ББК 32.973.26-018.2.75
Ш58
УДК 681.3.07

Издательский дом "Вильямс"
Зав. редакцией *А.В. Слепцов*

Перевод с английского *О.О. Лисового, В.В. Марченко, А.В. Назаренко, М.Л. Шафира*

Под редакцией *А.В. Назаренко*

По общим вопросам обращайтесь в Издательский дом "Вильямс" по адресу:
info@williamspublishing.com, http://www.williamspublishing.com

Шиллер, Йоган.

Ш58 Мобильные коммуникации. : Пер. с англ. — М. : Издательский дом "Вильямс", 2002. — 384 с. : ил. — Парал. тит. англ.

ISBN 5-8459-0353-X (рус.) 4/5621

Книга является введением в область мобильных коммуникаций и посвящена вопросам цифровой передачи данных. В ней предпринимается попытка охватить различные аспекты мобильной связи, причем с точки зрения компьютерных наук. Более того, в книге рассматриваются общие свойства различных технических решений, показана интеграция услуг и приложений для стационарных сетей в сети, поддерживающие беспроводный доступ и мобильность конечных систем.

Книга предназначена для студентов факультетов электроники и компьютерных наук, изучающих курс компьютерных сетей и коммуникаций; для инженеров, работающих со стационарными сетями и интересующихся будущим сетевых технологий; а также для тех менеджеров, которые нуждаются в хорошем и достаточно полном обзоре существующих средств мобильной связи.

ББК 32.973.26-018.2.75

Все названия программных продуктов являются зарегистрированными торговыми марками соответствующих фирм.

Никакая часть настоящего издания ни в каких целях не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами, будь то электронные или механические, включая фотокопирование и запись на магнитный носитель, если на это нет письменного разрешения издательства Addison-Wesley UK.

Authorized translation from the English language edition published by Pearson Education Limited, Copyright © 2000

All rights reserved. No part of this book may be reproduced, stored in retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without either the prior written permission o the Publisher.

Russian language edition published by Williams Publishing House according to the Agreement with R&I Enterprises International, Copyright © 2002

ISBN 5-8459-0353-X (рус.)
ISBN 0-2013-9836-2 (англ.)

© Издательский дом "Вильямс", 2002
© Pearson Education Limited, 2000

Оглавление

Об авторе	12
Предисловие	13
Благодарности	15
Глава 1. Введение	17
Глава 2. Беспроводная передача	37
Глава 3. Управление доступом к среде	75
Глава 4. Системы телекоммуникаций	97
Глава 5. Спутниковые системы	143
Глава 6. Широковещательные системы	157
Глава 7. Беспроводные локальные сети	171
Глава 8. Беспроводная ATM	219
Глава 9. Мобильный сетевой уровень	255
Глава 10. Мобильный транспортный уровень	287
Глава 11. Поддержка мобильности	303
Приложение А. Сокращения	355
Предметный указатель	368

Содержание

Об авторе 12

Предисловие 13

Благодарности 15

Глава 1. Введение 17

1.1. Основные применения мобильной связи 18
1.1.1. Транспортные средства 18
1.1.2. Чрезвычайные ситуации 19
1.1.3. Бизнес 20
1.1.4. Замена кабельных соединений 20
1.1.5. Информационная поддержка 20
1.1.6. Услуги, зависящие от местонахождения 21
1.1.7. Мобильные и беспроводные устройства 22
1.2. Краткая история беспроводной связи 23
1.3. Рынок мобильных коммуникаций 28
1.4. Некоторые еще нерешенные задачи 28
1.5. Упрощенная эталонная модель 30
1.6. Общий обзор содержания книги 31
1.7. Упражнения 34
1.8. Рекомендуемая литература 35

Глава 2. Беспроводная передача 37

2.1. Частоты радиопередачи 37
2.1.1. Регулирование использования частот 39
2.2. Сигналы 41
2.3. Антенны 43
2.4. Распространение сигналов 45
2.4.1. Потери на пути распространения 46
2.4.2. Дополнительное влияние условий на распространение сигналов 47
2.4.3. Многолучевое распространение 48
2.5. Уплотнение 50
2.5.1. Уплотнение с пространственным разделением 50
2.5.2. Уплотнение с частотным разделением 52
2.5.3. Уплотнение с временным разделением 52
2.5.4. Уплотнение с кодовым разделением 53
2.6. Модуляция 55
2.6.1. Амплитудная манипуляция 57
2.6.2. Частотная манипуляция 57
2.6.3. Фазовая манипуляция 58
2.6.4. Улучшенная частотная манипуляция 58
2.6.5. Улучшенная фазовая манипуляция 59
2.6.6. Модуляция с несколькими несущими 60
2.7. Расширение спектра 61
2.7.1. Расширение спектра методом прямой последовательности 64
2.7.2. Расширенный спектр с перескоком частоты 66

2.8. Сотовые системы 68
2.9. Резюме 71
2.10. Упражнения 72
2.11. Рекомендуемая литература 73

Глава 3. Управление доступом к среде 75

3.1. Специализированные схемы MAC как необходимость 76
3.1.1. Скрытые и открытые терминалы 76
3.1.2. Ближние и дальние терминалы 77
3.2. SDMA 78
3.3. FDMA 78
3.4. TDMA 79
3.4.1. Фиксированная схема TDM 80
3.4.2. Классическая схема Aloha 81
3.4.3. Aloha с разделением 81
3.4.4. Множественный доступ с детектированием несущей 82
3.4.5. Множественный доступ с распределением по запросу 83
3.4.6. Множественный доступ PRMA с резервированием пакетов 83
3.4.7. Схема TDMA с резервированием 84
3.4.8. Множественный доступ с исключением конфликтов 84
3.4.9. Опрос 87
3.4.10. Множественный доступ с детектированием подавления 87
3.5. CDMA 87
3.5.1. Множественный доступ с расширением по схеме Aloha 92
3.6. Сравнение схем S/T/F/CDMA 93
3.7. Упражнения 95
3.8. Рекомендуемая литература 96

Глава 4. Системы телекоммуникаций 97

4.1. Система GSM 98
4.1.1. Мобильные службы 99
4.1.2. Архитектура системы 101
4.1.3. Радиоинтерфейс 105
4.1.4. Протоколы 109
4.1.5. Локализация и вызовы 112
4.1.6. Переключение 115
4.1.7. Безопасность 118
4.1.8. Новые службы передачи данных 120
4.2. Система DECT 126
4.2.1. Архитектура системы 127
4.2.2. Архитектура протоколов 128
4.3. TETRA 130
4.4. UMTS и IMT-2000 132
4.4.1. Основы архитектуры системы UMTS 134
4.4.2. Режим UTRA FDD 134
4.4.3. Режим UTRA TDD 136
4.5. Резюме 137
4.6. Упражнения 138
4.7. Рекомендуемая литература 140

Глава 5. Спутниковые системы 143
5.1. История 143
5.2. Сферы применения 144
5.3. Основы 146
5.3.1. Спутники GEO 150
5.3.2. Спутники LEO 150
5.3.3. Спутники MEO 151
5.4. Маршрутизация 151
5.5. Определение местонахождения 152
5.6. Переключение 152
5.7. Примеры 153
5.8. Резюме 154
5.9. Упражнения 155
5.10. Рекомендуемая литература 155

