato de que uma área geográfica é dividida em várias áreas de cobertura geográfica, conhecidas como células, ilustradas como hexágonos à esquerda da Figura 6.18. Assim como estudamos o padrão Wi-Fi 802.11 na Seção 6.3.1, o GSM possui sua nomenclatura específica. Cada célula contém uma estação-base de transceptor (BTS) que transmite e recebe sinais de estações móveis dentro de sua célula. A área de cobertura de uma célula depende de muitos fatores, incluindo potência da BTS, potência de transmissão de aparelhos do usuário, obstáculos na célula, como prédios, e altura das antenas da estação-base. Embora a Figura 6.18 mostre cada célula com uma estação-base de transceptor posicionada no seu meio, hoje, muitos sistemas posicionam a BTS em pontos de interseção de três células, de modo que uma única BTS com antenas direcionais possa atender a três células.

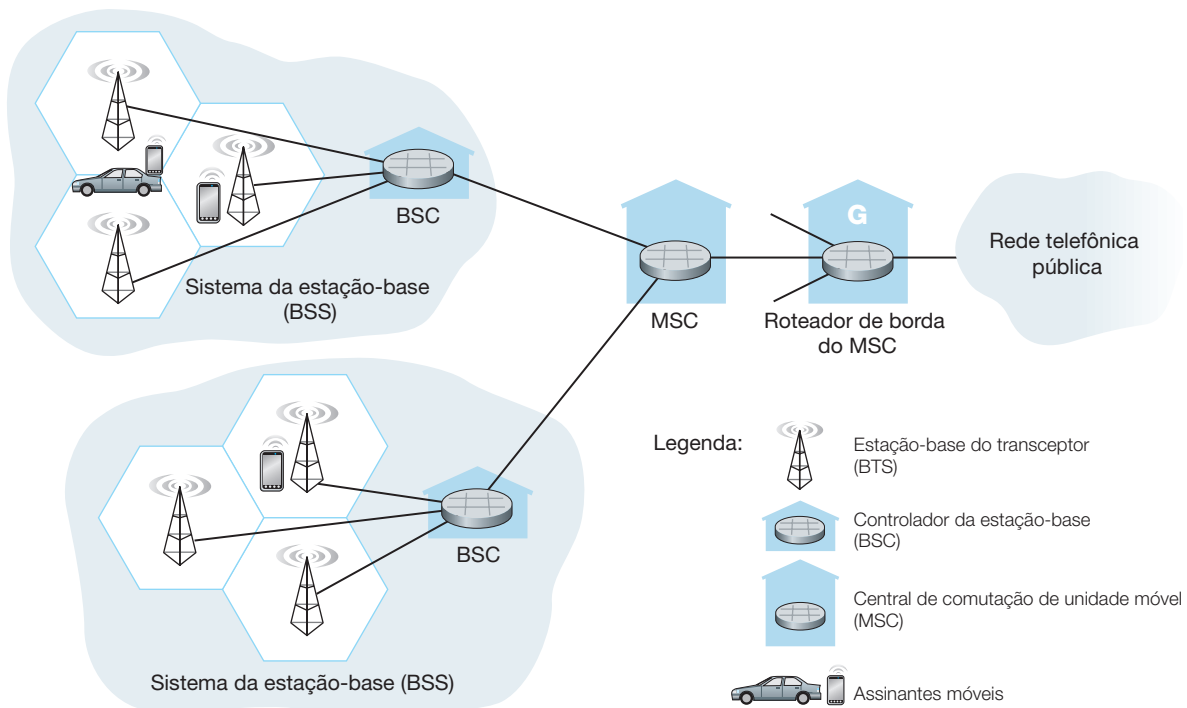
O padrão GSM para sistemas celulares 2G utiliza FDM/TDM (rádio) combinados para a interface ar. Lembre-se de que vimos no Capítulo 1 que, com FDM puro, o canal é dividido em uma série de bandas de frequência, e cada banda se dedica a uma chamada. Também no Capítulo 1, vimos que, com FDM puro, o tempo é dividido em quadros, cada quadro é dividido em intervalos e cada chamada é destinada ao uso de um intervalo específico no quadro rotativo. Nos sistemas combinados FDM/TDM, o canal é dividido em uma série de sub-bandas de frequência; dentro de cada sub-banda, o tempo é dividido em quadros e intervalos. Desse modo, para um sistema combinado FDM/TDM, se o canal for dividido em F sub-bandas e o tempo em T intervalos, então o canal poderá suportar $F \cdot T$ chamadas simultâneas. Lembre-se de que vimos, na Seção 5.3.4, que as redes de acesso a cabo também usam uma técnica combinada FDM/TDM. Os sistemas GSM consistem em bandas de frequência de 200-kHz e cada banda suporta oito chamadas TDM. O GSM codifica a voz a 13 kbits/s e 12,2 kbits/s.

Um **controlador de estação-base (BSC)** da rede GSM normalmente prestará serviço a dezenas de estações-base do transceptor. O papel da BSC é alocar os canais de rádio da BTS a assinantes móveis, realizar **paginação** (encontrar a célula na qual reside um usuário móvel) e realizar transferência de usuários móveis — um assunto que abordaremos, em poucas palavras, na Seção 6.7.2. O controlador de estação-base e suas estações-base de transceptor, coletivamente, constituem um **sistema de estação-base (BSS)** GSM.

Como veremos na Seção 6.7, a **central de comutação móvel (MSC)** desempenha o papel central na contabilidade e autorização do usuário (por exemplo, determinar se um aparelho móvel tem permissão para se conectar à rede celular), estabelecimento e interrupção de chamada, e transferências. Em geral, uma única MSC conterá até cinco BSCs, resultando em cerca de 200K assinantes por MSC. A rede de um provedor de celular terá diversas MSCs, com MSCs especiais conhecidas como roteadores de borda das MSCs, conectando a rede celular do provedor à rede telefônica pública mais ampla.

6.4.2 Redes de dados celulares 3G: estendendo a Internet para assinantes de celular

Nossa discussão na Seção 6.4.1 foi concentrada na conexão de usuários de voz celular à rede telefônica pública. Porém, claro, quando estamos nos movimentando, também queremos ler e-mails, acessar a Web, obter serviços dependentes do local (por exemplo, mapas e recomendações de restaurantes) e talvez ainda assistir a vídeos. Para isso, nosso smartphone precisará executar uma pilha de protocolos TCP/IP completa (incluindo as camadas física, de enlace, rede, transporte e aplicação) e conectar-se à Internet por meio da rede de dados celular. O assunto de redes de dados celulares é uma coleção um tanto confusa de padrões concorrentes e em evolução, à medida que uma geração (e meia geração) substitui a anterior e introduz novas tecnologias e serviços, com novos acrônimos. Para piorar as coisas ainda mais, não há um único órgão oficial que defina os requisitos para as tecnologias 2,5G, 3G, 3,5G ou 4G, tornando mais difícil apontar as diferenças entre os padrões concorrentes. Em nossa discussão a seguir, vamos nos concentrar nos padrões 3G do UMTS (*Universal Mobile Telecommunications*

FIGURA 6.18 COMPONENTES DA ARQUITETURA DE REDE CELULAR 2G GSM

Service), desenvolvidos pelo Projeto de Parceria da 3ª Geração (3GPP — *3rd Generation Partnership Project*) [3GPP 2012], uma tecnologia 3G bastante difundida.

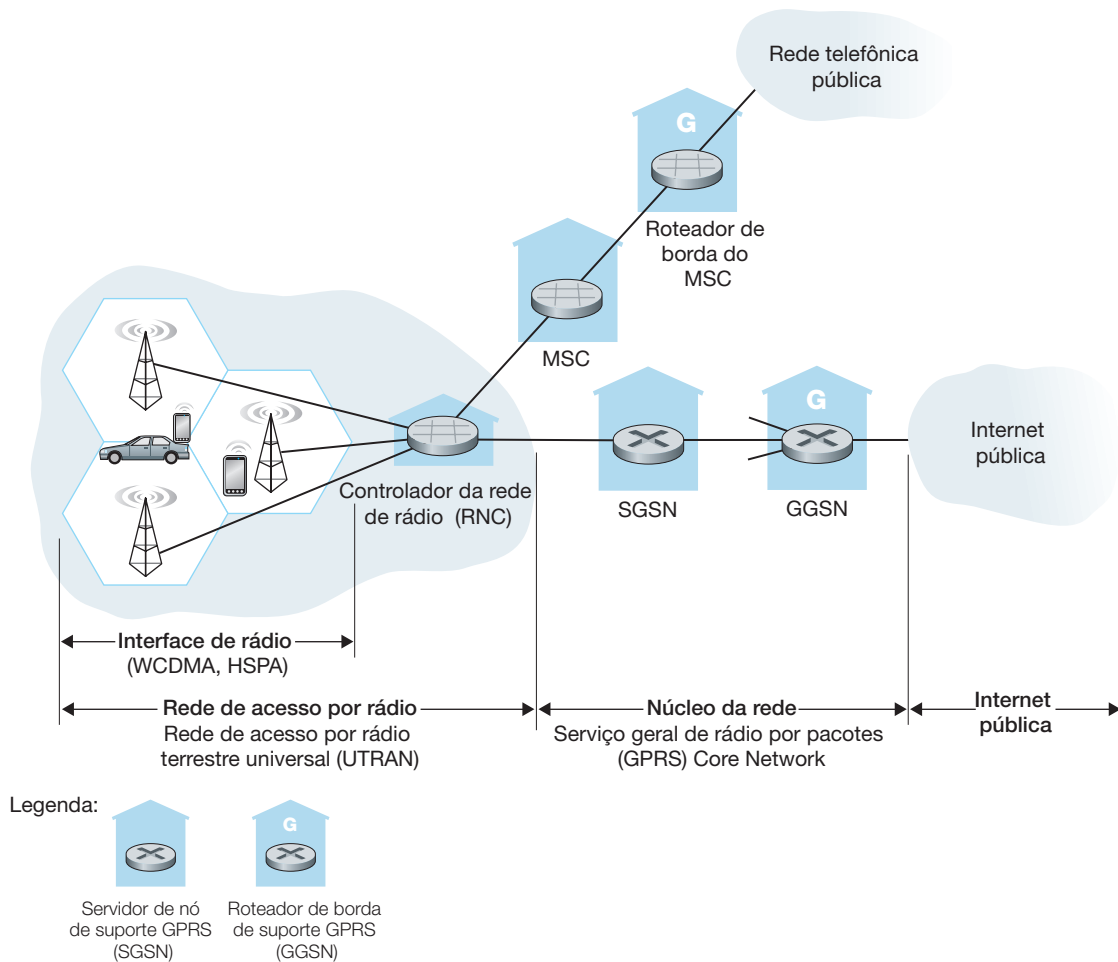
Vamos fazer uma análise de cima para baixo da arquitetura de rede de dados celular 3G mostrada na Figura 6.19.

O núcleo da rede 3G

O núcleo da rede de dados celular 3G conecta as redes de acesso por rádio à Internet pública. O núcleo da rede opera em conjunto com os componentes da rede celular de voz existente (particularmente, o MSC), que encontramos anteriormente na Figura 6.18. Dada a quantidade considerável de infraestrutura existente (e serviços lucrativos!) na rede celular de voz, o método utilizado pelos projetistas dos serviços de dados 3G é evidente: *deixe o núcleo da rede celular de voz GSM intacto, acrescentando funcionalidade de dados por celular em paralelo à rede de voz existente*. A alternativa — integrar novos serviços de dados diretamente no núcleo da rede celular de voz — teria gerado os mesmos desafios encontrados na Seção 4.4.4, na qual discutimos sobre a integração das tecnologias IPv6 (nova) e IPv4 (legada) na Internet.

Existem dois tipos de nós no núcleo da rede 3G: **Servidor de Nó de Suporte GPRS (SGSN — Serving GPRS Support Nodes)** e **Roteador de borda de suporte GPRS (GGSN — Gateway GPRS Support Nodes)**. (O acrônimo GPRS significa *Generalized Packet Radio Service* — serviço de rádio por pacotes generalizado —, um antigo serviço celular de dados em redes 2G; aqui, discutimos a versão aperfeiçoada do GPRS nas redes 3G.) Um SGSN é responsável por entregar datagramas de e para os nós móveis na rede de acesso por rádio à qual o SGSN está ligado. O SGSN interage com o MSC da rede celular de voz para essa área, oferecendo autorização do usuário e transferência, mantendo informações de local (célula) sobre nós móveis ativos e realizando repasse de datagramas entre os nós móveis na rede de acesso por rádio e um GGSN. O GGSN atua como um roteador de borda (*gateway*), conectando vários SGSNs à Internet maior. Um GGSN, portanto, é a última parte da infraestrutura 3G que um datagrama originado do nó móvel encontra antes de entrar na Internet. Para o

FIGURA 6.19 ARQUITETURA DO SISTEMA 3G



mundo exterior, o GGSN é semelhante a qualquer outro roteador de borda; a mobilidade dos nós 3G dentro da rede do GGSN fica oculta do mundo exterior, por trás do GGSN.

A rede de acesso por rádio 3G: a borda sem fio

A **rede de acesso por rádio 3G** é a rede do primeiro salto sem fio que vemos como usuários do 3G. O **controlador da rede de rádio (RNC)** em geral controla várias estações-base transceptoras da célula, semelhantes às estações-base que encontramos nos sistemas 2G (mas, no linguajar do UMTS 3G, conhecida oficialmente como “Node Bs” — um nome nada descritivo!). O enlace sem fio de cada célula opera entre os nós móveis e uma estação-base transceptora, assim como nas redes 2G. O RNC se conecta à rede de voz do celular por comutação de circuitos via um MSC, e à Internet por comutação de pacotes via um SGSN. Assim, embora os serviços de voz e dados por celular 3G utilizem núcleos de rede diferentes, eles compartilham uma rede comum de acesso por rádio do primeiro e último salto.