Глава 6. Широковещательные системы 157
6.1. Обзор 157
6.2. Циклическое повторение данных 159
6.3. Цифровое аудиовещание (DAB) 160
6.3.1. Протокол передачи мультимедийных объектов 162
6.4. Цифровое видеовещание (DVB) 164
6.5. Резюме 168
6.6. Упражнения 168
6.7. Рекомендуемая литература 169

Глава 7. Беспроводные локальные сети 171
7.1. Инфракрасный диапазон против радиопередачи 173
7.2. Сети с инфраструктурой и эпизодические сети 174
7.3. Стандарт IEEE 802.11 177
7.3.1. Архитектура беспроводных сетей 177
7.3.2. Протокольная архитектура 178
7.3.3. Физический уровень 179
7.3.4. Уровень управления доступом к среде 182
7.3.5. Управление MAC 190
7.3.6. Дальнейшее развитие 195
7.4. HIPERLAN 196
7.4.1. Протокольная архитектура 197
7.4.2. Физический уровень 198
7.4.3. Подуровень управления доступом к каналу 200
7.4.4. Подуровень управления доступом к среде 204
7.4.5. Информационные базы и организация сетей 207
7.5. Bluetooth 209
7.5.1. Пользовательские сценарии 210
7.5.2. Физический уровень 211
7.5.3. Уровень MAC 212
7.5.4. Организация сетей 214
7.5.5. Безопасность 215
7.5.6. Управление каналом 216
7.6. Резюме 216
7.7. Упражнения 217
7.8. Рекомендуемая литература 218

Глава 8. Беспроводная ATM 219
8.1. Мотивы разработки WATM 220
8.2. Рабочая группа беспроводной ATM 220
8.3. Службы WATM 221
8.4. Эталонная модель 223
8.4.1. Примеры конфигураций 223
8.4.2. Общая эталонная модель 224
8.5. Функции 226
8.5.1. Функции беспроводного мобильного терминала 226
8.5.2. Функции сети 228
8.6. Уровень радиодоступа 229
8.6.1. Требования 229
8.6.2. BRAN 230
8.7. Переключение 232
8.7.1. Эталонная модель переключения 233
8.7.2. Требования к переключению 233
8.7.3. Типы переключений 234
8.7.4. Сценарии переключения 235
8.7.5. Обратное переключение 236
8.7.6. Прямое переключение 241
8.8. Управление местоположением 242
8.8.1. Требования к управлению местоположением 243
8.8.2. Процедуры и объекты 244
8.9. Адресация 247
8.10. Мобильное качество обслуживания 247
8.11. Протокол управления точками доступа 248
8.12. Резюме 250
8.13. Упражнения 252
8.14. Рекомендуемая литература 254

Глава 9. Мобильный сетевой уровень 255
9.1. Mobile IP 255
9.1.1. Цели, предположения и требования 256
9.1.2. Объекты и терминология 258
9.1.3. Доставка пакетов IP 260
9.1.4. Объявление и открытие агента 261
9.1.5. Регистрация 262
9.1.6. Туннелирование и инкапсуляция 264
9.1.7. Оптимизация 268
9.1.8. Обратное туннелирование 270
9.1.9. IPv6 271
9.2. Протокол динамической конфигурации хоста 271
9.3. Эпизодические сети 273
9.3.1. Маршрутизация 274
9.3.2. Дистанционно-векторная маршрутизация с нумерацией и оценкой пути к пункту назначения 277
9.3.3. Динамическая маршрутизация от источника 278
9.3.4. Иерархические алгоритмы 280
9.3.5. Альтернативные метрики 281
9.4. Резюме 283
9.5. Упражнения 283

9.6. Рекомендуемая литература 285

Глава 10. Мобильный транспортный уровень 287

10.1. Традиционный TCP 287
10.1.1. Управление перегрузкой 287
10.1.2. Плавный старт 288
10.1.3. Быстрая повторная передача/быстрое восстановление 289
10.1.4. Влияние мобильности 289
10.2. Косвенный TCP 290
10.3. Подслушивающий TCP 293
10.4. Мобильный TCP 294
10.5. Быстрая повторная передача/быстрое восстановление 296
10.6. Замораживание передачи/времени ожидания 297
10.7. Выборочная повторная передача 297
10.8. TCP, ориентированный на транзакции 298
10.9. Резюме 298
10.10. Упражнения 300
10.11. Рекомендуемая литература 300

Глава 11. Поддержка мобильности 303

11.1. Файловые системы 304
11.1.1. Согласованность 304
11.1.2. Примеры 305
11.2. World Wide Web 308
11.2.1. Протокол передачи гипертекстовых файлов 308
11.2.2. Язык гипертекстовой разметки 311
11.2.3. Некоторые подходы, которые могут способствовать беспроводному доступу 312
11.2.4. Системные архитектуры 314
11.3. Протокол беспроводных приложений 317
11.3.1. Архитектура 318
11.3.2. Протокол дейтаграмм беспроводных сетей 320
11.3.3. Безопасность транспортного уровня в беспроводных сетях 322
11.3.4. Протокол транзакций в беспроводных сетях 324
11.3.5. Протокол сеансов связи для беспроводной сети 328
11.3.6. Среда беспроводных приложений 336
11.3.7. Язык разметки для беспроводных сетей 337
11.3.8. WMLScript 339
11.3.9. Приложения беспроводной телефонии 341
11.3.10. Примеры стеков WAP 348
11.4. Резюме 349
11.5. Упражнения 349
11.6. Рекомендуемая литература 350

Приложение А. Сокращения 355

Предметный указатель 368

Об авторе

Йоган Шиллер (Jochen Shiller) получил ученую степень по вычислительной технике в Германии в университете Карлсруэ. Во время своей рабочей поездки в Швецию в университет Упсалы участвовал в европейском проекте HIPPARCH, занимаясь, в частности, обеспечением качества обслуживания (QoS) крупных WWW-серверов.

В настоящее время Йоган Шиллер работает в Институте Телематики (университет Карлсруэ), где проводит исследования телекоммуникационных информационных потоков и механизмов обратной связи для структур QoS и мобильных коммуникаций. Кроме того, является доцентом университета Карлсруэ, одного из ведущих технических европейских вузов, где читает лекции по компьютерным сетям и мобильным коммуникациям. Также д-р Шиллер преподает в университете Упсалы (Швеция), является членом IEEE (Institute of Electrical and Electronics Engineers) и консультирует несколько компаний по вопросам связи и сетей.

Предисловие

В настоящее время наблюдается бурное развитие систем мобильной связи, сопровождаемое все возрастающим интересом общества к мобильной связи и отменой государственного регулирования в прошлом сильно монополизированных рынков сбыта. В то время, как обычная связь имеет дело со стационарными сетями, мобильность порождает ряд новых вопросов, технических приемов и решений. Во многих странах мобильные коммуникации являются вообще единственным приемлемым решением из-за отсутствия инфраструктуры, необходимой для обеспечения стационарной связи. Подобная ситуация создает все возрастающую потребность в квалифицированных специалистах, понимающих принципы функционирования и существующие возможности новейших разработок в области мобильной связи. То, что мы видим сегодня — это только начало. В исследовательских лабораториях сейчас разрабатывается много новых захватывающих вещей. В будущем будет появляться все больше и больше мобильных устройств, произойдет слияние речевой связи и технологий передачи данных, использование Internet распространится на мобильные и беспроводные устройства. Новые приложения и новые мобильные сети выведут на массовый рынок мультимедийную обработку данных; радиоприемники, электронные записные книжки, переносные компьютеры и мобильные телефоны будут объединены в единое устройство, способное выполнять множество различных функций.