Uma mudança significativa no UMTS 3G em relação às redes 2G é que, em vez de usar o método FDMA/TDMA do GSM, o UMTS emprega uma técnica CDMA denominada *Direct Sequence Wideband CDMA* (CDMA banda larga de sequência direta) (DS-WCDMA) [Dahlman, 1998] dentro de intervalos TDMA; estes, por sua vez, são acessíveis em frequências múltiplas — uma utilização interessante de todos os três métodos de compartilhamento de canal dedicado que identificamos no Capítulo 5 e semelhante à técnica usada nas redes de acesso a cabo (veja Seção 5.3.4). Essa mudança exige uma nova rede celular 3G de acesso sem fio operando paralelamente

com a rede de rádio BSS 2G mostrada na Figura 6.18. O serviço de dados associado à especificação do WCDMA é conhecido como Acesso a Pacote em Alta Velocidade (HSP — *High Speed Packet Access*) e promete taxas de dados de até 14 Mbits/s. Mais detalhes referentes às redes 3G podem ser encontrados no site do Projeto de Parceria da 3ª Geração (3GPP) [3GPP, 2012].

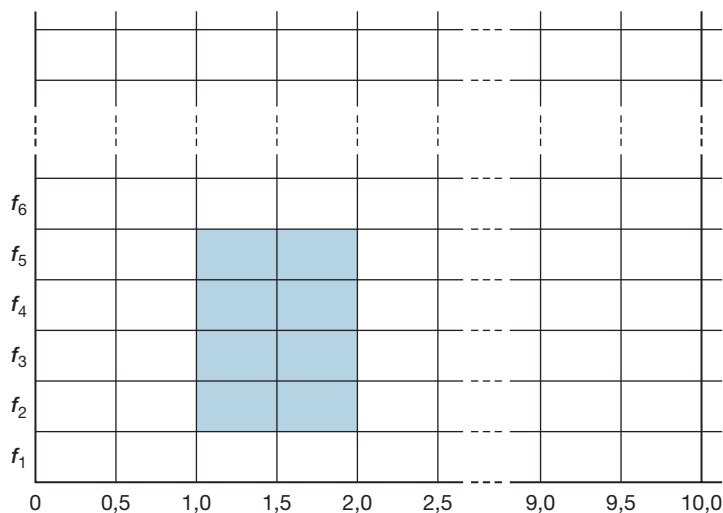
6.4.3 No caminho para o 4G: LTE

Com os sistemas 3G agora sendo utilizados no mundo inteiro, será que os sistemas 4G estão bem atrasados? Decerto não! Na realidade, projeto, teste e implantação inicial dos sistemas 4G já estão sendo realizados. O padrão 4G Long-Term Evolution (LTE) apresentado pelo 3GPP tem duas inovações importantes em relação aos sistemas 3G:

- **Núcleo de pacote desenvolvido (EPC — *Evolved Packet Core*)** [3GPP Network Architecture, 2012]. O EPC é uma rede de núcleo simplificado em IP, que unifica a rede celular de voz comutada por circuitos separada e a rede celular de dados comutada por pacotes mostrada na Figura 6.19. Esta é uma rede “toda em IP”, pois voz e dados serão transportados em datagramas IP. Como vimos no Capítulo 4 e estudaremos com mais detalhes no Capítulo 7, o modelo de serviço de “melhor esforço” do IP não é inerentemente bem adequado para os requisitos de desempenho exigentes do tráfego VoIP (voz sobre IP), a menos que os recursos da rede sejam controlados com cuidado para evitar (em vez de reagir a) congestionamento. Assim, uma tarefa importante do EPC é controlar os recursos da rede para oferecer essa alta qualidade de serviço. O EPC também faz uma separação nítida entre os planos de controle da rede e dados do usuário, com muitos dos recursos de suporte a mobilidade que estudaremos na Seção 6.7 sendo executados no plano de controle. O EPC permite que vários tipos de redes de acesso por rádio, incluindo redes de acesso 2G e 3G legadas, se conectem ao núcleo da rede. Duas introduções bastante claras ao EPC são [Motorola, 2007; Alcatel-Lucent, 2009].
- **Rede de acesso por rádio LTE.** O padrão LTE usa uma combinação de multiplexação por divisão de frequência e multiplexação por divisão de tempo no canal descendente, conhecida como multiplexação por divisão de frequência ortogonal (OFDM) [Rohde, 2008; Ericsson, 2011]. (O termo “ortogonal” vem do fato de que os sinais enviados em diferentes canais de frequência são criados de modo que interfiram muito pouco uns nos outros, mesmo quando as frequências de canal são pouco espaçadas.) No LTE, cada nó móvel ativo recebe um ou mais intervalos de tempo de 0,5 ms em uma ou mais das frequências do canal. A Figura 6.20 mostra a alocação de oito intervalos de tempo sobre quatro frequências. Recebendo cada vez mais intervalos de tempo (seja na mesma frequência ou em frequências diferentes), um nó móvel pode alcançar velocidades de transmissão cada vez mais altas. A (re)alocação de intervalo entre os nós móveis pode ser realizada até mesmo a cada milissegundo. Diferentes esquemas de modulação também podem ser usados para alterar a taxa de transmissão; veja nossa discussão anterior da Figura 6.3 e a seleção dinâmica de esquemas de modulação em redes Wi-Fi. Outra inovação na rede de rádio LTE é o uso de sofisticadas antenas de entrada múltipla, saída múltipla (MIMO). A taxa de dados máxima para um usuário LTE é 100 Mbits/s na direção descendente e 50 Mbits/s na direção ascendente, quando estiver usando 20 MHz do espectro sem fio.

A alocação em particular de intervalos de tempo a nós móveis não é exigida pelo padrão LTE. Em vez disso, a decisão de quais nós móveis terão permissão para transmitir em dado intervalo de tempo em dada frequência é determinada pelos algoritmos de escalonamento fornecidos pelo fornecedor de equipamento LTE e/ou operador da rede. Com o escalonamento oportunista [Bender, 2000; Kolding, 2003; Kulkarni, 2005], combinando o protocolo da camada física e as condições do canal entre remetente e destinatário e escolhendo os destinatários aos quais os pacotes serão enviados com base nas condições do canal, permite que o controlador da rede de rádio faça o melhor uso do meio sem fio. Além disso, as prioridades do usuário e os níveis de

FIGURA 6.20 VINTE INTERVALOS DE 0,5 MILISSEGUNDOS ORGANIZADOS EM QUADROS DE 10 MILISSEGUNDOS EM CADA FREQUÊNCIA. A REGIÃO SOMBREADA INDICA UMA ALOCAÇÃO DE OITO INTERVALOS



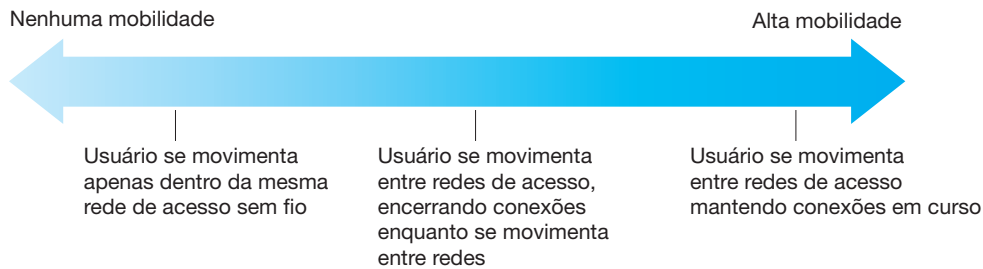
serviço contratados (por exemplo, prata, ouro ou platina) podem ser usados no escalonamento das transmissões descendentes de pacotes. Além das capacidades do LTE descritas aqui, LTE-Advanced permite larguras de banda descendentes de centenas de Mbits/s, com a alocação de canais agregados a um nó móvel [Akyildiz, 2010].

Outra tecnologia 4G sem fio — WiMAX (World Interoperability for Microwave Access) — é uma família de padrões IEEE 802.16 que diferem bastante do LTE. Ainda não sabemos se a tecnologia 4G preferida será LTE ou WiMAX, mas no momento (2012), LTE parece ter um impulso mais significativo. Uma discussão detalhada do WiMAX poderá ser encontrada no site deste livro.

6.5 GERENCIAMENTO DA MOBILIDADE: PRINCÍPIOS

Após estudarmos a natureza *sem fio* dos enlaces de comunicação em uma rede sem fio, é hora de voltarmos nossa atenção à *mobilidade* que esses enlaces sem fio possibilitam. No sentido mais amplo, um nó móvel é aquele que muda seu ponto de conexão com a rede ao longo do tempo. Como o termo *mobilidade* adquiriu muitos significados nos campos da computação e da telefonia, será proveitoso, antes de tudo, considerarmos diversas dimensões da mobilidade com algum detalhe.

- *Do ponto de vista da camada de rede, até que ponto um usuário é móvel?* Um usuário fisicamente móvel apresentará à camada de rede conjuntos de desafios muito diferentes dependendo de como ele se movimenta entre pontos de conexão com a rede. Em uma extremidade do espectro na Figura 6.21, um usuário pode carregar consigo um notebook equipado com uma placa de interface de rede sem fio dentro de um prédio. Como vimos na Seção 6.3.4, esse usuário *não* é móvel do ponto de vista da camada de rede. Além do mais, se ele se associar com o mesmo ponto de acesso independentemente do local, então não será móvel nem mesmo do ponto de vista da camada de enlace.
- Na outra extremidade do espectro, considere o usuário que está dentro de um BMW, correndo pela estrada a 150 km/h, passando por várias redes de acesso sem fio e querendo manter uma conexão TCP ininterrupta com uma aplicação remota durante a viagem. Esse usuário é *definitivamente* móvel! Entre esses extremos está um usuário que leva seu notebook de um local (por exemplo, escritório ou quarto de dormir) a outro (por exemplo, lanchonete, biblioteca) e quer se conectar à rede no novo local. Ele também é móvel (embora menos do que o motorista da BMW!), mas não precisa manter uma conexão

FIGURA 6.21 VÁRIOS GRAUS DE MOBILIDADE DO PONTO DE VISTA DA CAMADA DE REDE

ativa enquanto se movimenta entre pontos de conexão com a rede. A Figura 6.21 ilustra esse espectro de mobilidade do usuário do ponto de vista da camada de rede.

- *Qual é a importância de o endereço do nó móvel permanecer sempre o mesmo?* Com a telefonia móvel, o número de seu telefone — basicamente, o endereço de camada de rede do seu aparelho — permanece o mesmo quando você transita entre uma operadora de rede de telefonia móvel e outra. Um notebook também terá de manter o mesmo endereço IP enquanto se movimenta entre redes IP?

A resposta a essa pergunta dependerá muito da aplicação que está sendo executada. Para o motorista da BMW que quer manter uma conexão TCP ininterrupta com uma aplicação remota enquanto voa pela estrada, seria conveniente manter o mesmo endereço IP. Lembre-se de que dissemos, no Capítulo 3, que uma aplicação de Internet precisa conhecer o endereço IP e o número de porta da entidade remota com a qual está se comunicando. Se uma entidade móvel puder manter seu endereço IP enquanto estiver em trânsito, a mobilidade torna-se invisível do ponto de vista da aplicação. Essa transparência é de grande valor — uma aplicação não precisa se preocupar com uma potencial mudança de endereço IP e o mesmo código de aplicação servirá igualmente a conexões móveis e não móveis. Na próxima seção veremos que o IP móvel oferece essa transparência, permitindo que um nó móvel mantenha seu endereço IP permanente enquanto se movimenta entre redes.

Por outro lado, um usuário móvel menos sofisticado poderia querer apenas desligar seu notebook no escritório, levá-lo para casa, ligá-lo de novo e continuar trabalhando. Se o notebook funciona principalmente como um cliente em aplicações cliente-servidor (por exemplo, enviar/receber e-mails, navegar pela web, usar Telnet com um hospedeiro remoto), o endereço IP particular utilizado pelo notebook não é tão importante. Em particular, é possível passar muito bem com um endereço temporariamente alocado ao notebook pelo ISP que atende a residência. Na Seção 4.4, vimos que o DHCP já oferece essa funcionalidade.

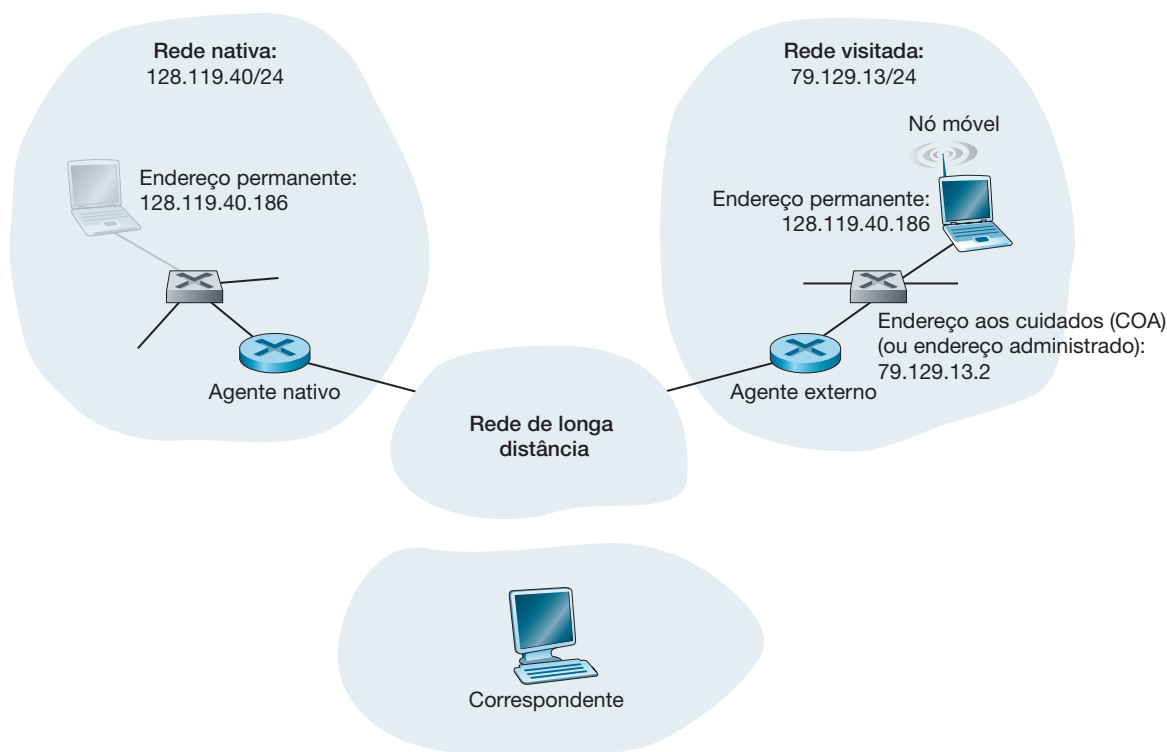
- *Qual é a infraestrutura cabeada de suporte disponível?* Em todos os quatro cenários descritos, admitimos de modo implícito haver uma infraestrutura fixa à qual o usuário móvel pode se conectar — por exemplo, a rede do ISP que atende a residência, a rede de acesso sem fio no local de trabalho ou as redes de acesso sem fio na rodovia. E se tal infraestrutura não existir? Se dois usuários estiverem a uma distância que permita comunicação, eles poderão estabelecer uma conexão de rede na ausência de qualquer outra infraestrutura de camada de rede? As redes *ad hoc* oferecem exatamente essas capacidades. Essa área, que está em rápido desenvolvimento e representa a pesquisa de ponta em redes móveis, está fora do escopo deste livro. Perkins [2000] e as páginas Web do grupo de trabalho Mobile *Ad hoc* Network (manet) do IETF apresentam um tratamento aprofundado do assunto.