Данная книга является введением в область мобильных коммуникаций и сосредоточена на цифровой передаче данных. Она предназначена для студентов факультетов электроники и компьютерных наук, изучающих курс компьютерных сетей и коммуникаций; для инженеров, работающих со стационарными сетями и интересующихся будущим сетевых технологий; а также для тех менеджеров, которые нуждаются в хорошем и достаточно полном обзоре существующих средств мобильной связи. Предполагается, что читатель понимает основы построения и функционирования коммуникационных линий и немного знаком с общими принципами работы компьютерных сетей и Internet. Поскольку существует достаточно много книг по конкретным технологиям, в этой книге предпринимается попытка охватить различные аспекты мобильной связи, причем с точки зрения компьютерных наук. Более того, в книге рассматриваются общие свойства различных технических решений, показана интеграция служб и приложений для стационарных сетей в сети, поддерживающие беспроводный доступ и мобильность конечных систем. Для читателей, заинтересованных в более детальной информации по какой-либо теме, приведены ссылки на соответствующие публикации или Web-узлы.

Эта книга будет также полезна преподавателям курса, предназначенного для изложения основ передачи данных или принципов функционирования компьютерных сетей. Ее можно использовать как часть более общего курса при использовании других книг по стационарным или скоростным сетям. Вместе с данной книгой для курса можно использовать также материал, доступный по следующему Web-адресу:

`http://telematik.informatik.uni-karlsruhe.de/~schiller/`

Этот материал включает все иллюстрации, около 370 слайдов в формате PDF и PowerPoint™, список всех акронимов и множество ссылок на соответствующие Web-узлы. Кроме того, вопросы, приведенные в книге, студенты могут использовать для самоконтроля.

Книга предназначена людям, которые хотят знать, как работают системы мобильной телефонной связи, интересуются будущим спутниковых коммуникаций и влиянием мобильности на приложения, безопасность и компьютерные сети. Специалисты, работающие со стационарными сетями, найдут здесь описание путей перехода к смешанным стационарно-мобильным сетям.

В книге освещен широкий круг вопросов мобильной связи — от сигналов и протоколов доступа до требований к приложениям и безопасности. При этом не делается акцента на отдельных темах в ущерб всем остальным. Поскольку автор видит будущее за такими цифровыми системами, как CDMA, GSM, DECT, DAB, основное внимание сосредоточено на цифровых мобильных коммуникационных системах. Также рассмотрены новые и важные вопросы высших уровней связи, в частности, WAP — протокол, предназначенный для распространения информационных материалов по Internet.

В главе 1 рассмотрены основные понятия мобильной и беспроводной связи, представлена краткая история предмета и указаны вопросы, требующие дополнительных исследований. В главу также включен обзор рынка мобильных технологий, который иллюстрирует их потенциал. Глава 2 посвящена классическим аспектам коммуникационных систем и изложению основ беспроводной технологии с точки зрения компьютерных наук. Тема этой главы — распространение, уплотнение и модуляция сигналов. Хотя глубокие электротехнические знания здесь не потребуются, необходимо все же понимать основные принципы беспроводной передачи данных. В главе 3 представлено несколько схем сетевого доступа. Тут же объясняется, почему стандартные схемы стационарных сетей оказываются непригодными в беспроводном случае.

Главы 4–8 посвящены различным существующим беспроводным коммуникационным системам и их можно читать в любом порядке. Общей чертой упомянутых систем является возможность беспроводного доступа к сети и осуществление передачи произвольных данных. В главе 4 рассматриваются глобальная система мобильной связи (global system for mobile communication — GSM) как наиболее успешная на сегодняшний день система мобильной телефонной связи, радиомагистрали, беспроводные телефонные технологии и будущие разработки универсальной мобильной телекоммуникационной системы (universal mobile telecommunications system — UMTS). Спутниковые системы рассмотрены в главе 5, в главе 6 обсуждаются такие цифровые широковещательные системы, как DAB. Они могут быть частью большей системы связи, предоставляющей конечным пользователям массовые данные. Глава 7 посвящена беспроводным локальным сетям, позволяющим обойтись без укладки кабеля в зданиях. Примерами здесь являются технологии IEEE 802.11, HIPERLAN и Bluetooth. В главе 10 рассмотрена технология беспроводного асинхронного режима передачи (ATM). Особенностью беспроводной технологии ATM является обеспечение качества обслуживания (QoS) за счет контроля пропускной способности и частоты появления ошибок.

Большая часть главы 9 посвящена Mobile IP — расширению протокола Internet (IP) для мобильной связи. Также обсуждаются соответствующие эпизодические сети с их требованиями к специфическим протоколам маршрутизации. Следующий уровень, транспортный, рассмотрен в главе 10. Здесь обсуждаются несколько попыток приспособления известного протокола TCP к специфическим требованиям систем мобильной связи. И, наконец, в главе 11 представлен WAP — стандарт, позволяющий мобильным и беспроводным устройствам использовать ресурсы Internet.

Основой книги является курс, прочитанный в университете Карлсруэ, состоящий из 14 полуторачасовых лекций. Материал, изложенный в книге, практически используют уже больше 150 студентов. Еще раз напомним, что весь этот материал есть и в Internet.

Любые замечания и пожелания относительно книги или материала курса присылайте по адресу `schiller@computer.org`. Я надеюсь, что книга читателям понравится и они простят мне некоторые упрощения. Дальнейшие исследования в области мобильных коммуникаций могут изменить многие упомянутые в книге детали и стандарты.

Благодарности

В первую очередь я хотел бы поблагодарить тех многочисленных студентов прочитанного мной курса мобильных коммуникаций, которые подтолкнули меня к написанию этой книги, требуя больше информации по беспроводной и мобильной связи. Студенческие вопросы во многом определили содержание данной книги.

Также я признателен многим моим коллегам по университету Карлсруэ за полезные замечания и предварительное обсуждение глав книги. Особенно я благодарен Марку Бехлеру (Marc Bechler), Стефану Дреслеру (Stefan Dresler), Йогану Сайтцу (Jochen Seitz) и Гюнтеру Шаферу (Gunter Schafer). Я в долгу перед Хартмутом Риттером (Hartmut Ritter), который взял на себя часть моей повседневной работы. На ранних этапах подготовки данного курса мне помогали Верена Роуз (Verena Rose) и Эльмар Дорнер (Elmar Dorner).

Я хочу поблагодарить Пэра Гуннингберга (Per Gunningberg) из университета Упсалы за многочисленные проницательные замечания. Я глубоко признателен Анжелике Райдер (Angelika Rieder) и Керстину Риссе (Kerstin Risse) за помощь с переводом книги на английский язык.

Я благодарен сотрудникам отделения издательства Addison-Wesley в Великобритании, и особенно Пенелопе Олпорт (Penelope Allport), Элисон Бертвел (Alison Birtwell) и Стиву Темблетту (Steve Temblett) за поддержку, оказанную мне при написании книги.

И, наконец, я хочу выразить благодарность директору Института Телематики проф. Герхарду Крюгеру (Gerhard Kruger) за предоставленную мне свободу и возможность подготовки курса в наиболее благоприятных условиях.

Введение 1

Какими будут **компьютеры** через десять лет, через двадцать лет? Вряд ли можно точно ответить на этот вопрос, но определенно можно сказать, что большинство компьютеров будут **портативными**. Как пользователи через компьютеры или другие средства связи будут получать **доступ** к сетям? Количество пользователей, использующих **беспроводный** доступ, будет постоянно увеличиваться. Как **люди** будут проводить время на работе и в отпуске? Многие люди будут **мобильными** — это уже является одной из ключевых характеристик современного общества. Возьмем, к примеру, самолет на 800 мест. В современных лайнерах сейчас обеспечивается ограниченный доступ к сети, а в самолетах следующего поколения пассажиры будут иметь свободный доступ к Internet. В этом случае единственным средством передачи данных будет мобильная сеть, движущаяся над землей с огромной скоростью. Далее, представим себе автомобили с доступом к Internet и миллиарды встроенных процессоров, связанных с фотокамерами, мобильными телефонами, проигрывателями компакт-дисков, головными стереодинамиками, клавиатурами, дорожными знаками и различными датчиками.

Перед рассмотрением других применений необходимо определить понятия "мобильный" и "беспроводный", которые мы будем постоянно использовать в этой книге. Существует два различных типа мобильности: мобильность пользователя и портативность устройства. **Мобильность пользователя** подразумевает, что он может пользоваться одинаковыми или похожими телекоммуникационными службами из разных мест, т.е. при перемещении пользователя службы как бы "следуют" за ним. Примерами таких механизмов являются телефон или рабочий стол компьютера, поддерживающие роуминг (т.е. рабочий стол выглядит одинаково независимо от того, с какого компьютера пользователь зашел в сеть).

Когда говорят о **портативности устройства**[1], подразумевается, что средство связи может перемещаться (вместе с пользователем или без него). В этом случае в сети и внутри устройства необходимы механизмы, которые во время перемещения постоянно будут подтверждать, что связь все еще возможна. Типичным примером системы, поддерживающей портативность устройств, является система мобильной телефонной связи, которая может переключить устройство с одного радиопередатчика (также называемого базовой станцией) на другой, когда сигнал становится слишком слабым. Большинство систем, описанных в этой книге, характеризуются и мобильностью пользователей, и портативностью устройств одновременно.

---

[1] Помимо термина "портативность", в отношении устройств могут использоваться и несколько других терминов (например, "мобильный" в отношении "мобильных телефонов"). В этой книге обычно различаются беспроводный доступ к сетям и мобильность пользователя устройства, как его ключевая характеристика.

Понятие **беспроводное** используется по отношению к устройствам. Оно обозначает лишь возможность доступа к сети или к другим партнерам по связи без кабеля. Кабель заменяет передача электромагнитных волн "по воздуху" (хотя в сущности беспроводная передача не требует вообще никакой среды).

Итак, устройство связи может обладать следующими характеристиками.

- **Стационарное с кабелем**. Эта конфигурация описывает обычный компьютер в офисе.
- **Мобильное с кабелем**. В эту категорию попадают, например, переносные компьютеры — пользователь может переносить их из одной гостиницы в другую, подключаясь к сети своей компании через модем и телефонную сеть.
- **Стационарное беспроводное**. Этот режим используется для установки сетей в зданиях, имеющих историческую ценность, чтобы избежать их повреждений при укладке кабеля.
- **Мобильное беспроводное**. Это самый интересный случай. Не ограниченный никаким кабелем пользователь может переходить из одной беспроводной сети в другую. Большинство технологий, описанных в данной книге, работают с устройствами именно этого типа, а также с поддерживающими их сетями.

В следующем разделе приведен обзор типичных устройств мобильной и беспроводной связи и приведены их некоторые возможные применения. Однако, следует иметь в виду, что и устройства, и их применения в будущем могут измениться, поскольку рынок мобильных и беспроводных технологий развивается очень быстро. Рассмотрим кратко историю развития беспроводной связи на протяжении последних 200 лет. В разделе 1.3 обсуждаются перспективы рынка беспроводных и мобильных коммуникаций. Несмотря на то, что уже миллионы людей используют беспроводные устройства, потенциал этого рынка по-прежнему огромен.

Раздел 1.4 посвящен нескольким открытым проблемам, обусловленным фундаментальными различиями между проводной и беспроводной связью. Эталонная модель коммуникационных систем, которая используется в данной книге, рассмотрена в разделе 1.5, а завершает главу беглый обзор содержания всей книги.

1.1. Основные применения мобильной связи

Несмотря на то, что беспроводным сетям и мобильным коммуникациям может быть найдено множество применений, часть из них можно назвать сразу же. Некоторые из этих применений приведем в следующих разделах, — а расширить этот список читатель может самостоятельно.

1.1.1. Транспортные средства

Будущие автомобили будут объединять множество беспроводных коммуникационных систем и мобильных приложений. Музыка, новости, состояние дороги, прогноз погоды и другая широковещательная информация передаются через цифровую радиовещательную систему (DAB) со скоростью 1,5 Мбит/с. Возможность разговаривать и передавать данные со скоростью 384 Кбит/с обеспечивается глобальной системой мобильной связи (GSM). Для удаленного доступа можно использовать спутниковую связь, а с помощью глобальной системы навигации и определения положения (GPS) можно определять точное местоположение автомобиля. Для быстрого обмена информацией в аварийных ситуациях близко находящиеся автомобили объединяются в эпизодичную локальную сеть. Это помогает водителям соблюдать безопасное расстояние, а в случае аварии приводится в действие воздушная подушка и автоматически вызываются полиция и скорая помощь. Такие автомобили существуют уже сейчас. Будущие автомобили будут также оповещать другие автомобили об аварии, чтобы те сумели вовремя остановиться. Автобусы, грузовики и поезда уже сейчас передают текущую информацию о сво-

ем состоянии и перемещении на собственные базы, что помогает координировать организацию перевозок, а значит повысить их эффективность и обеспечить экономию времени и денег.

На рис. 1.1 приведена типичная схема мобильных коммуникаций со многими беспроводными устройствами. Сети с фиксированной инфраструктурой (например, сотовые телефоны) связаны с магистральными радиосистемами (TETRA) и беспроводными ЛВС (WLAN). Кроме того, могут использоваться спутниковые каналы связи. Сети внутри и между автомобилями, скорее всего, будут эпизодическими. Беспроводная сеть внутри машины может объединять электронные записные книжки, переносные компьютеры и мобильные телефоны, например, с помощью технологии Bluetooth.