Para ilustrar as questões que envolvem permitir que um usuário mantenha conexões em curso enquanto se movimenta entre redes, vamos fazer uma analogia com seres humanos. Um adulto de 20 e poucos anos que sai da casa dos pais torna-se móvel, pois passa a morar em uma série de quartos e/ou apartamentos e está sempre mudando de endereço. Se uma velha amiga quiser entrar em contato com ele, como conseguirá o endereço de seu amigo móvel? Uma maneira comum de fazer isso é entrar em contato com a família, já que um jovem

móvel costuma informar seus novos endereços (nem que seja só para que os pais possam lhe enviar dinheiro para ajudar a pagar o aluguel!). A residência da família, com seu endereço permanente, torna-se o único lugar a que outros podem se dirigir como uma primeira etapa para estabelecer comunicação com o jovem móvel. As comunicações posteriores da velha amiga podem ser indiretas (por exemplo, pelo envio de uma carta primeiro à casa dos pais, que a encaminharão ao jovem móvel) ou diretas (por exemplo, utilizar o endereço informado pelos pais e enviar uma carta diretamente ao amigo móvel).

Em um ambiente de rede, a residência permanente de um nó móvel (tal como um notebook ou um smartphone) é conhecida como **rede nativa** (*home network*) e a entidade dentro dessa rede que executa o gerenciamento de funções de mobilidade em nome do nó móvel é conhecida como **agente nativo** (*home agent*). A rede na qual o nó móvel está residindo é conhecida como **rede externa** (*foreign network*) ou **rede visitada** (*visited network*), e a entidade dentro da rede externa que auxilia o nó móvel no gerenciamento das funções de mobilidade discutidas adiante é conhecida como **agente externo** (*foreign agent*). No caso de profissionais móveis, suas redes nativas possivelmente seriam as redes das empresas em que trabalham, enquanto a rede visitada poderia ser a rede de um colega que estão visitando. Um **correspondente** é a entidade que quer se comunicar com o nó móvel. A Figura 6.22 ilustra esses conceitos, bem como conceitos de endereçamento considerados mais adiante. Observe que, na Figura 6.22, os agentes são colocados junto dos roteadores (por exemplo, como processos que rodam em roteadores), mas, como alternativa, poderiam estar rodando em outros hospedeiros ou servidores na rede.

FIGURA 6.22 ELEMENTOS INICIAIS DE UMA ARQUITETURA DE REDE MÓVEL



6.5.1 Endereçamento

Já observamos que, para a mobilidade do usuário ser transparente para aplicações de rede, é desejável que um nó móvel mantenha seu endereço quando transita de uma rede para outra. Quando um nó móvel residir em uma rede externa, todo o tráfego enviado ao endereço permanente do nó agora precisará ser roteado para a rede

externa. Como isso pode ser feito? Uma opção é a rede externa anunciar para todas as outras que o nó móvel agora reside em sua rede, o que poderia ser feito mediante a costumeira troca de informações de roteamento interdomínio e intradomínio, exigindo poucas mudanças na infraestrutura de roteamento. A rede externa poderia apenas anunciar a seus vizinhos que tem uma rota altamente específica para o endereço permanente do nó móvel (isto é, em essência, informar a outras redes que ela tem o caminho correto para rotear datagramas para o endereço permanente do nó móvel; ver Seção 4.4). Esses vizinhos então propagariam a informação de roteamento por toda a rede como parte do procedimento normal de atualização de informações de roteamento e tabelas de repasse. Quando o nó móvel saísse de uma rede externa e se juntasse a outra, a nova rede externa anunciaria uma nova rota, altamente específica, até o nó móvel e a antiga rede externa retiraria suas informações de roteamento referentes ao nó móvel.

Esse procedimento resolve dois problemas de uma vez só e o faz sem promover mudanças significativas na infraestrutura da camada de rede. Outras redes conhecem a localização do nó móvel e é fácil rotear datagramas para o nó móvel, visto que as tabelas de repasse dirigirão datagramas à rede externa. Uma desvantagem significativa, contudo, é a da facilidade de expansão. Se a responsabilidade pelo gerenciamento da mobilidade tivesse de recair sobre os roteadores da rede, eles teriam de manter registros em tabelas de repasse para potencialmente milhões de nós móveis e atualizar esses registros à medida que os nós se movimentassem. Algumas desvantagens adicionais serão exploradas nos problemas ao final deste capítulo.

Um método alternativo (que tem sido adotado na prática) é passar a funcionalidade de mobilidade do núcleo da rede para a borda — um tema recorrente em nosso estudo da arquitetura da Internet. Um modo natural de fazer isso é por meio da rede nativa do nó móvel. De maneira muito semelhante ao modo como os pais daquele jovem de 20 e poucos anos monitoram a localização do filho, o agente nativo na rede nativa do nó móvel pode monitorar a rede externa na qual o nó móvel reside. Decerto será preciso um protocolo (ou um agente externo representando o nó móvel) entre o nó móvel e o agente nativo para atualizar a localização do nó móvel.

Agora vamos considerar o agente externo com mais detalhes. A técnica conceitualmente mais simples, mostrada na Figura 6.22, é localizar agentes externos nos roteadores de borda na rede externa. Um dos papéis do agente externo é criar o denominado **endereço aos cuidados** (*care-of-address* — **COA**) ou **endereço administrativo** para o nó móvel, sendo que a parte da rede do endereço COA combinaria com a parte da rede do endereço da rede externa. Assim, há dois endereços associados a um nó móvel, seu **endereço permanente** (semelhante ao endereço da família do nosso jovem móvel) e seu endereço COA, às vezes denominado **endereço externo** (semelhante ao endereço da casa onde nosso jovem móvel estiver residindo). No exemplo da Figura 6.22, o endereço permanente do nó móvel é 128.119.40.186. Quando está visitando a rede 79.129.13/24, o nó móvel tem um COA 79.129.13.2. Um segundo papel desempenhado pelo agente externo é informar ao agente nativo que o nó móvel está residindo em sua rede (a rede do agente externo) e tem o endereço COA informado. Veremos, em breve, que o COA pode ser utilizado para redirecionar datagramas para o nó móvel através de seu agente externo.

Embora tenhamos separado a funcionalidade do nó móvel e do agente externo, vale a pena observar que o nó móvel também pode assumir as responsabilidades do agente externo. Por exemplo, o nó móvel poderia obter um COA na rede externa (por exemplo, utilizando um protocolo como o DHCP) e ele mesmo informar seu COA ao agente nativo.

6.5.2 Roteamento para um nó móvel

Agora já vimos como um nó móvel obtém um COA e como é possível informar esse endereço ao agente nativo. Mas conseguir que o agente nativo conheça o COA resolve apenas parte do problema. Como datagramas devem ser endereçados e repassados para o nó móvel? Visto que só o agente nativo (e não os roteadores no âmbito da rede) conhece a localização do nó móvel, já não será mais suficiente apenas endereçar um datagrama para o endereço permanente do nó móvel e enviá-lo para a infraestrutura da camada de rede. É preciso fazer algo mais. Duas técnicas podem ser identificadas, as quais denominaremos roteamento indireto e roteamento direto.

Roteamento indireto para um nó móvel

Vamos considerar primeiro um correspondente que quer enviar um datagrama a um nó móvel. Na abordagem de **roteamento indireto** o correspondente apenas endereça o datagrama ao endereço permanente do nó móvel, envia o datagrama para a rede e nem precisa saber se o nó móvel reside em sua rede nativa ou está visitando uma rede externa; assim, a mobilidade é completamente transparente para o correspondente. Esses datagramas são primeiro roteados, como sempre, para a rede local do nó móvel. Isso é ilustrado na etapa 1 da Figura 6.23.

Agora vamos voltar nossa atenção ao agente nativo. Além de ser responsável por interagir com um agente externo para monitorar o COA do nó móvel, ele tem outra função muito importante. Sua segunda tarefa é ficar à espreita de datagramas que chegam e são endereçados a nós cuja rede nativa é a rede do agente nativo, mas que estão residindo no momento em uma rede externa. O agente nativo intercepta esses datagramas e então os repassa a um nó móvel por um processo de duas etapas. Primeiro, o datagrama é repassado para o agente externo usando o COA do nó móvel (etapa 2 na Figura 6.23), e depois do agente externo para o nó móvel (etapa 3 na Figura 6.23).

É instrutivo considerar esse redirecionamento com mais detalhes. O agente nativo precisará endereçar o datagrama usando o COA do nó móvel, de modo que a camada de rede roteará o datagrama para a rede externa. Por outro lado, é desejável deixar intacto o datagrama do correspondente, pois a aplicação que recebe o datagrama deve desconhecer que este foi repassado por meio do agente nativo. Ambas as metas podem ser cumpridas fazendo o agente nativo **encapsular** o datagrama original completo do correspondente dentro de um novo (e maior) datagrama. Este é endereçado e entregue ao COA do nó móvel. O agente externo “proprietário” do COA receberá e desencapsulará o datagrama — isto é, removerá o datagrama original do correspondente de dentro daquele datagrama maior de encapsulamento e repassará o datagrama original (etapa 3 na Figura 6.23) para o nó móvel. A Figura 6.24 mostra um datagrama original de um correspondente sendo enviado para a rede nativa, um datagrama encapsulado sendo enviado ao agente externo e o datagrama original sendo entregue ao nó móvel. O leitor atento notará que o encapsulamento/desencapsulamento descrito aqui é idêntico à noção de implementação de túnel discutida no Capítulo 4, no contexto do IP de transmissão para um grupo (*multicast*) e do IPv6.

FIGURA 6.23 ROTEAMENTO INDIRETO PARA UM NÓ MÓVEL

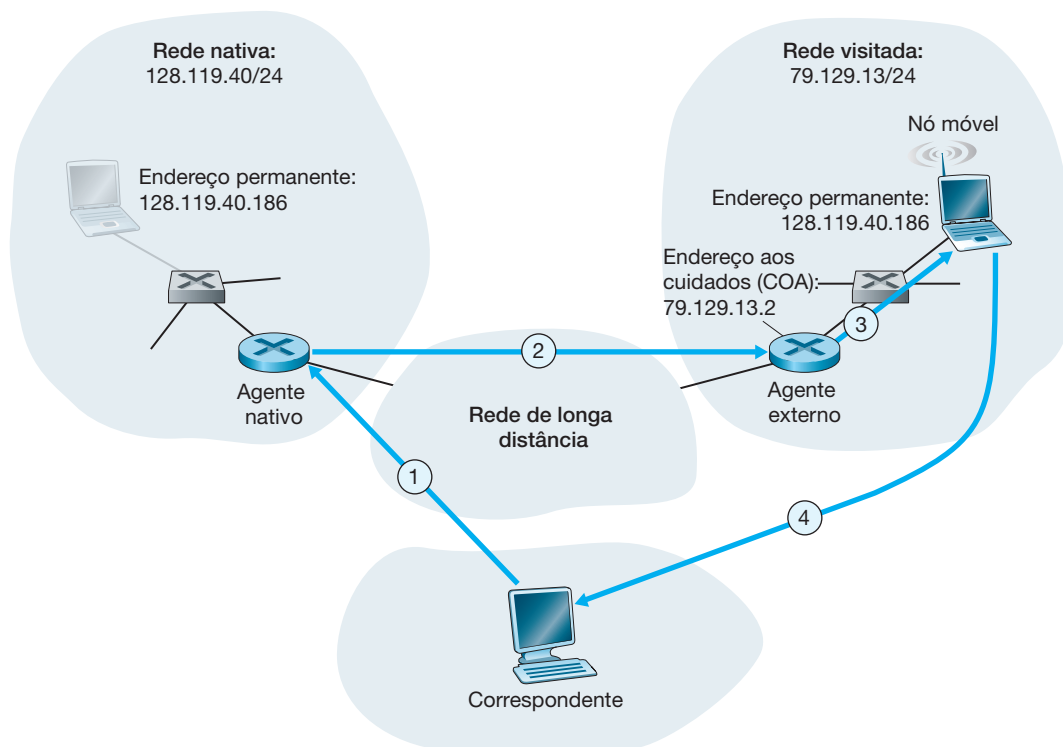
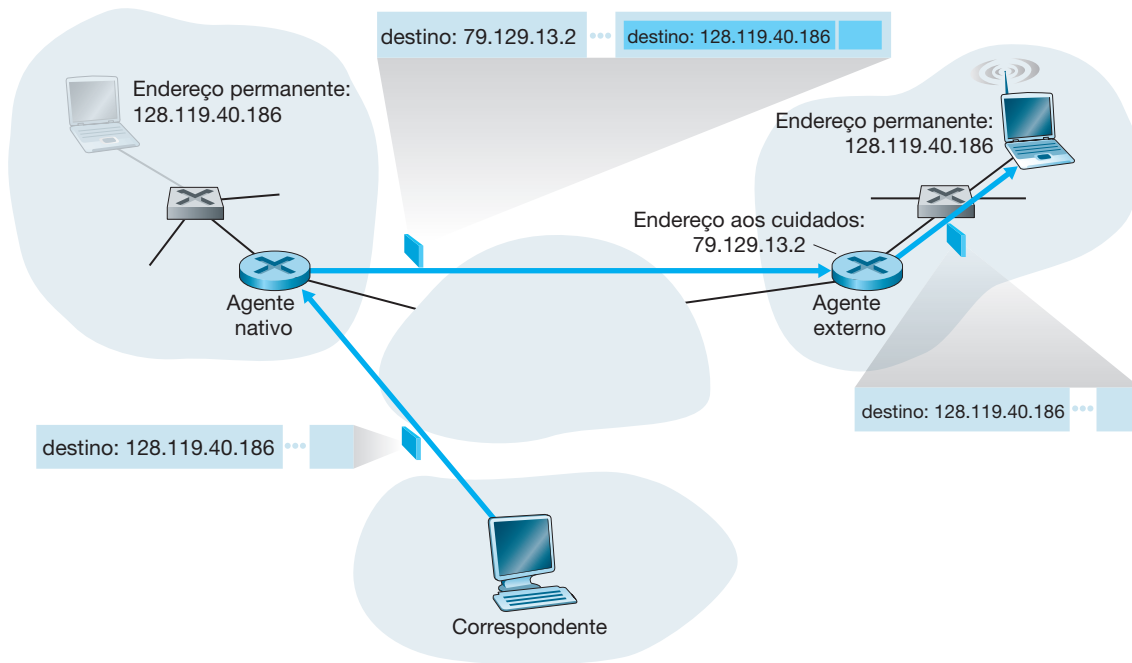