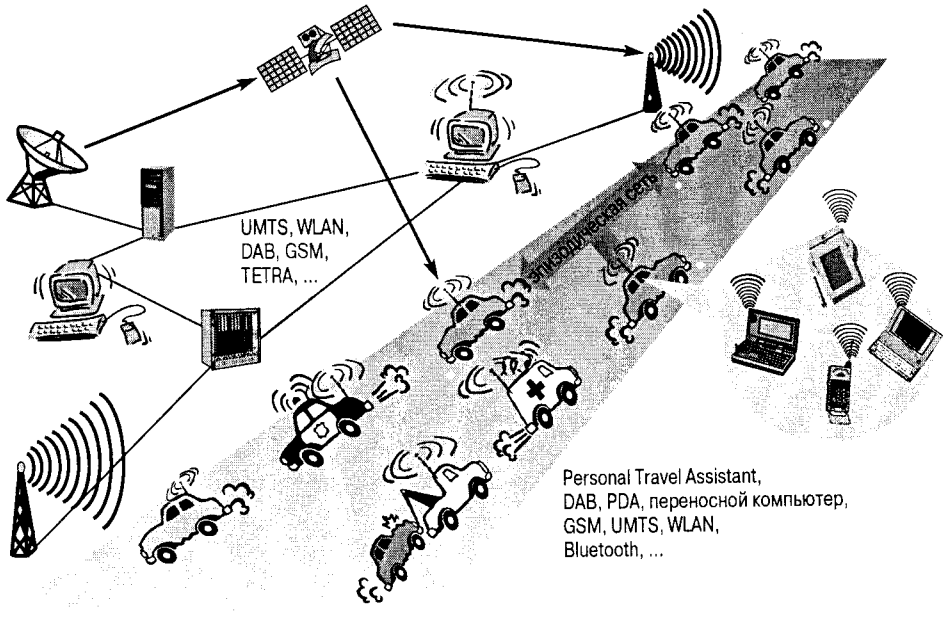

Рис. 1.1. Типичная схема применения мобильных коммуникаций в дорожном движении

Кроме графического содержания схема на рис 1.1 включает много **аббревиатур** — эта черта вообще характерна для всего, что связано с мобильными коммуникациями. В этой книге аббревиатуры будут встречаться весьма часто. Для удобства читателя их полный список вынесен в приложение А.

Похожие схемы можно придумать и для железнодорожного или воздушного движения. При этом, правда, мы столкнемся с различными дополнительными трудностями из-за высокой скорости перемещения — особенно в последнем случае. Скорость современных поездов может достигать 350 км/ч, а авиалайнеров — 900 км/ч и выше. В то же время многие технологии не работают, если относительная скорость мобильных устройств превышает, например, 250 км/ч (для GSM) или 100 км/ч (для AMPS). Лишь некоторые технологии (например, DAB) могут работать при скорости перемещения, достигающей 900 км/ч.

1.1.2. Чрезвычайные ситуации

Представим себе возможности высококачественного беспроводного соединения машины скорой помощи с больницей. При несчастном случае жизненно важную информацию относительно пострадавших можно немедленно передать в больницу. Там сразу же будет произведена соответствующая подготовка, а при необходимости можно будет заранее привлечь специалистов для постановки предварительного диагноза. Более того, беспроводные сети являются фак-

тически единственными средствами сообщения в случае таких стихийных бедствий, как ураганы или землетрясения. В самых худших случаях смогут "выжить" лишь специальные децентрализованные беспроводные сети. Нарушение всех кабельных соединений означает крах не только обычной телефонной связи, но и всех мобильных телефонных систем, для которых необходимы базовые станции!

1.1.3. Бизнес

Современному коммивояжеру нужен постоянный доступ к базе данных компании для того, чтобы убедиться, что файлы на его переносном компьютере отражают действительное состояние дел, чтобы обеспечить компании возможность контролировать деятельность своих служащих и т. д. При беспроводном доступе переносной компьютер превращается в настоящий мобильный офис.

1.1.4. Замена кабельных соединений

В некоторых случаях беспроводные сети заменяют кабельные соединения, например, в исторических зданиях, на торговых выставках или при дистанционном сборе данных. По экономическим причинам часто невозможно установить кабельное соединение с удаленными датчиками, необходимыми, например, для составления прогнозов погоды, выдачи предупреждений о возможном землетрясении или для сбора данных о состоянии окружающей среды. Выход из этой ситуации предоставляют беспроводные соединения через спутник. Для торговых выставок необходима очень динамичная инфраструктура, а укладка кабеля занимает много времени, и это часто оказывается неудобным. Поэтому на многих компьютерных выставках вместо кабеля используются беспроводные ЛВС (WLAN). Еще одной возможностью применения беспроводных сетей являются компьютеры, датчики и информационные дисплеи в зданиях, имеющих историческую ценность, где очень важно не повредить стены и пол, что неизбежно при укладке кабеля.

1.1.5. Информационная поддержка

Возможна ли Internet повсюду? С беспроводными сетями — да! Представьте себе путеводитель по городу. Статическую информацию можно получить с CD-ROM, DVD или даже дома через Internet. Но беспроводные сети смогут обеспечить нас нужными данными в любом месте (рис. 1.2).

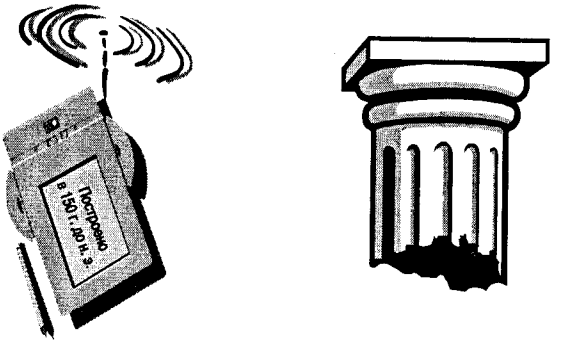

Рис. 1.2. Учитывающий местонахождение путеводитель с беспроводным доступом к сети

Экскурсовод сможет рассказать вам что-нибудь об истории здания (зная через систему GPS, где именно вы находитесь), пока ваш компьютер через локальную беспроводную сеть будет загружать информацию о сегодняшнем концерте в этом же здании. Когда информация будет загружена, вы сможете выбрать себе место в зале, оплатить билет с помощью электронной кредитной карточки, а за-

тем сбросить эту информацию поставщику сетевых услуг. Еще одной быстро растущей областью применения беспроводных технологий являются развлечения и игры, т.е. создание специальных игровых сетей, объединяющих людей, изъявивших желание принять участие в общей игре.

1.1.6. Услуги, зависящие от местонахождения

Много усилий разработчиков мобильных и беспроводных сетей направлено на то, чтобы скрыть изменение способа доступа к сети (скажем, с мобильного телефона на беспроводную ЛВС или между различными точками доступа к сети), а также то, что беспроводная связь более подвержена ошибкам, чем обычная кабельная. В разных главах этой книги можно найти многочисленные соответствующие примеры — например, протокол Mobile IP пытается скрыть факт изменения точек доступа к сети, осуществляя перенаправление пакетов и удерживая за счет этого тот же самый IP-адрес пользователя (раздел 9.1). Для того, чтобы приложения, разработанные для стационарных сетей, могли успешно работать в беспроводных сетях, многие протоколы улучшают качество соединения за счет механизмов кодирования или повторной передачи.

Однако, в некоторых случаях приложению или пользователю для успешного проведения дальнейших действий все же необходима информация о его местонахождении. В частности, ниже перечислено несколько типов услуг, зависящих от текущего местонахождения клиента.