FIGURA 6.24 ENCAPSULAMENTO E DESENCAPSULAMENTO



A seguir, vamos considerar como um nó móvel envia datagramas a um correspondente. Isso é muito simples, pois o nó móvel pode endereçar seu datagrama *diretamente* ao correspondente (usando o próprio endereço permanente como o de origem e o do correspondente como o endereço de destino). Visto que o nó móvel conhece o endereço do correspondente, não há necessidade de rotear o datagrama de volta por meio do agente nativo. Isso é mostrado como etapa 4 na Figura 6.23.

Vamos resumir o que discutimos sobre roteamento indireto relacionando as novas funcionalidades da camada de rede exigidas para dar suporte à mobilidade.

- *Um protocolo de nó móvel ao agente externo.* O nó móvel fará seu registro no agente externo ao se conectar à rede externa. De modo semelhante, um nó móvel cancelará seu registro no agente externo quando sair da rede externa.
- *Um protocolo de registro do agente externo ao agente nativo.* O agente externo registrará o COA do nó móvel no agente nativo. Um agente externo não precisa cancelar explicitamente um COA quando um nó móvel sai da rede porque o registro subsequente de um novo COA se encarregará disso quando o nó móvel passar para uma nova rede.
- *Um protocolo de encapsulamento de datagrama para o agente nativo.* Encapsulamento e repasse do datagrama original do correspondente dentro de um datagrama endereçado ao COA.
- *Um protocolo de desencapsulamento para o agente externo.* Extração do datagrama original do correspondente de dentro daquele que o encapsulou e repasse do datagrama original ao nó móvel.

A discussão anterior provê todas as peças — agentes externos, agente nativo e repasse indireto — de que um nó móvel necessita para manter uma conexão em curso enquanto transita entre redes. Para ilustrar como essas peças se encaixam, suponha que o nó móvel está ligado à rede externa A, registrou um COA no agente local na rede A e está recebendo datagramas que estão sendo roteados indiretamente por meio de seu agente nativo. O nó móvel agora passa para a rede externa B e se registra no agente externo na rede B, que informa ao agente nativo o novo COA do nó móvel. Desse ponto em diante, o agente nativo redirecionará datagramas para a rede externa B. No que diz respeito ao correspondente, a mobilidade é transparente — datagramas são roteados por meio do mesmo agente nativo antes e depois de o nó móvel mudar de rede. Quanto ao agente nativo, não há qualquer ruptura no fluxo de datagramas — datagramas que chegam são repassados primeiro para a rede externa A; após a

mudança no COA, são repassados para a rede externa B. Mas o nó móvel verá um fluxo de datagramas interrompido ao se movimentar entre redes? Contanto que seja pequeno o tempo transcorrido entre o desligamento do nó móvel da rede A (ponto em que ele não pode mais receber datagramas via A) e sua conexão com a rede B (ponto em que registrará um novo COA no agente nativo da rede), poucos datagramas serão perdidos. Lembre-se de que dissemos, no Capítulo 3, que conexões fim a fim podem sofrer perda de datagramas por causa do congestionamento na rede. Por conseguinte, a perda ocasional de datagramas dentro de uma conexão quando um nó se move entre redes não é, de maneira alguma, um problema catastrófico. Se for preciso uma comunicação livre de perdas, mecanismos de camadas superiores recuperarão a perda de dados, quer elas resultem do congestionamento da rede ou da mobilidade do usuário.

Uma abordagem indireta de roteamento é utilizada no padrão IP móvel [RFC 5944], como discutiremos na Seção 6.6.

Roteamento direto para um nó móvel

A abordagem do roteamento indireto ilustrada na Figura 6.23 sofre de uma ineficiência conhecida como problema do roteamento triangular — datagramas endereçados ao nó móvel devem ser roteados primeiro para o agente nativo e em seguida para a rede externa, mesmo quando existir uma rota muito mais eficiente entre o correspondente e o nó móvel. No pior caso, imagine um usuário móvel que está visitando a rede externa de um colega. Os dois estão sentados lado a lado e trocando dados pela rede. Datagramas do correspondente (nesse caso, do colega do visitante) são roteados para o agente nativo do usuário móvel e, então, novamente de volta para a rede externa!

O roteamento direto supera a ineficiência do roteamento triangular, mas o faz à custa de complexidade adicional. Na abordagem do roteamento direto, um agente correspondente na rede do correspondente primeiro aprende o COA do nó móvel. Isso pode ser realizado fazendo o agente correspondente consultar o agente nativo, admitindo que (como é o caso no roteamento indireto) o nó móvel tem um valor atualizado para seu COA registrado no seu agente nativo. Também é possível que o próprio correspondente execute a função do agente correspondente, assim como um nó móvel poderia executar a função do agente externo. Tal situação é ilustrada como as etapas 1 e 2 na Figura 6.25. O agente correspondente, então, executa um túnel para os datagramas diretamente até o COA do nó móvel, de modo semelhante ao túnel executado pelo agente nativo, etapas 3 e 4 da Figura 6.25.

Embora supere o problema do roteamento triangular, o roteamento direto introduz dois importantes desafios adicionais:

- É preciso um **protocolo de localização de usuário móvel** para o agente correspondente consultar o agente nativo de modo a obter o COA do nó móvel (etapas 1 e 2 da Figura 6.25).
- Quando o nó móvel passa de uma rede externa para outra, como os dados são repassados, agora, para a nova rede? No caso do roteamento indireto, esse problema era facilmente resolvido atualizando-se o COA mantido pelo agente nativo. Todavia, com roteamento direto, o agente correspondente consulta o COA junto ao agente nativo apenas uma vez, no início da sessão. Assim, atualizar o COA do agente nativo, embora necessário, não será suficiente para resolver o problema do roteamento de dados para a nova rede externa do nó móvel.

Uma solução seria criar um novo protocolo para notificar a mudança de COA ao correspondente. Uma solução alternativa, que, como veremos, é adotada na prática em redes GSM, funciona da seguinte maneira: suponha que estão sendo repassados dados correntemente para o nó móvel na rede externa onde ele estava localizado quando a sessão foi iniciada (etapa 1 na Figura 6.26). O agente externo naquela rede externa onde o nó móvel foi encontrado pela primeira vez será denominado agente externo âncora. Quando o nó móvel passar para uma nova rede externa (etapa 2 na Figura 6.26), ele se registrará com o novo agente externo (etapa 3), o qual fornecerá ao agente externo âncora o novo COA do nó móvel (etapa 4). Quando o agente externo âncora receber um datagrama encapsulado para um nó móvel que já deixou a rede, ele então poderá reencapsular o datagrama e repassá-lo

para o nó móvel (etapa 5) usando o novo COA. Se, mais tarde, o nó móvel passar para mais alguma outra rede externa, o agente externo nessa nova rede visitada então contatará o agente externo âncora para estabelecer re-passe para essa nova rede externa.

FIGURA 6.25 ROTEAMENTO DIRETO PARA UM USUÁRIO MÓVEL

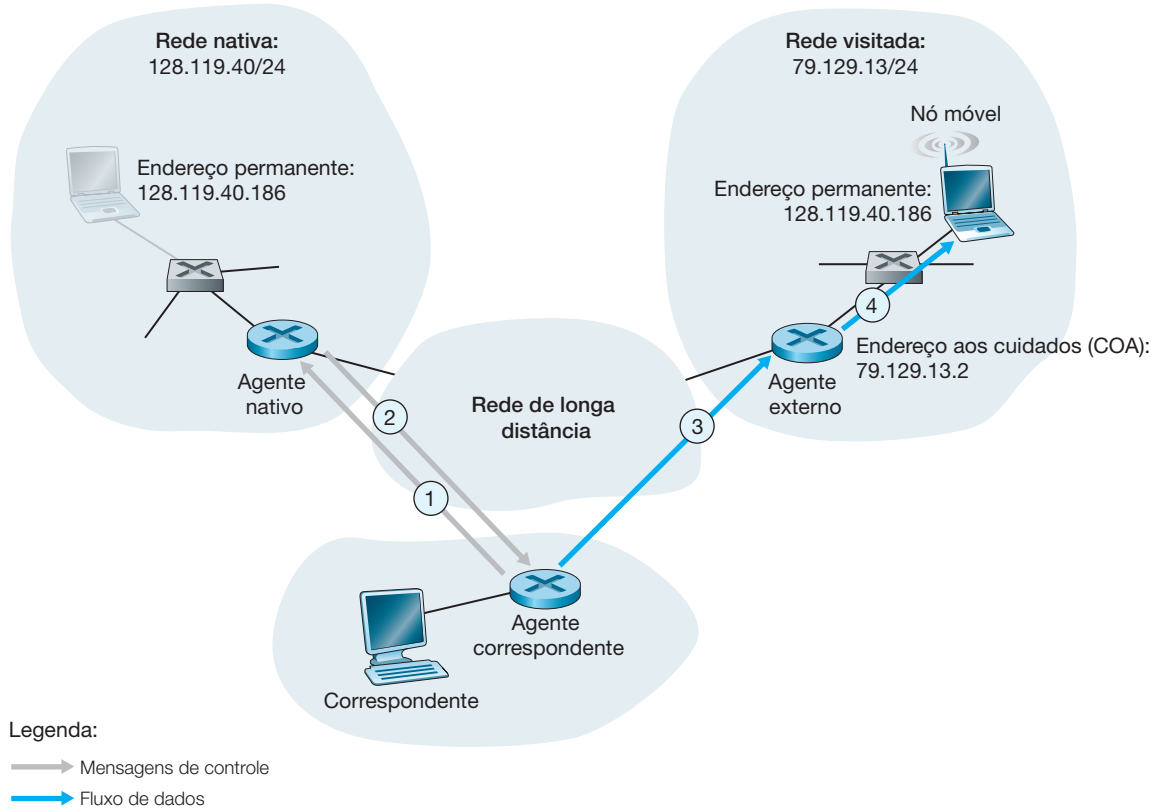
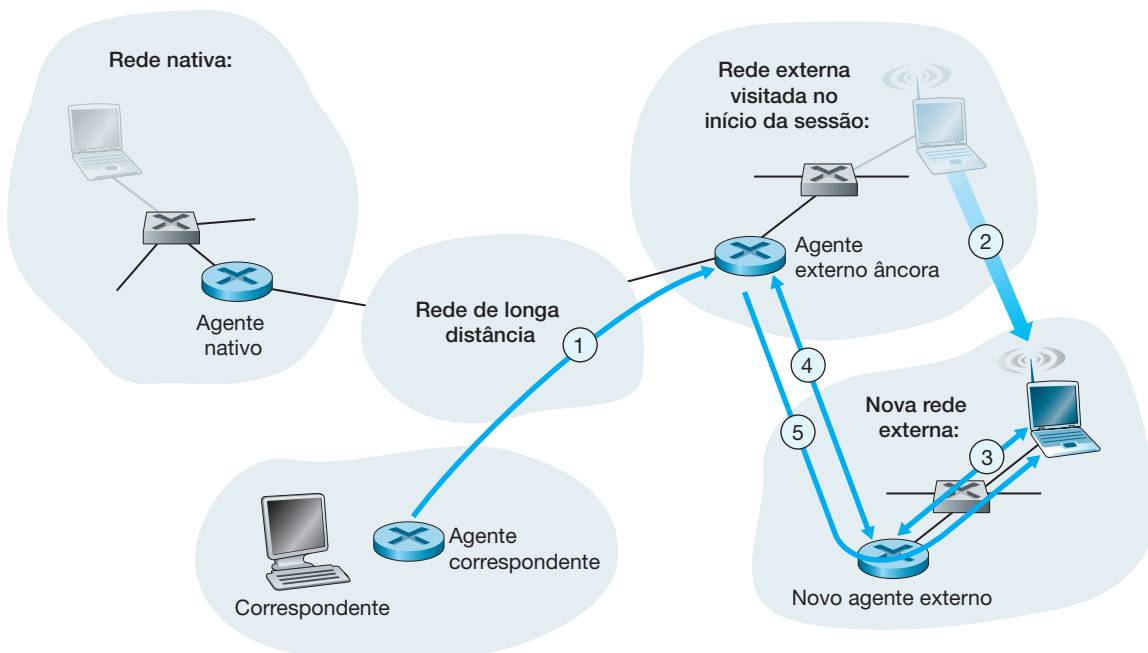


FIGURA 6.26 TRANSFERÊNCIA MÓVEL ENTRE REDES COM ROTEAMENTO DIRETO



6.6 IP MÓVEL

A arquitetura e os protocolos da Internet para suporte de mobilidade, conhecidos como IP móvel, estão definidos principalmente no RFC 5944 para IPv4. O IP móvel é um protocolo flexível que suporta muitos modos de operação diferentes (por exemplo, operação com ou sem um agente externo), várias maneiras para agentes e nós móveis descobrirem uns aos outros, utilização de um único COA ou de vários COAs e diversas formas de encapsulamento. Por isso, o IP móvel é um protocolo complexo, cuja descrição detalhada exigiria um livro inteiro; um desses livros é Perkins [1998b]. Aqui, nossa modesta meta é oferecer uma visão geral dos aspectos mais importantes do IP móvel e ilustrar sua utilização em alguns cenários comuns.