- **Сопровождающие услуги**. Функция перенаправления звонков на текущее местонахождение пользователя нам хорошо знакома по старой доброй телефонной системе. Где бы вы ни были, достаточно передать временный телефонный номер на ваш телефон, и он будет автоматически перенаправлять все входящие звонки[2]. При использовании мобильных компьютеров сопровождающие услуги обеспечивают, например, одинаковый вид рабочего стола окна операционной системы, в каком бы уголке планеты ни находился пользователь. Электронная почта будет автоматически пересылаться, все изменения на рабочем столе и в документах будут сохраняться и автоматически направляться в главный офис компании. Если кто-то попытается получить к вам доступ, пользуясь системой электронных конференций, поступивший вызов также будет переправлен в то место, где вы в данный момент находитесь.
- **Услуги, учитывающие местонахождение**. Предположим, что вы находитесь в вестибюле гостиницы и хотите распечатать документ с помощью своего переносного компьютера. Если просто перетащить документ на пиктограмму принтера, то где в этом случае, по вашему мнению, должен быть напечатан данный документ? Безусловно, не на принтере в центральном офисе! Однако, без дополнительной информации о текущем окружении только этого и можно ожидать. Следовательно, службам операционной системы необходимо предоставить дополнительные сведения о возможностях текущего окружения пользователя. Например, службы сети гостиницы должны известить вас о том, что в вестибюле есть общедоступный лазерный принтер. В свою очередь, ваш компьютер должен будет передать окружению сведения о вас и использованных вами услугах, чтобы плата за эти услуги была внесена в ваш счет.
- **Конфиденциальность**. Два рассмотренных типа услуг затрагивают и вопросы возможности достижения определенной конфиденциальности. К примеру, вы хотите, чтобы во время обеда электронная почта вам по-прежнему пересылалась, но не желаете отвечать на входящие вызовы видеоконференций. Иначе говоря, бывают случаи, когда некоторые услуги необходимо исключить. Более того, иногда бывает необходимо воспользоваться услугами, зависящими от местонахождения, но без предоставления окружению точной информации о себе.
- **Информационные службы**. Во время прогулки по городу беспроводный путеводитель можно использовать для получения самой разнообразной информации, например, для ответа на

[2] В действительности эта функция реализована в телефонной сети, а собственно телефон лишь обрабатывает поступающие сигналы.

вопрос "Где находится ближайший мексиканский ресторан?". В то же время служба сама может активно выводить информацию в ваш путеводитель, посылая сообщение о скидках в том мексиканском ресторане, который находится за углом.

- **Службы поддержки**. Наконец, можно интегрировать много дополнительных мелких механизмов поддержки мобильного устройства. Скажем, промежуточные результаты вычислений или содержимое кэша могут "следовать" за ним по стационарной сети. Вся эта информация сразу же окажется доступной при каждом новом подключении мобильного устройства, что позволит уменьшить задержку доступа и снизить загрузку стационарной сети. Альтернативой такому подходу является централизованное размещение информации пользователя и постоянный доступ к ней через сеть (возможно, очень большую и перегруженную) — что обычно и имеет место в настоящее время.

1.1.7. Мобильные и беспроводные устройства

Хотя уже есть много мобильных и беспроводных устройств, в будущем их будет намного больше. Пока нет строгой классификации таких устройств по форме, размерам, весу или вычислительной мощности. В данный момент вершиной мобильных технологий являются переносные компьютеры (лэптопы)[3]. В приведенном ниже списке рассмотрено несколько мобильных и беспроводных устройств в порядке возрастания их производительности (CPU, память, дисплей, устройства ввода и т. д.).

- **Датчик**. Это очень простое беспроводное устройство, представляющее собой некоторый датчик, передающий информацию о своем состоянии. В качестве примера можно привести датчик на двери офиса. Когда дверь заперта, датчик передает это состояние на мобильный телефон внутри офиса, и мобильный телефон не принимает входящие звонки. Таким образом сведения о том, что дверь заперта, без какого-либо вмешательства пользователя оказывают влияние на режим приема телефонных звонков.
- **Встроенные средства управления**. У многих приборов есть средства управления, иногда простые, иногда более сложные. Здесь примерами могут быть клавиатуры, мыши, наушники, стиральные машины, кофемолки, фены или телевизоры. Почему бы не превратить фен в простое мобильное беспроводное устройство (с точки зрения коммуникаций), которое может взаимодействовать с мобильным телефоном? Тогда телефон мог бы выключать фен, как только кто-то позвонит — по-моему, это прекрасное применение мобильной связи!
- **Пейджер**. Это очень простой приемник, который может показывать короткие текстовые сообщения на крохотном дисплее, но сам никаких сообщений отправлять не может. Иногда пейджеры встраиваются даже в часы.
- **Мобильные телефоны**. Обычный мобильный телефон может передавать речь или короткие сообщения и имеет простой черно-белый текстовый дисплей. Однако, современные мобильные телефоны все больше приближаются к электронным записным книжкам. Теперь они имеют цветной графический дисплей, сенсорный экран и встроенный Internet-броузер.
- **Электронная записная книжка (PDA)**. Эти устройства предлагают пользователю простые версии офисных программ (календарь, блокнот, почта, поисковые программы). Типичным устройством ввода является ручка, а встроенные средства распознавания превращают написанный от руки текст в последовательность символов. Для этих устройств также существуют специальные версии Web-броузеров и других популярных программных продуктов.

---

[3] Если мэйнфрейм установить на грузовик, то это вовсе не сделает его мобильным компьютером.

- **Карманный компьютер**. Следующий шаг — это карманные компьютеры, которые имеют цветной дисплей и небольшую клавиатуру. На них доступны упрощенные версии обычных программ для персональных компьютеров (обработка текста, процессор электронных таблиц и т.п.).
- **Ноутбук/переносной компьютер**. Наконец, переносные компьютеры обладают примерно такими же характеристиками и показателями производительности, как и обычные персональные компьютеры. На них используется такое же программное обеспечение, а технически они отличаются лишь размерами, весом и возможностью автономной работы на батареях.

В будущем мобильные и беспроводные устройства будут предоставлять пользователям и другие возможности, и при этом они будут легче и мощнее. Однако, здесь существует до сих пор не решенная проблема энергообеспечения. Чем больше возможностей у устройства, тем больше энергии ему необходимо. Чем выше его характеристики, тем быстрее "садятся" батарейки. Кроме того, много энергии потребляет собственно беспроводная передача данных.

Несмотря на то, что область мобильных коммуникаций очень быстро развивается, для типичных современных устройств, кроме проблемы энергообеспечения, характерны и другие существенные недостатки в сравнении с обычными персональными компьютерами. Для того, чтобы устройство было портативным, его интерфейс должен быть достаточно маленьким. Поэтому используются небольшие клавиатуры, на которых часто просто неудобно печатать из-за малых размеров клавиш. Более того, иногда маленькие графические дисплеи оказываются бесполезными, так как решающим фактором в этом случае является не графическое разрешение дисплея, а разрешающая способность глаз человека. Таким образом, необходимы новые механизмы взаимодействия устройства с пользователем, например, распознавание и синтез речи.

Объединение телекоммуникаций и компьютерных сетей сильно повлияло и на мобильную связь. Вряд ли можно точно описать, каким будет телефон будущего, но скорее всего, это будет компьютер. Даже сейчас обычные и мобильные телефоны далеки от тех простых "средств передачи голоса", какими они были в прошлом[4]. Такие разработки, как "речь через IP" и общая тенденция к переходу на пакетно-ориентированные сети, стимулируют изменения функциональных возможностей телефонов. Хотя нельзя точно предсказать будущее коммуникационных устройств, совершенно понятно, что все равно будет существовать много стационарных систем, дополненных мириадами беспроводных устройств, распределенных по всему миру.

1.2. Краткая история беспроводной связи

Чтобы лучше понять особенности современных беспроводных систем и этапы их развития, в этом разделе кратко рассмотрим общую историю беспроводной связи. Конечно, здесь не удастся охватить все изобретения и открытия, поэтому остановимся на тех из них, которые больше всего повлияли на современные системы.