A arquitetura do IP móvel contém muitos dos elementos que acabamos de considerar, incluindo os conceitos de agentes nativos, agentes externos, endereços administrados e encapsulamento/descapsulamento. O padrão atual [RFC 5944] especifica a utilização de roteamento indireto para o nó móvel.

O padrão IP móvel consiste em três partes principais:

- *Descoberta de agente.* O IP móvel define os protocolos utilizados por um agente nativo ou por um agente externo para anunciar seus serviços a nós móveis e protocolos para que os nós móveis solicitem os serviços de um agente externo ou nativo.
- *Registro no agente nativo.* O IP móvel define os protocolos usados pelo nó móvel e/ou agente externo para registrar e anular os registros de COAs no agente local de um nó móvel.
- *Roteamento indireto de datagramas.* O padrão também define a maneira pela qual datagramas são repassados para nós móveis por um agente nativo, incluindo regras para repassar datagramas, regras para manipular condições de erro e diversas formas de encapsulamento [RFC, 2003; RFC, 2004].

Considerações de segurança têm destaque em todo o padrão IP móvel. Por exemplo, a autenticação de um nó móvel é claramente necessária para impedir que um usuário mal-intencionado registre no agente nativo um falso endereço administrado, o que poderia fazer datagramas endereçados a um endereço IP serem direcionados ao usuário mal-intencionado. O IP móvel consegue segurança usando muitos dos mecanismos que examinaremos no Capítulo 8, portanto, não consideraremos a questão da segurança de endereços na discussão a seguir.

Descoberta de agente

Um nó IP móvel que está chegando a uma nova rede, esteja ele se conectando a uma rede externa ou retornando à sua rede nativa, tem de aprender a identidade do correspondente externo ou do agente nativo. De fato, é a descoberta de um novo agente externo, com um novo endereço de rede, que permite à camada de rede em um nó móvel descobrir que ele passou para uma nova rede externa. Esse processo é conhecido como **descoberta de agente**. A descoberta de agente pode ser realizada de duas maneiras: via anúncio de agente ou via solicitação de agente

No caso do anúncio de agente, um agente externo ou nativo anuncia seus serviços usando uma extensão do protocolo de descoberta de roteador existente [RFC 1256]. O agente transmite periodicamente por difusão uma mensagem ICMP tendo 9 no campo de tipo (descoberta de roteador) em todos os enlaces aos quais está conectado. A mensagem de descoberta de roteador contém o endereço IP do roteador (isto é, o agente), e isso permite que um nó móvel descubra o endereço IP do agente. A mensagem de descoberta de roteador também contém uma extensão de anúncio de agente de mobilidade que guarda informações adicionais de que o nó móvel necessita. Entre os campos mais importantes na extensão estão os seguintes:

- *Bit do agente nativo (H).* Indica que o agente é um agente nativo para a rede na qual reside.
- *Bit de agente externo (F).* Indica que o agente é um agente externo para a rede na qual reside.
- *Bit de registro obrigatório (R).* Indica que um usuário móvel nessa rede *deverá* se registrar em um agente externo. Em particular, um usuário móvel não pode obter um endereço administrado na rede externa

(por exemplo, usando DHCP) e assumir a funcionalidade de agente externo para si mesmo sem se registrar no agente externo.

- *Bits de encapsulamento M, G.* Indicam se será utilizada uma forma de encapsulamento que não seja o encapsulamento IP em IP.
- *Campos de endereços aos cuidados (COA).* Uma lista de um ou mais endereços aos cuidados, fornecida pelo agente externo. No exemplo logo a seguir, o COA estará associado com o agente externo, que receberá datagramas enviados ao COA e então os repassará para o nó móvel adequado. O usuário móvel selecionará um desses endereços como seu COA ao se registrar no seu agente nativo.

A Figura 6.27 ilustra alguns dos principais campos na mensagem de anúncio de agente.

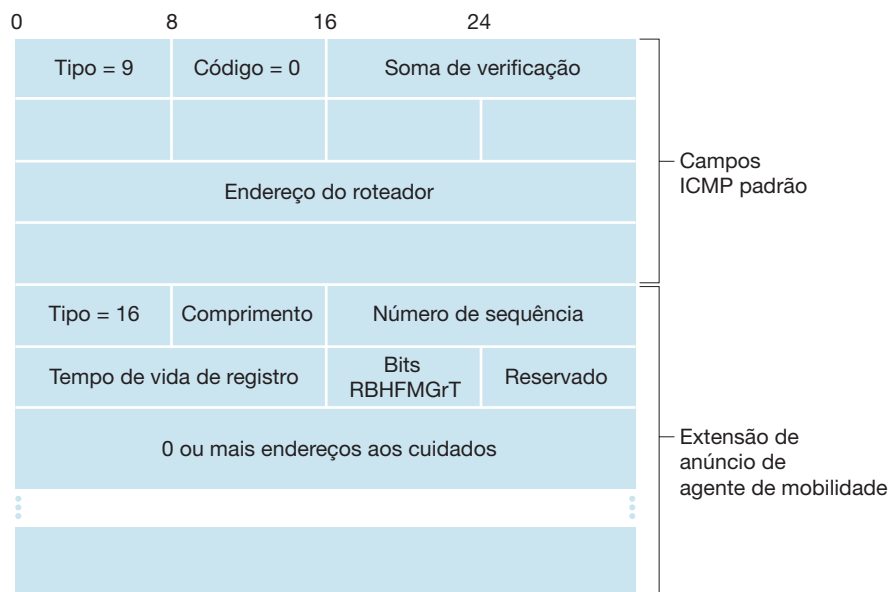
Com solicitação de agente, um nó móvel que quer conhecer agentes sem esperar para receber um anúncio de agente pode transmitir uma mensagem de solicitação em difusão, que é nada mais que uma mensagem ICMP cujo valor de tipo é 10. Ao receber a solicitação, um agente transmite um anúncio individual de agente diretamente ao nó móvel, que, então, procederá como se tivesse recebido um anúncio não solicitado.

Registro no agente nativo

Tão logo um nó móvel IP tenha recebido um COA, o endereço deve ser registrado no agente nativo, o que pode ser feito por meio do agente externo (que, então, registra o COA no agente nativo) ou diretamente pelo próprio nó móvel IP. A seguir, consideramos o primeiro caso. Há quatro etapas envolvidas.

1. Após o recebimento de um anúncio de agente externo, um nó móvel envia uma mensagem de registro IP móvel ao agente externo. A mensagem de registro é transportada dentro de um datagrama UDP e enviada à porta 434. A mensagem de registro leva um COA anunciado pelo agente externo, o endereço do agente nativo (HA), o endereço permanente do nó móvel (MA), o tempo de vida de registro requerido e uma identificação de registro de 64 bits. O tempo de vida de registro requerido é o número de segundos durante os quais o registro será válido. Se o registro não for renovado no agente nativo dentro do tempo de vida especificado, ele se tornará inválido. O identificador de registro age como um número de sequência e serve para combinar uma resposta de registro recebida com uma solicitação de registro, como discutiremos mais adiante.

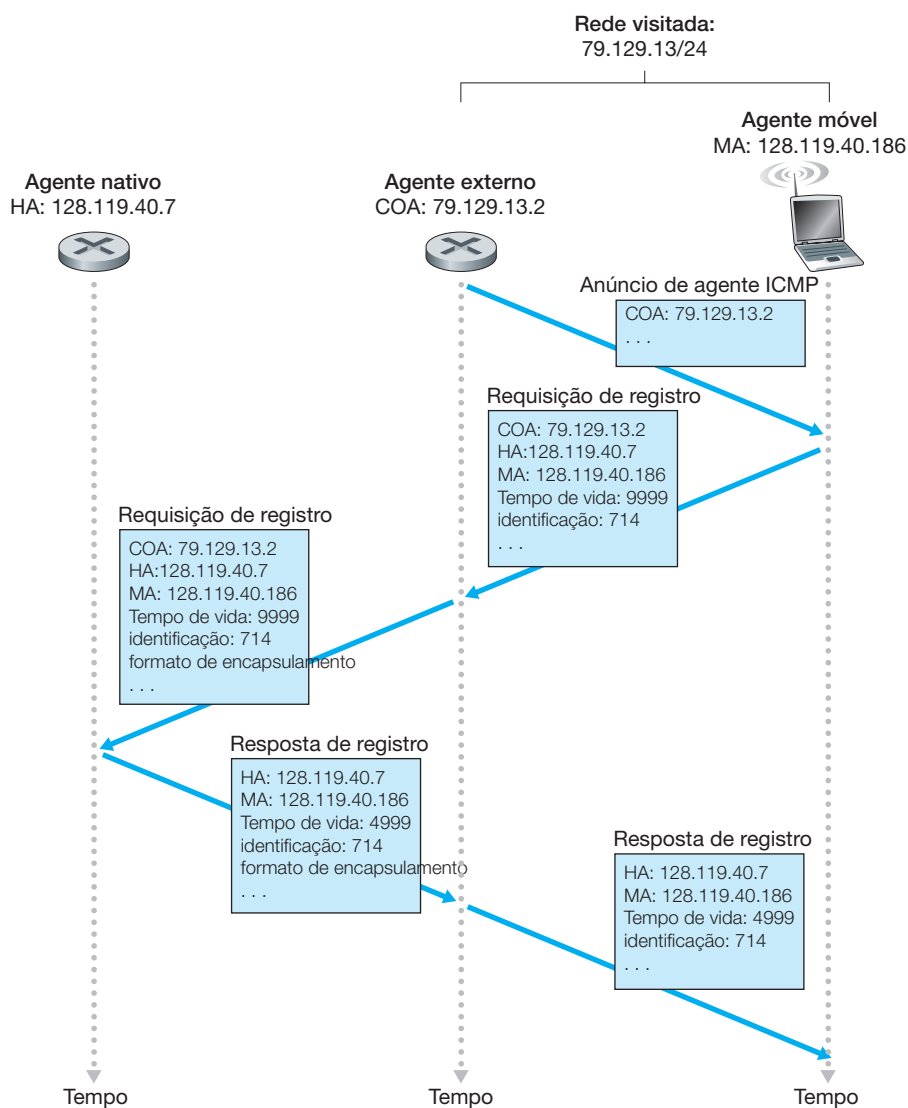
FIGURA 6.27 MENSAGEM ICMP DE DESCOBERTA DE ROTEADOR COM EXTENSÃO DE ANÚNCIO DE AGENTE DE MOBILIDADE



2. O agente externo recebe a mensagem de registro e registra o endereço IP permanente do nó móvel. Agora o agente externo sabe que deve procurar datagramas que contenham um datagrama encapsulado cujo endereço de destino combine com o endereço permanente do nó móvel. O agente externo, então, envia uma mensagem de registro do IP móvel (novamente, dentro de um datagrama UDP) à porta 434 do agente nativo. A mensagem contém o COA, o HA, o MA, o formato de encapsulamento requisitado, o tempo de vida de registro requisitado e a identificação do registro.
3. O agente nativo recebe a requisição de registro e verifica sua autenticidade e exatidão. O agente nativo vincula o endereço IP permanente do nó móvel ao COA; no futuro, datagramas que chegarem ao agente nativo e endereçados ao nó móvel serão encapsulados e enviados por túnel até o COA. O agente nativo envia uma resposta de registro IP móvel contendo o HA, o MA, o tempo de vida de registro vigente e a identificação de registro da solicitação que está sendo atendida com essa resposta.
4. O agente externo recebe a resposta de registro e então a repassa ao nó móvel.

Nesse ponto o registro está concluído e o nó móvel pode receber datagramas enviados a seu endereço permanente. A Figura 6.28 ilustra essas etapas. Note que o agente nativo especifica um tempo de vida menor do que aquele requisitado pelo nó móvel.

FIGURA 6.28 ANÚNCIO DE AGENTE E REGISTRO IP MÓVEL



Um agente externo não precisa anular explicitamente um registro de COA quando um nó móvel sai da rede. Isso ocorrerá automaticamente quando o nó móvel passar para uma nova rede (seja outra rede externa, seja sua rede nativa) e registrar um novo COA.

O padrão IP móvel permite muitos cenários e capacidades adicionais além das que acabamos de descrever. O leitor interessado deve consultar Perkins [1998b]; RFC [5944].

6.7 GERENCIAMENTO DE MOBILIDADE EM REDES CELULARES

Agora que acabamos de examinar como a mobilidade é gerenciada em redes IP, vamos voltar nossa atenção para redes cujo histórico de suporte à mobilidade é ainda mais longo — redes de telefonia celular. Enquanto na Seção 6.4 focalizamos o enlace sem fio do primeiro salto em redes celulares, aqui focalizaremos a mobilidade utilizando a arquitetura de rede celular GSM [Goodman, 1997; Mouly, 1992; Scourias, 2012; Kaaranen, 2001; Korhonen, 2003; Turner, 2012] como objeto de nosso estudo, visto que é uma tecnologia madura e amplamente disponibilizada. Como no caso do IP móvel, veremos que vários dos princípios fundamentais que identificamos da Seção 6.5 estão incorporados à arquitetura de rede do GSM.