Использование света для беспроводной связи известно с древнейших времен. В прошлом световые сигналы "модулировались" с помощью системы зеркал, которая позволяла включать и выключать световой луч ("амплитудная модуляция"). В других случаях для передачи закодированных в сигналах слов использовались флаги ("амплитудная и частотная модуляция", см. главу 2). Употребление для связи дымовых сигналов упоминается Полибиусом, Греция, еще в 150 г. до н.э. Световые сигналы и флаги использовались на флоте для беспроводной связи вплоть до изобретения радиосвязи. Однако и сейчас каждый матрос должен уметь передавать некоторые сигналы с помощью флагов — на случай неисправности других средств беспроводной связи.

---

[4] В главе 4 мы познакомимся с дополнительными возможностями современных мобильных телефонных систем, имеющих следствием все возрастающие требования к расширению полосы пропускания, вызванные стремлением использовать на мобильных "телефонах" типичные Internet-приложения.

В конце 18-го века **Клод Шапе** (Claude Shappe) изобрел оптический телеграф (**1794**), который положил начало беспроводной связи на далекие расстояния. Линии оптического телеграфа использовались до конца девятнадцатого века.

Первая коммерческая телеграфная линия была проложена между Вашингтоном и Балтимором в 1843 году, а в 1876 году **Александр Грэхэм Белл** (Alexander Graham Bell) изобрел и начал продавать телефоны. Это пробовали делать и раньше, но безуспешно (например, **Филип Райс** (Philip Reis, 1834–1874), который изобрел механизм телефонной связи в **1861** году). Телефонная служба появилась в Берлине в **1881** году. В 1936 году была осуществлена передача звука и изображения (мультимедиа!) между Берлином и Лейпцигом.

Общим недостатком систем оптической передачи была высокая частота несущего светового сигнала. Любая мелкая преграда отклоняла сигнал, а дождь или туман делали связь практически невозможной. Кроме того, в те времена было трудно сфокусировать свет в достаточной степени (сейчас это делается с помощью лазера). Поэтому очень важными для беспроводной связи оказались открытие электромагнитных волн и разработка оборудования для их модуляции. Все началось с того, что в 1831 году **Майкл Фарадей** (Michael Faraday) и примерно в то же время **Джозеф Генри** (Joseph Henry) открыли явление электромагнитной индукции. Теоретические основы электромагнетизма заложил **Джеймс К. Максвелл** (James C. Maxwell, 1831–1879), который в **1864** году записал свои знаменитые уравнения. Наконец, в **1886** году **Генрих Герц** (Heinrich Hertz, 1857–1894) экспериментально проверил справедливость уравнений Максвелла, доказав волновую природу электрической передачи в пространстве. Сегодня о его открытии нам напоминает название единицы измерения частоты (Гц). Вскоре **Николе Тесла** (Nikola Tesla, 1856–1943) удалось увеличить дальность электромагнитной передачи.

Успех беспроводных технологий тесно связан с именем **Гильермо Маркони** (Guglielmo Marconi, 1874–1937). В **1895** году он впервые продемонстрировал возможности беспроводного телеграфа, использовав длинноволновый передатчик с очень большой мощностью (больше 200 кВт). Первая трансатлантическая передача была осуществлена в 1901 году, а уже через шесть лет, в **1907** году, были созданы первые **коммерческие трансатлантические соединения**. Для этого потребовались огромные базовые станции с тридцатью стометровыми антеннами по обе стороны Атлантического океана. Примерно в это же время состоялась **Всемирная конференция по управлению радио** (WARC), целью которой была координация использования радиочастот. Первая **радиопередача**, в которой прозвучала рождественская музыка, была проведена в **1906** году **Реджинальдом Фессенденом** (Reginald A. Fessenden, 1866–1932). В 1915 году была впервые осуществлена беспроводная передача речи между Нью-Йорком и Сан-Франциско. Первая **коммерческая радиостанция** начала свою работу в **1920** году (KDKA, Питтсбург). При этом и передатчику, и приемнику были равно необходимы громадные антенны и высокая мощность передачи.

Все кардинально изменилось после открытия тем же Маркони **коротких волн** в **1920** году. Для беспроводной связи важно следующее достоинство коротких волн: они отражаются от ионосферы. Это позволяет посылать короткие радиоволны в любую точку земного шара, причем данная технология находит применение и сегодня. Изобретение **электронно-вакуумных приборов** в **1906** году, сделанное **Ли ДиФорестом** (Lee DeForest, 1873–1961) и **Робертом фон Либеном** (Robert von Lieben, 1878–1913), позволило уменьшить размеры передатчиков и приемников. Вакуумные приборы также используются до сих пор, например, для усиления выходного сигнала передатчика в современных радиостанциях. Один из первых **"мобильных" передатчиков** находился на борту дирижабля (**1911**). Наконец, в **1926** году был установлен первый **телефон в поезде**, следовавшем по маршруту Берлин–Гамбург, причем вместо антенны использовались линии электропередачи вдоль железной дороги. **Радиоприемники для автомобилей** поступили в продажу в **1927** году, хотя еще в 1922 году 18-летнему Джорджу Фросту (George Frost) из Чикаго удалось встроить радио в автомобиль марки "Форд Т".

1928 год ознаменовал начало эры **телевещания**. **Джон Берд** (John L. Baird, 1888–1946) изобрел цветное телевидение и осуществил передачу телесигнала через Атлантику. Станция WGY (Нью-Йорк) начала **регулярное телевещание** и трансляцию **теленовостей**. Первые **телекурсы** появились в **1932** году на станции **CBS**. До этого времени беспроводные коммуникации использовали исключительно амплитудную модуляцию (см. раздел 2.6). При этом наличие различных помех сильно ухудшало качество принимаемого сигнала. Большой шаг вперед в этом отношении был сделан после изобретения в **1933** году **Эдвином Армстронгом** (Edwin H. Armstrong, 1890–1954) **частотной модуляции**. Обе схемы модуляции до сих пор используются в радиовещании. В начале 30-х годов было уже много радиостанций, которые транслировали свои передачи по всему миру.

После второй мировой войны многие национальные и международные проекты в области беспроводной связи были прекращены. Первая немецкая сеть A-Netz (1958) была аналоговой и использовала несущую частоту 160 МГц. Установка соединения была возможна только с мобильной станции, переключение базовой станции было недоступно. В 1971 году эта система покрывала около 80 процентов территории страны и имела 11 000 пользователей. В 1972 году в Германии появилась сеть B-Netz, работающая на той же самой частоте в 160 МГц. Новой чертой этой сети была возможность установки соединения со станции в стационарной телефонной сети. Тем не менее, для этого необходимо было знать текущее местонахождение мобильного приемника. Передатчик и приемник были тяжелыми и обычно встраивались в автомобили. Эта система распространялась также на Австрию, Голландию и Люксембург и в 1979 году насчитывала 13 000 пользователей.

В то же время в северных европейских странах — Дании, Финляндии, Норвегии и Швеции (колыбели современной мобильной связи) была создана **Скандинавская мобильная телефонная система (NMT)**. Аналоговая сеть NMT работала на несущей частоте в 450 МГц и долгое время являлась единственным средством мобильной связи в некоторых отдаленных местах северной Европы (следующая система NMT, использующая частоту 900 МГц, появилась в 1986 году). Было разработано еще несколько национальных стандартов. В результате, в начале восьмидесятых годов в Европе действовало несколько различных и полностью несовместимых стандартов аналоговой мобильной телефонной связи. В **1982** году, в соответствии с общими принципами Европейского союза, было решено разработать общеевропейский мобильный телефонный стандарт. Новая система должна была работать на частоте 900 МГц и обеспечивать роуминг по всей Европе. Кроме того, было решено, что эта система должна быть цифровой, способной предоставлять услуги по передаче речи и данных. Для разработки этой системы была создана группа **"Groupe Spéciale Mobile" (GSM)**.