Do mesmo modo que o IP móvel, o GSM adota uma abordagem de roteamento indireto (veja Seção 6.5.2), primeiro roteando a chamada do correspondente para a rede nativa do nó móvel e daí para a rede visitada. Em terminologia GSM, a rede nativa do nó móvel é denominada **rede pública terrestre móvel nativa (PLMN nativa)**. Visto que o acrônimo PLMN é um pouco complicado, e sempre firmes na nossa decisão de evitar uma sopa de letrinhas, denominaremos a rede PLMN nativa GSM apenas **rede nativa**. A rede nativa é a operadora de celular da qual o usuário móvel é assinante (isto é, a operadora que cobra mensalmente do usuário pelo serviço celular). A PLMN visitada, que denominaremos **rede visitada**, é a rede na qual o nó móvel estiver residindo.

Como no caso do IP móvel, as responsabilidades das redes nativas e visitadas são bastante diferentes.

- A rede nativa mantém um banco de dados conhecido como **registro nativo de localização** (*home location register* — HLR), que contém o número permanente do telefone celular e as informações do perfil do assinante para cada assinante. O importante é que o HLR também contém informações sobre as localizações atuais desses assinantes. Isto é, se um usuário móvel estiver em trânsito pela rede celular de outra operadora, o HLR conterá informações suficientes para obter (por meio de um processo que descreveremos em breve) um endereço na rede visitada para o qual deverá ser roteada uma chamada ao usuário móvel. Como veremos, quando é feita uma chamada para um usuário móvel, um comutador especial na rede nativa, conhecido como **Central de Comutação para Portal de Serviços Móveis** (*Gateway Mobile Services Switching Center* — GMSC), é contatado por um correspondente. Mais uma vez, conforme nossa decisão de evitar a sopa de letrinhas, denominaremos a GMSC por um termo mais descritivo, uma **MSC nativa**.
- A rede visitada mantém um banco de dados conhecido como **registro de localização de visitantes** (*visitor location register* — VLR). O VLR contém um registro para cada usuário móvel que está *atualmente* na parte da rede atendida pelo VLR. Assim, os registros do VLR vêm e vão à medida que usuários entram e saem da rede. Um VLR normalmente está localizado juntamente com a central de comutação móvel (MSC) que coordena o estabelecimento de uma chamada de e para a rede visitada.

Na prática, a rede celular de uma operadora servirá como uma rede nativa para seus assinantes e como uma rede visitada para usuários móveis que são assinantes de outras operadoras de serviços celulares.

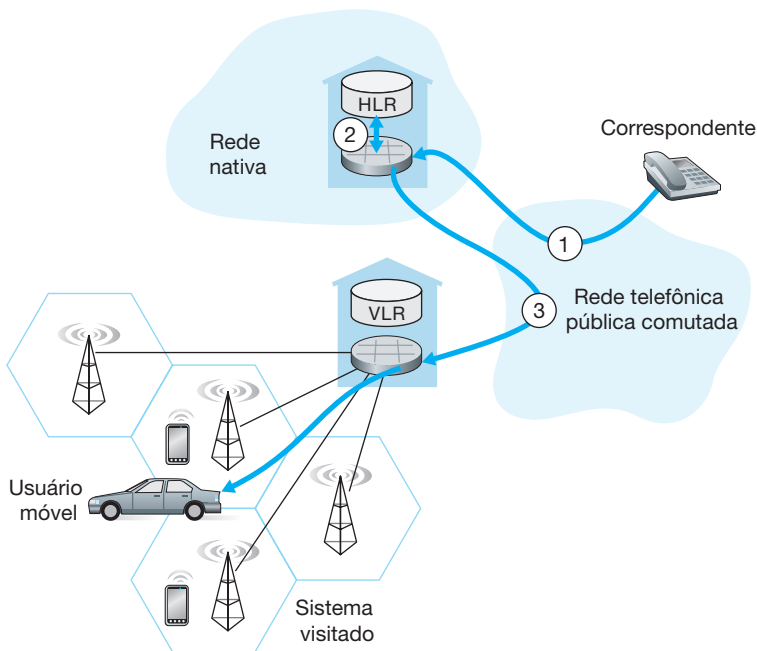
6.7.1 Roteando chamadas para um usuário móvel

Agora estamos prontos para descrever como é estabelecida uma chamada para um usuário GSM em uma rede visitada. A seguir, consideraremos um exemplo simples; cenários mais complexos são descritos em Mouly [1992]. As etapas, como ilustradas na Figura 6.29, são as seguintes:

1. O correspondente discar o número do telefone do usuário móvel. Esse número, em si, não se refere a uma determinada linha telefônica ou localização (afinal, o número do telefone é fixo, mas o usuário é móvel!). Os dígitos iniciais do número são suficientes para identificar globalmente a rede nativa do usuário móvel. A chamada é roteada desde o correspondente, passa através do PSTN e chega até a MSC na rede nativa do usuário móvel. Esse é o primeiro trecho da chamada.
2. A MSC nativa recebe a chamada e interroga o HLR para determinar a localização do usuário móvel. No caso mais simples, o HLR retorna o **número roaming da estação móvel** (*mobile station roaming number* — MSRN), que denominaremos **número de roaming**. Note que esse número é diferente do número permanente do telefone móvel, que é associado com a rede nativa do usuário móvel. O número de *roaming* é efêmero: é designado temporariamente a um usuário móvel quando ele entra em uma rede visitada. O número de *roaming* desempenha um papel semelhante ao do endereço COA no IP móvel e, da mesma forma, é invisível para o correspondente e para o usuário móvel. Se o HLR não tiver o número de *roaming*, ele retornará ao endereço do VLR na rede visitada. Nesse caso (que não é mostrado na Figura 6.29), a MSC nativa precisará consultar o VLR para obter o número de *roaming* do nó móvel. Mas, antes de tudo, como o HLR obtém o número de *roaming* ou o endereço VLR? O que acontece a esses valores quando o usuário móvel passa para outra rede visitada? Em breve consideraremos essas perguntas importantes.
3. Dado o número de *roaming*, a MSC nativa estabelece o segundo trecho da chamada através da rede até a MSC na rede visitada. A chamada está concluída — foi roteada do correspondente até a MSC nativa e daí para a MSC visitada, e desta até a estação-base que atende ao usuário móvel.

Uma questão não resolvida na etapa 2 é como o HLR obtém informação sobre a localização do usuário móvel. Quando um telefone móvel é ligado ou entra em uma parte de uma rede visitada que é coberta por um novo VLR, ele deve se registrar na rede visitada. Isso é feito por meio da troca de mensagens de sinalização entre o usuário móvel e o VLR. O VLR visitado, por sua vez, envia uma mensagem de requisição de atualização de localização ao HLR do usuário móvel. Essa mensagem informa ao HLR o número de *roaming* no qual o usuário móvel pode ser contatado ou o endereço do VLR (que então pode ser consultado mais tarde para obter esse número). Como parte dessa troca, o VLR também obtém do HLR informações sobre o assinante do usuário móvel e determina quais serviços (se houver algum) devem ser prestados a ele pela rede visitada.

FIGURA 6.29 ESTABELECENDO UMA CHAMADA PARA UM USUÁRIO MÓVEL: ROTEAMENTO INDIRETO



6.7.2 Transferências (*handoffs*) em GSM

Uma **transferência** (*handoff*) ocorre quando uma estação móvel muda sua associação de uma estação-base para outra durante uma chamada. Como ilustra a Figura 6.30, uma chamada de telefone móvel é roteada de início (antes da transferência) para o usuário móvel por meio de uma estação-base (a qual denominaremos antiga estação-base) e, após, por meio de outra estação-base (a qual denominaremos nova estação-base). Note que uma transferência entre estações-base resulta não apenas em transferência/recepção de/para um telefone móvel e uma nova estação-base, mas também no redirecionamento da chamada em curso de um ponto de comutação dentro da rede para a nova estação-base. Vamos admitir inicialmente que as estações-base antiga e nova compartilham a mesma MSC e que o redirecionamento ocorre nessa MSC.

Há diversas razões possíveis para ocorrer transferência, incluindo (1) o sinal entre a estação-base corrente e o usuário móvel pode ter-se deteriorado a tal ponto que a chamada corre perigo de “cair” e (2) uma célula pode ter ficado sobrecarregada, manipulando grande número de chamadas. Esse congestionamento pode ser aliviado transferindo usuários móveis para células próximas, menos congestionadas.

Enquanto está associado com uma estação-base, um usuário móvel mede periodicamente a potência de um sinal de sinalização emitido por sua estação-base corrente, bem como de sinais de sinalização emitidos por estações-base próximas que ele pode “ouvir”. Essas medições são passadas uma ou duas vezes por segundo para a estação-base corrente do usuário móvel. A transferência em GSM é iniciada pela estação-base antiga com base nessas medições, nas cargas correntes de usuários móveis em células próximas e outros fatores [Mouly, 1992]. O padrão GSM não determina o algoritmo específico a ser utilizado por uma estação-base para decidir se realiza ou não uma transferência.

A Figura 6.31 ilustra as etapas envolvidas quando uma estação-base decide transferir um usuário móvel:

1. A antiga estação-base (BS) informa à MSC visitada que deve ser feita uma transferência e a BS (ou possível conjunto de BSs) para a qual o usuário móvel deve ser transferido.
2. A MSC visitada inicia o estabelecimento do caminho até a nova BS, alocando os recursos necessários para carregar a chamada redirecionada e sinalizando à nova BS que uma transferência está prestes a ocorrer.
3. A nova BS reserva e ativa um canal de rádio para ser utilizado pelo usuário móvel.
4. A nova BS devolve um sinal à MSC visitada e à antiga BS, indicando que foi estabelecido um caminho entre a MSC visitada e a nova BS e que o usuário móvel deve ser informado da transferência iminente. A nova BS provê todas as informações de que o usuário móvel necessitará para se associar com a nova BS.
5. O usuário móvel é informado de que deve realizar uma transferência. Note que, até esse ponto, ele está totalmente desavisado de que a rede estava preparando o terreno para uma transferência (por exemplo, reservando um canal na nova BS e alocando um caminho entre a MSC visitada e a nova BS).
6. O usuário móvel e a nova BS trocam uma ou mais mensagens para ativar totalmente o novo canal na nova BS.

FIGURA 6.30 CENÁRIO DE TRANSFERÊNCIA ENTRE ESTAÇÕES-BASE QUE TÊM UMA MSC EM COMUM

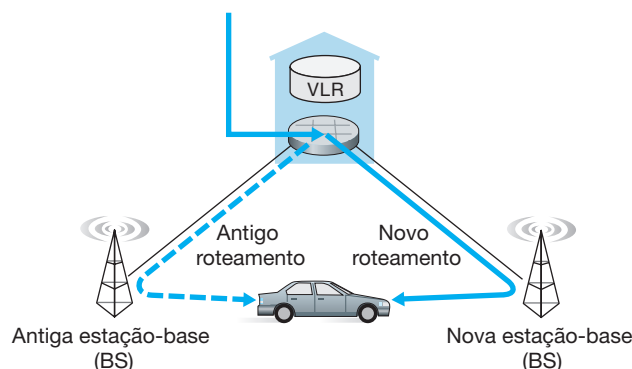
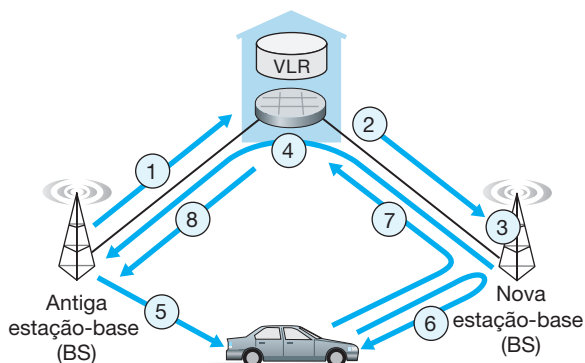


FIGURA 6.31 ETAPAS DA EXECUÇÃO DE UMA TRANSFERÊNCIA ENTRE ESTAÇÕES-BASE QUE TÊM UMA MSC EM COMUM

7. O móvel envia à nova BS uma mensagem de conclusão de transferência que é repassada para a MSC visitada. Então, a MSC visitada redireciona a chamada em curso para o usuário móvel, por meio da nova BS.
8. Os recursos reservados ao longo do caminho até a antiga BS são liberados.

Vamos concluir nossa discussão sobre transferência considerando o que acontece quando o usuário móvel passa para uma BS associada com uma MSC *diferente* da antiga BS e o que acontece quando essa transferência entre MSCs ocorre mais de uma vez. Como ilustra a Figura 6.32, o GSM define a noção de uma MSC âncora. A MSC âncora é a MSC visitada pelo usuário móvel logo no início de uma chamada; assim, ela não muda durante a chamada. Por toda a duração da chamada e independentemente do número de transferências entre MSCs realizadas pelo usuário móvel, a chamada é roteada desde a MSC nativa até a âncora e, então, desta até a MSC visitada, onde o usuário móvel está localizado no momento. Quando um usuário móvel passa da área de cobertura de uma MSC para a área de cobertura de outra, a chamada em curso é redirecionada desde a MSC âncora até a nova MSC visitada, que contém a nova estação-base. Assim, durante o tempo todo há, no máximo, três MSCs

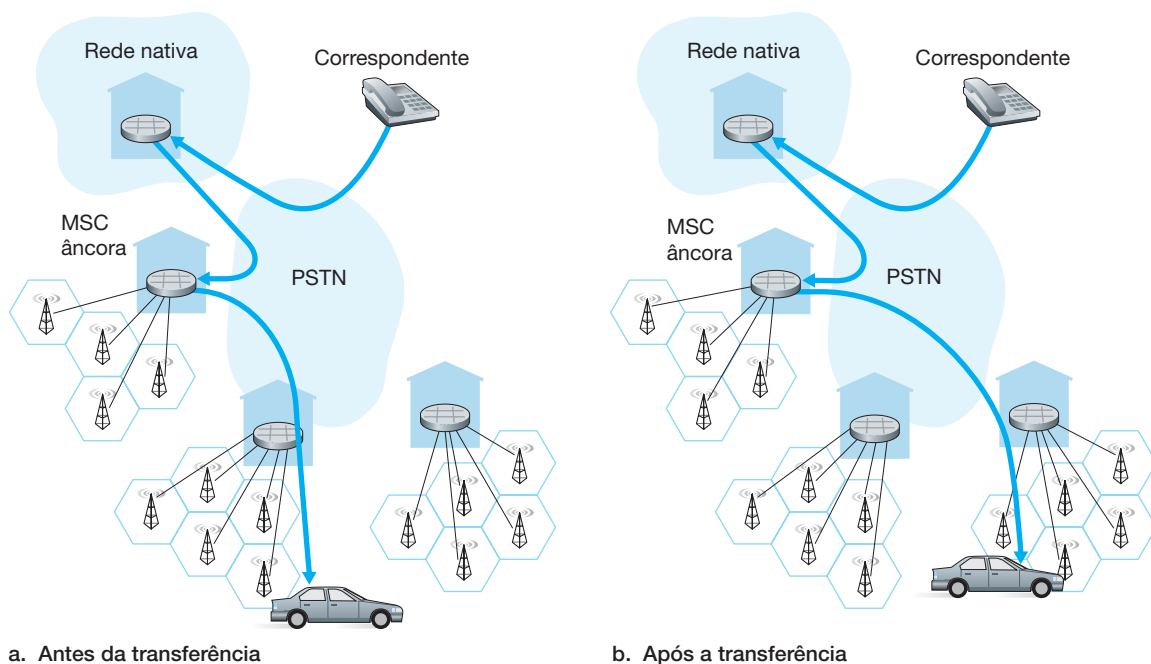
FIGURA 6.32 REDIRECIONAMENTO VIA A MSC ÂNCORA

TABELA 6.2 SEMELHANÇAS ENTRE A MOBILIDADE EM IP MÓVEL E EM GSM

Elemento do GSM	Comentário sobre o elemento do GSM	Elemento do IP móvel
Sistema nativo	Rede à qual pertence o número de telefone permanente do usuário.	Rede nativa
Central de comutação de unidade móvel ou simplesmente MSC nativa, Registro nativo de localização (HLR)	MSC nativa: ponto de contato para obter endereço roteável de usuário móvel. HLR: banco de dados no sistema nativo que contém número de telefone permanente, informações de perfil, localização corrente de usuário móvel, informações de assinatura.	Agente nativo
Sistema visitado	Rede, exceto o sistema nativo, onde o usuário móvel está residindo.	Rede visitada
Central de serviços de comutação de unidade móvel visitada, Registro de Localização de Visitante (VLR)	MSC visitada: responsável por estabelecer chamadas de/ para nós móveis em células associadas com MSC. VLR: registro temporário em banco de dados em sistema visitado, contendo informações de assinatura para cada usuário móvel visitado.	Agente externo
Número de <i>roaming</i> de estação móvel (MSRN) ou simplesmente número de roaming	Endereço roteável para segmento de chamada telefônica entre MSC nativa e MSC visitada, que não é visível nem para o usuário móvel nem para o correspondente.	Endereço administrado

(a nativa, a âncora e a visitada) entre o correspondente e o usuário móvel. A Figura 6.32 ilustra o roteamento de uma chamada entre as MSCs visitadas por um usuário móvel.

Em vez de manter um único salto desde a MSC âncora até a MSC corrente, uma técnica alternativa seria encadear as MSCs visitadas pelo móvel, fazendo uma MSC antiga transmitir a chamada em curso até a nova MSC cada vez que o móvel passa para uma nova MSC. Esse encadeamento de MSC pode de fato ocorrer em redes celulares IS-41, com uma etapa opcional de minimização de caminho para remover MSCs entre a âncora e a nova visitada [Lin, 2001].

Vamos encerrar nossa discussão sobre o gerenciamento da mobilidade GSM fazendo uma comparação entre gerenciamento de mobilidade em GSM e em IP Móvel. A comparação apresentada na Tabela 6.2 indica que, embora redes IP e celulares sejam em essência diferentes em muitos aspectos, compartilham um número surpreendente de elementos funcionais comuns e abordagens gerais para o tratamento da mobilidade.

6.8 SEM FIO E MOBILIDADE: IMPACTO SOBRE PROTOCOLOS DE CAMADAS SUPERIORES

Neste capítulo, vimos que redes sem fio são significativamente diferentes de suas contrapartes cabeadas tanto na camada de enlace (como resultado de características de canais sem fio como desvanecimento, propagação multivias e terminais ocultos) quanto na camada de rede (como resultado de usuários móveis que mudam seus pontos de conexão com a rede). Mas há diferenças importantes nas camadas de transporte e de aplicação? É tentador pensar que essas diferenças seriam pequenas, visto que a camada de rede provê o mesmo modelo de serviço de entrega de melhor esforço às camadas superiores tanto em redes cabeadas quanto em redes sem fio. De modo semelhante, se protocolos como TCP ou UDP são usados para oferecer serviços da camada de transporte a aplicações tanto em redes cabeadas como em redes sem fio, então a camada de aplicação também deve permanecer inalterada. Nossa intuição está certa em um sentido — TCP e UDP podem operar, e de fato operam, em redes com enlaces sem fio. Por outro lado, protocolos de transporte em geral e o TCP em particular às vezes podem ter desempenhos muito diferentes em redes cabeadas e em redes sem fio, e é neste particular, em termos de desempenho, que as diferenças se manifestam. Vejamos por quê.

Lembre-se de que o TCP retransmite um segmento que é perdido ou corrompido no caminho entre remetente e destinatário. No caso de usuários móveis, a perda pode resultar de congestionamento de rede (estouro de buffer de roteador) ou de transferência (por exemplo, de atrasos no redirecionamento de segmentos para um novo ponto de conexão do usuário à rede). Em todos os casos, o ACK do destinatário ao remetente do TCP indica apenas que um segmento não foi recebido intacto; o remetente não sabe se o segmento foi perdido por congestionamento, durante a transferência, ou por erros de bits detectados. Em todos os casos, a resposta do remetente é a mesma — retransmitir o segmento. A resposta do controle de congestionamento do TCP *também* é a mesma em todos os casos — o TCP reduz sua janela de congestionamento, como discutimos na Seção 3.7. Reduzindo de modo incondicional sua janela de congestionamento, o TCP admite implicitamente que a perda de segmento resulta de congestionamento e não de corrupção ou transferência. Vimos na Seção 6.2 que erros de bits são muito mais comuns em redes sem fio do que nas cabeadas. Quando ocorrem esses erros de bits ou quando há perda na transferência, na realidade não há razão alguma para que o remetente TCP reduza sua janela de congestionamento (reduzindo, assim, sua taxa de envio). Na verdade, é bem possível que os buffers de roteador estejam vazios e que pacotes estejam fluindo ao longo de caminhos fim a fim desimpedidos, sem congestionamento.

Entre o início e meados da década de 1990, pesquisadores perceberam que, dadas as altas taxas de erros de bits em enlaces sem fio e a possibilidade de perdas pela transferência de usuários, a resposta do controle de congestionamento do TCP poderia ser problemática em um ambiente sem fio. Há duas classes gerais de abordagens possíveis para tratar esse problema:

- *Recuperação local.* Os protocolos de recuperação local recuperam erros de bits quando e onde (por exemplo, no enlace sem fio) eles ocorrem (por exemplo, o protocolo ARQ 802.11, que estudamos na Seção 6.3, ou técnicas mais sofisticadas que utilizam ARQ e também FEC [Ayanoglu, 1995]).
- *Remetente TCP ciente de enlaces sem fio.* Em técnicas de recuperação locais, o remetente TCP fica completamente desavisado de que seus segmentos estão atravessando um enlace sem fio. Uma técnica alternativa é o remetente e o destinatário ficarem cientes da existência de um enlace sem fio, para distinguir entre perdas por congestionamento na rede cabeada e corrupção/perdas no enlace sem fio, e invocar o controle de congestionamento somente em resposta a perdas por congestionamento na rede cabeada. Balakrishnan [1997] investiga vários tipos de TCP, supondo que sistemas finais possam fazer essa distinção. Liu [2003] investiga técnicas para distinguir entre perdas nos segmentos cabeado e sem fio de um caminho fim a fim.
- *Técnicas de conexão dividida.* Nesta técnica de conexão dividida [Bakre, 1995], a conexão fim a fim entre o usuário móvel e o outro ponto terminal é dividida em duas conexões da camada de transporte: uma do hospedeiro móvel ao ponto de acesso sem fio, e uma do ponto de acesso sem fio ao outro ponto terminal de comunicação (admitiremos, aqui, um usuário cabeado). A conexão fim a fim é, então, formada por uma concatenação de uma parte sem fio e uma parte cabeada. A camada de transporte sobre um segmento sem fio pode ser uma conexão-padrão TCP [Bakre, 1995], ou principalmente um protocolo de recuperação de erro personalizado em cima do UDP. Yavatkar [1994] analisa o uso de um protocolo de repetição seletiva da camada de transporte por uma conexão sem fio. As medidas relatadas em Wei [2006] indicam que conexões TCP divididas são bastante usadas em redes de dados celulares, e que aperfeiçoamentos significativos podem ser feitos com o uso dessas conexões.

Aqui, nosso tratamento do TCP em enlaces sem fio foi necessariamente breve. Estudos mais aprofundados dos desafios e soluções do TCP nessas redes podem ser encontrados em Hanabali [2005]; Leung [2006]. Aconselhamos o leitor a consultar as referências se quiser mais detalhes sobre essa área de pesquisa em curso.

Agora que já consideramos protocolos de camada de transporte, vamos analisar em seguida o efeito do sem fio e da mobilidade sobre protocolos da camada de aplicação. Uma consideração importante aqui é que enlaces sem fio muitas vezes têm larguras de banda relativamente baixas, como vimos na Figura 6.2. Por conseguinte, aplicações que operam por enlaces sem fio, em particular por enlaces celulares sem fio, devem

tratar a largura de banda como uma mercadoria escassa. Por exemplo, um servidor Web que serve conteúdo a um navegador Web que está rodando em um telefone 3G talvez não consiga prover o mesmo conteúdo rico em imagens que oferece a um navegador que está rodando sobre uma conexão cabeada. Embora enlaces sem fio proponham desafios na camada de aplicação, a mobilidade que eles criam também torna possível um rico conjunto de aplicações cientes de localização e de contexto [Chen, 2000; Baldauf, 2007]. Em termos mais gerais, redes sem fio e redes móveis desempenharão um papel fundamental na concretização dos ambientes de computação onipresentes do futuro [Weiser, 1991]. É justo dizer que vimos somente a ponta do iceberg quando se trata do impacto de redes sem fio e móveis sobre aplicações em rede e seus protocolos!

6.9 RESUMO

Redes sem fio e móveis revolucionaram a telefonia e também estão causando um impacto cada vez mais profundo no mundo das redes de computadores. Com o acesso à infraestrutura da rede global que oferecem — desimpedido, a qualquer hora, em qualquer lugar — estão não só aumentando a onipresença do acesso a redes, mas também habilitando um novo conjunto muito interessante de serviços dependentes de localização. Dada a crescente importância das redes sem fio e móveis, este capítulo focalizou os princípios, as tecnologias de enlace e as arquiteturas de rede para suportar comunicações sem fio e móveis.

Iniciamos o capítulo com uma introdução às redes sem fio e móveis, traçando uma importante distinção entre os desafios propostos pela natureza *sem fio* dos enlaces de comunicação desse tipo de rede, e pela *mobilidade* que permitem. Isso nos possibilitou isolar, identificar e dominar melhor os conceitos fundamentais em cada área. Focalizamos primeiro a comunicação sem fio, considerando as características de um enlace sem fio na Seção 6.2. Nas seções 6.3 e 6.4 examinamos os aspectos de camada de enlace do padrão IEEE 802.11 (Wi-Fi) para LANs sem fio, duas redes pessoais IEEE 802.15 (Bluetooth e Zigbee) e acesso à Internet pela rede celular 3G e 4G. Depois, voltamos nossa atenção para a questão da mobilidade. Na Seção 6.5, identificamos diversas formas de mobilidade e verificamos que há pontos nesse espectro que propõem desafios diferentes e admitem soluções diferentes. Consideramos os problemas de localização e roteamento para um usuário móvel, bem como técnicas para transferir o usuário móvel que passa dinamicamente de um ponto de conexão com a rede para outro. Examinamos como essas questões foram abordadas no padrão IP móvel e em GSM nas seções 6.6 e 6.7, respectivamente. Por fim, na Seção 6.8, consideramos o impacto causado por enlaces sem fio e pela mobilidade sobre protocolos de camada de transporte e aplicações em rede.

Embora tenhamos dedicado um capítulo inteiro ao estudo de redes sem fio e redes móveis, seria preciso todo um livro (ou mais) para explorar completamente esse campo tão animador e que está se expandindo tão depressa. Aconselhamos o leitor a se aprofundar mais nesse campo consultando as muitas referências fornecidas neste capítulo.

EXERCÍCIOS DE FIXAÇÃO E PERGUNTAS

Questões de revisão do Capítulo 6

SEÇÃO 6.1

- R1. O que significa para uma rede sem fio estar operando no “modo de infraestrutura”? Se a rede não estiver nesse modo, em qual modo ela está e qual é a diferença entre esse modo de operação e o de infraestrutura?
- R2. Quais são os quatro tipos de redes sem fio identificadas em nossa taxonomia na Seção 6.1? Quais desses tipos de rede sem fio você usou?

SEÇÃO 6.2

- R3. Quais são as diferenças entre os seguintes tipos de falhas no canal sem fio: atenuação de percurso, propagação multivias, interferência de outras fontes?
- R4. Um nó móvel se distancia cada vez mais de uma estação-base. Quais são as duas atitudes que uma estação-base poderia tomar para garantir que a probabilidade de perda de um quadro transmitido não aumente?

SEÇÕES 6.3-6.4

- R5. Descreva o papel dos quadros de sinalização em 802.11.
- R6. Verdadeiro ou falso: antes de uma estação 802.11 transmitir um quadro de dados, ela deve primeiro enviar um quadro RTS e receber um quadro CTS correspondente.
- R7. Por que são usados reconhecimentos em 802.11, mas não em Ethernet cabeada?
- R8. Verdadeiro ou falso: Ethernet e 802.11 usam a mesma estrutura de quadro.
- R9. Descreva como funciona o patamar RTS.
- R10. Suponha que os quadros RTS e CTS IEEE 802.11 fossem tão longos quanto os padronizados DATA e ACK. Haveria alguma vantagem em usar os quadros CTS e RTS? Por quê?
- R11. A Seção 6.3.4 discute mobilidade 802.11, na qual uma estação sem fio passa de um BSS para outro dentro da mesma sub-rede. Quando os APs estão interconectados com um comutador, um AP pode precisar enviar um quadro com um endereço MAC fingido para fazer o comutador transmitir quadros adequadamente. Por quê?
- R12. Quais são as diferenças entre o dispositivo mestre em uma rede Bluetooth e uma estação-base em uma rede 802.11?
- R13. O que significa um superquadro no padrão Zigbee 802.15.4?
- R14. Qual é a função do “núcleo da rede” na arquitetura de dados celular 3G?
- R15. Qual é a função do RNC na arquitetura da rede de dados celular 3G? Que função o RNC desempenha na rede celular de voz?

SEÇÕES 6.5-6.6

- R16. Se um nó tem uma conexão sem fio à Internet, ele precisa ser móvel? Explique. Suponha que um usuário portando um notebook ande com ele pela casa, e sempre acesse a Internet por meio do mesmo ponto de acesso. Em relação à rede, este é usuário móvel? Explique.
- R17. Qual é a diferença entre um endereço permanente e um endereço aos cuidados (COA)? Quem determina o endereço aos cuidados?
- R18. Considere uma conexão TCP através de um IP móvel. Falso ou verdadeiro: a fase da conexão TCP entre o correspondente e o hospedeiro móvel percorre a rede doméstica móvel, mas a fase de transferência de dados está diretamente entre o correspondente e o hospedeiro móvel, pulando a rede doméstica.

SEÇÃO 6.7

- R19. Quais são os objetivos do HLR e VLR nas redes GSM? Quais elementos de IP móvel são semelhantes ao HLR e ao VLR?
- R20. Qual é o papel da MSC âncora em redes GSM?

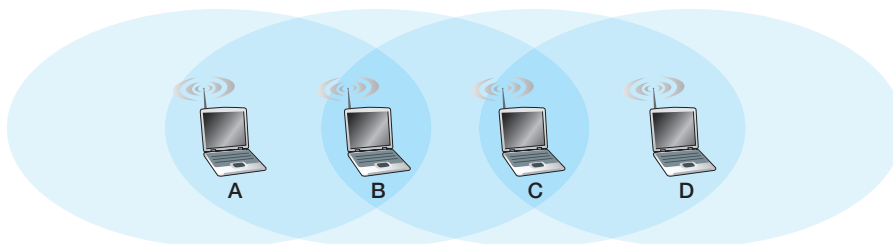
SEÇÃO 6.8

- R21. Quais são os três métodos que podem ser realizados para evitar que um único enlace sem fio reduza o desempenho de uma conexão TCP fim a fim da camada de transporte?

PROBLEMAS

- P1. Considere o exemplo do remetente CDMA único na Figura 6.5. Qual seria a saída do remetente (para os 2 bits de dados mostrados) se o código do remetente CDMA fosse $(1, -1, 1, -1, 1, -1, 1, -1)$?
- P2. Considere o remetente 2 na Figura 6.6. Qual é a saída do remetente para o canal (antes de ser adicionada ao sinal vindo do remetente 1), $Z_{i,m}^2$?
- P3. Suponha que o receptor na Figura 6.6 queira receber os dados que estão sendo enviados pelo remetente 2. Mostre (por cálculo) que o receptor pode, na verdade, recuperar dados do remetente 2 do sinal agregado do canal usando o código do remetente 2.
- P4. Para o exemplo sobre dois remetentes, dois destinatários, dê um exemplo de dois códigos CDMA contendo os valores 1 e -1, que não permitem que dois destinatários extraiam os bits originais transmitidos por dois remetentes CDMA.
- P5. Suponha que dois ISPs fornecem acesso Wi-Fi em um determinado local, e que cada um deles opera seu próprio AP e tem seu próprio bloco de endereços IP.
- Suponha ainda mais, que, por acidente, cada ISP configurou seu AP para operar no canal 11. O protocolo 802.11 falhará totalmente nessa situação? Discuta o que acontece quando duas estações, cada uma associada com um ISP diferente, tentam transmitir ao mesmo tempo.
 - Agora suponha que um AP opera no canal 1 e outro no canal 11. Como você mudaria suas respostas?
- P6. Na etapa 4 do protocolo CSMA/CA, uma estação que transmite um quadro com sucesso inicia o protocolo CSMA/CA para um segundo quadro na etapa 2, e não na 1. Quais seriam as razões que os projetistas do CSMA/CA provavelmente tinham em mente para fazer essa estação não transmitir o segundo quadro de imediato (se o canal fosse percebido como ocioso)?
- P7. Suponha que uma estação 802.11b seja configurada para sempre reservar o canal com a sequência RTS/CTS. Imagine que essa estação de repente queira transmitir 1.000 bytes de dados e que todas as outras estações estão ociosas nesse momento. Calcule o tempo requerido para transmitir o quadro e receber o reconhecimento como uma função de SIFS e DIFS, ignorando atraso de propagação e admitindo que não haja erros de bits.
- P8. Considere o cenário mostrado na Figura 6.33, no qual existem quatro nós sem fios, A, B, C e D. A cobertura de rádio dos quatro nós é mostrada pelas formas ovais mais escuras; todos os nós compartilham a mesma frequência. Quando A transmite, ele pode ser ouvido/recebido por B; quando B transmite, ele só pode ser ouvido/recebido por A e C; quando C transmite, B e D podem ouvir/receber de C; quando D transmite, somente C pode ouvir/receber de D.

FIGURA 6.33 CENÁRIO PARA O PROBLEMA P8



Agora suponha que cada nó possua um estoque infinito de mensagens que ele queira enviar para os outros nós. Se o destinatário da mensagem não for um vizinho imediato, então a mensagem deve ser retransmitida. Por exemplo, se A quer enviar para D, uma mensagem de A deve ser primeiro enviada a B, que, então, envia a mensagem a C, e este a D. O tempo é dividido em intervalos, com um tempo de transmissão de mensagem de

exatamente um intervalo de tempo, como em um *slotted* Aloha, por exemplo. Durante um intervalo, um nó pode fazer uma das seguintes opções: (i) enviar uma mensagem; (ii) receber uma mensagem (se, exatamente, uma mensagem estiver sendo enviada a ele), (iii) permanecer silencioso. Como sempre, se um nó ouvir duas ou mais transmissões simultâneas, ocorrerá uma colisão e nenhuma das mensagens transmitidas é recebida com sucesso. Você pode admitir aqui que não existem erros de bits e, dessa forma, se uma mensagem for enviada, ela será recebida corretamente pelos que estão dentro do raio de transmissão do emissor.

- a. Suponha que um controlador onisciente (ou seja, que sabe o estado de cada nó na rede) possa comandar cada nó a fazer o que ele (o controlador onisciente) quiser, isto é, enviar uma mensagem, receber uma mensagem, ou permanecer silencioso. Dado esse controlador onisciente, qual é a taxa máxima à qual uma mensagem de dados pode ser transferida de C para A, sabendo que não existem outras mensagens entre nenhuma outra dupla remetente/destinatária?
 - b. Suponha que A envie uma mensagem a B, e D envie uma mensagem a C. Qual é a taxa máxima combinada à qual as mensagens de dados podem fluir de A a B e de D a C?
 - c. Considere agora que A envie uma mensagem a B, e C envie uma mensagem a D. Qual é a taxa máxima combinada à qual as mensagens de dados podem fluir de A a B e de C a D?
 - d. Suponha agora que os enlaces sem fio sejam substituídos por enlaces cabeados. Repita as questões de “a” a “c” neste cenário cabeado.
 - e. Agora imagine que estamos de novo em um cenário sem fio e que para cada mensagem de dados enviada do remetente ao destinatário, este envie de volta uma mensagem ACK para o remetente (como no TCP, por exemplo). Suponha também que cada mensagem ACK possua um intervalo. Repita as questões de “a” a “c” para este cenário.
- P9. Descreva o formato do quadro Bluetooth 802.15.1. Você precisará de uma leitura complementar para encontrar essa informação. Existe algo no formato do quadro que basicamente limite o número de nós ativos para oito em uma rede 802.15.1? Explique.
- P10. Considere o seguinte cenário WiMAX ideal. O subquadro descendente (veja Figura 6.17) é dividido em intervalos de tempo, com N intervalos descendentes por subquadro, e todos os intervalos de tempo têm o mesmo comprimento. Existem quatro nós, A, B, C e D, alcançáveis da estação-base a taxas de 10 Mbits/s, 5 Mbits/s, 2,5 Mbits/s e 1 Mbit/s, respectivamente no canal descendente. A estação-base possui infinitos dados para enviar a cada nó e pode enviar para qualquer um dos quatro nós durante qualquer intervalo de tempo no subquadro descendente.
- a. Qual é a taxa máxima à qual a estação-base pode enviar aos nós, admitindo que ela pode enviar a qualquer nó de sua escolha durante cada intervalo de tempo? Sua solução é justa? Explique e defina o que você quis dizer com “justo”.
 - b. Se há requisito de equidade que todos os nós devem receber uma quantidade igual de dados durante cada quadro de *downstream*, qual é a taxa média de transmissão pela estação-base (para todos os nós) durante o subquadro de *downstream*? Explique como você chegou a essa resposta.
 - c. Suponha que, como critério de equidade, qualquer nó possa receber, no máximo, duas vezes tantos dados quanto qualquer outro nó durante o subquadro. Qual é a taxa média de transmissão pela estação-base (para todos os nós) durante o subquadro de *downstream*? Explique como você chegou a esta resposta.
- P11. Na Seção 6.5, uma solução proposta que permitia que usuários móveis mantivessem seu endereço IP à medida que transitavam entre redes externas era fazer uma rede externa anunciar ao usuário móvel uma rota altamente específica e usar a infraestrutura de roteamento existente para propagar essa informação por toda a rede. Uma das preocupações que identificamos foi a escalabilidade. Suponha que, quando um usuário móvel passa de uma rede para outra, a nova rede externa anuncie uma rota específica para o usuário móvel e a antiga rede externa retire sua rota. Considere como informações de roteamento se propagam em um algoritmo vetor de distâncias (em particular para o caso de roteamento interdomínios entre redes que abrangem o globo terrestre).

- a. Outros roteadores conseguirão rotear datagramas imediatamente para a nova rede externa tão logo essa rede comece a anunciar sua rota?
 - b. É possível que roteadores diferentes acreditem que redes externas diferentes contenham o usuário móvel?
 - c. Discuta a escala temporal segundo a qual outros roteadores na rede finalmente aprenderão o caminho até os usuários móveis.
- P12. Suponha que o correspondente na Figura 6.22 fosse móvel. Faça um desenho esquemático da infraestrutura adicional de camada de rede que seria necessária para rotear o datagrama do usuário móvel original até o correspondente (que agora é móvel). Mostre a estrutura do(s) datagrama(s) entre o usuário móvel original e o correspondente (agora móvel), como na Figura 6.23.
- P13. Em IP móvel, que efeito terá a mobilidade sobre atrasos fim a fim de datagramas entre a fonte e o destino?
- P14. Considere o exemplo de encadeamento discutido no final da Seção 6.7.2. Suponha que um usuário móvel visite as redes externas A, B e C, e que um correspondente inicie uma conexão com o usuário móvel enquanto este reside na rede externa A. Relacione a sequência de mensagens entre agentes externos e entre agentes externos e o agente nativo, enquanto o usuário passa da rede A para a rede B e para a rede C. Em seguida, suponha que não é executado encadeamento e que as mudanças no endereço administrado do usuário móvel devem ser notificadas explicitamente ao correspondente (bem como ao agente nativo). Relacione a sequência de mensagens que seria necessário trocar nesse segundo cenário.
- P15. Considere dois nós móveis em uma rede externa que tem um agente externo. É possível que esses nós utilizem o mesmo endereço aos cuidados em IP móvel? Explique sua resposta.
- P16. Quando discutimos como o VLR atualizava o HLR com informações sobre a localização corrente de usuários móveis, quais eram as vantagens e as desvantagens de fornecer ao HLR o MSRN em vez do endereço do VLR?

WIRESHARK LAB

No site de apoio do livro você encontrará um Wireshark Lab, em inglês, para este capítulo, que captura e estuda os quadros 802.11 trocados entre um notebook sem fio e um ponto de acesso.

ENTREVISTA



Deborah Estrin

Deborah Estrin é professora de ciência da computação na UCLA, Jon Postel Chair in Computer Networks, diretora do Center for Embedded Networked Sensing (CENS), e co-fundadora da organização sem fins lucrativos openmhealth.org. Ela recebeu doutorado (1985) em ciência da computação pelo MIT e bacharelado (1980) pela Universidade da Califórnia em Berkeley. O estudo inicial de Estrin foi voltado para o projeto de protocolos de rede, incluindo roteamento de transmissão para um grupo (*multicast*) e interdomínio. Em 2002, fundou o NSF-funded Science and Technology Center, CENS (<http://cens.ucla.edu>), para desenvolver e explorar tecnologias e aplicações de monitoração ambiental. Hoje, Estrin e seus colaboradores estão desenvolvendo sistemas de sensores participativos, aproveitando a programabilidade, proximidade e difusão dos telefones móveis; os principais contextos de implementação são saúde móvel (<http://openmhealth.org>), coleta de dados da comunidade e educação STEM (<http://mobilizingcs.org>). A professora Estrin é membro eleito da American Academy of Arts and Sciences (2007) e da National Academy of Engineering (2009). É *fellow* do IEEE, ACM e AAAS. Foi selecionada como a primeira ACM-W Athena Lecturer (2006), recebeu o prêmio Women of Vision Award for Innovation (2007) pelo Anita Borg Institute, entrou para o *hall* da fama do WITI (2008) e recebeu o prêmio Doctor Honoris Causa do EPFL (2008) e da Uppsala University (2011).