Примерно в то же время (**1983**) в США была создана система **AMPS** — аналоговая система мобильной телефонной связи, работающая на частоте 850 МГц (EIA, 1989). В **1984** году, после принятия стандарта **CT1** (его предшественником был стандарт **CT0** от **1980** года) впервые появились домашние беспроводные телефоны. Поскольку цифровых систем тогда еще не было, последовали и другие аналоговые стандарты, например, немецкая сеть C-Netz на частоте 450 МГц с аналоговой передачей речи. Здесь уже стало возможным переключение базовой станции; а сигнальная система была цифровой — в соответствии с общей тенденцией развития стационарных сетей (SS7). Также поддерживалось автоматическое определение местонахождения мобильного пользователя. Эта сеть используется до сих пор, главным образом из-за стопроцентного покрытия территории. Кроме передачи речи в сети предоставлялись следующие услуги: факс, передача данных через модем, поддержка протокола X.25 и электронная почта. Стандарт **CT2**, преемник CT1, был включен в список британских стандартов, опубликованный в **1987** году (DTI, 1987). Позже этот стандарт был утвержден ETSI как общеевропейский (ETS, 1994). Стандарт CT2 предусматривает работу на частоте 864 МГц, обеспечивая передачу данных со скоростью 32 Кбит/с.

В начале 90-х годов появились первые **полностью цифровые системы**. В **1991** году ETSI был принят **цифровой европейский беспроводный телефонный стандарт (DECT)**, регламентирующий организацию цифровой беспроводной телефонной связи (ETSI, 1998). Стандарт DECT использует частоту 1880–1900 МГц и интервал 100–500 м. Передача данных со скоростью до 1,2 Мбит/с происходит че-

рез сто двадцать дуплексных каналов. Новыми чертами этого стандарта являются кодирование речи и аутентификация пользователей. Данная система может поддерживать несколько десятков тысяч пользователей на квадратном километре и широко используется более чем в 40 странах мира. Недавно из маркетинговых соображений стандарт DECT был переименован и получил новую расшифровку названия — **цифровые расширенные беспроводные телекоммуникации**, что отражает возможность передачи мультимедийных данных. В **1991** году, после долгих лет обсуждений и испытаний, был наконец принят стандарт **GSM**. Его описание заняло больше 5 000 страниц. Эта первая версия GSM, которая сейчас называется **Глобальной системой мобильной связи**, работает на частоте 900 МГц и использует 124 полнодуплексных канала. Система GSM обеспечивает полный международный роуминг, автоматическое определение местонахождения, контроль доступа, кодирование в беспроводных каналах связи, эффективное взаимодействие с системами ISDN и относительно высокое качество звука. К тому же система GSM позволяет передавать данные со скоростью 9,6 Кбит/с и короткие сообщения (до 160 буквенно-цифровых символов). В зависимости от государства каналы могут использовать один или несколько операторов, могут применяться различные схемы оплаты и т. п. При этом все системы GSM остаются совместимыми. Стандарт GSM на частоте 900 МГц был принят операторами более чем в 130 странах мира.

Вскоре оказалось, что при высокой плотности пользователей в городах аналоговой сети AMPS (в США) и цифровой сети GSM (в Европе) недостаточно. В США для новой системы не нашлось новой частоты, а в Европе для этого была выбрана полоса в области 1800 МГц. Результат оказался следующим. В США несколькими компаниями были разработаны новые технологии работы в одной полосе частот в сети AMPS. Это привело к возникновению трех несовместимых систем: аналоговой узкополосной AMPS (IS-88, [TIA, 1993a]) и двух цифровых систем — **TDMA** (IS-136, [TIA, 1996]) и **CDMA** (IS-95, [TIA, 1993b]). В Европе же просто стали использовать систему GSM и на частоте 1800 МГц. Сейчас эту систему поддерживают больше десятка стран. Сети GSM-1800 (их также называют **DCS 1800** или цифровой сотовой системой) обеспечивают более высокое качество речи и работают с меньшими ячейками. Поэтому они лучше подходят для использования в больших городах. Сейчас систему GSM используют и в США (GSM-1900 или **PCS 1900**). Эта система, как и более поздние версии TDMA и CDMA, работает на частоте 1900 МГц.

В Европе верят в стандарты, в США — в рыночные механизмы. Стандарт GSM — это один из немногих примеров, когда более эффективным оказался подход стандартизации. Сейчас, когда Европа имеет единый стандарт, а роуминг возможен даже в Австралии или Сингапуре, в США до сих пор борются с различными несовместимыми системами. Однако, если рассмотреть системы, больше ориентированные на передачу данных (например, ЛВС), картина выглядит несколько иначе. Когда в **1996** году комитетом ETSI была стандартизована сеть HIPERLAN (выпущен сборник стандартов и рекомендаций), в мире уже существовало множество беспроводных ЛВС. Стандарт HIPERLAN 1-го типа использовал частоту 5,2 ГГц и обеспечивал передачу данных со скоростью до 23,5 Мбит/с. Для стандарта HIPERLAN 4-го типа, работающего на частоте 17 ГГц, скорость была повышена до 155 Мбит/с. Однако, самым лучшим для локальных сетей оказался стандарт IEEE **802.11**, появившийся позже (**1997**), чем HIPERLAN. Он использует свободную от лицензирования полосу частот для промышленного, научного и медицинского использования (ISM) 2,4 ГГц и инфракрасный диапазон, обеспечивая скорость передачи данных 2 Мбит/с (для специальных решений — до 10 Мбит/с). Несмотря на то, что стандарт HIPERLAN обладает более высокими рабочими характеристиками, большинство компаний работает с оборудованием, совместимым со стандартом IEEE 802.11.

1998 год ознаменовал начало эры спутниковых мобильных коммуникаций через систему **Iridium** (Iridium, 1998). До этого времени спутники использовались в качестве средств широковещания, а для работы с ними требовалось громоздкое оборудование. Мобильные спутниковые телефоны системы Iridium компактны и по-настоящему портативны. Система Iridium — это 66 спутников, расположенных на низких орбитах, а для мобильной связи используется частота 1,6 ГГц. Позднее, также в 1998

году, в Европе была разработана **Универсальная мобильная телекоммуникационная система** (UMTS). Она была предложена Международному телекоммуникационному союзу (ITU) в качестве основы для создания системы **IMT-2000 (международных мобильных телекоммуникаций)**. Система UMTS объединила сети стандарта GSM с наиболее эффективными решениями системы CDMA.

В рекомендациях IMT-2000 определяется единая всемирная структура мобильных коммуникаций будущего (ITU, 1998). Этот стандарт использует частоту 2 ГГц и охватывает базовые службы и архитектуру сетей, включая спутниковые коммуникации. Кроме того, он содержит рекомендации для развивающихся стран, требования к радиоинтерфейсам и соглашения по использованию частот. В нем описана структура системы безопасности и управления, а также предложены различные технологии передачи.

Сводный обзор нескольких упомянутых выше сетей приведен на рис. 1.3.

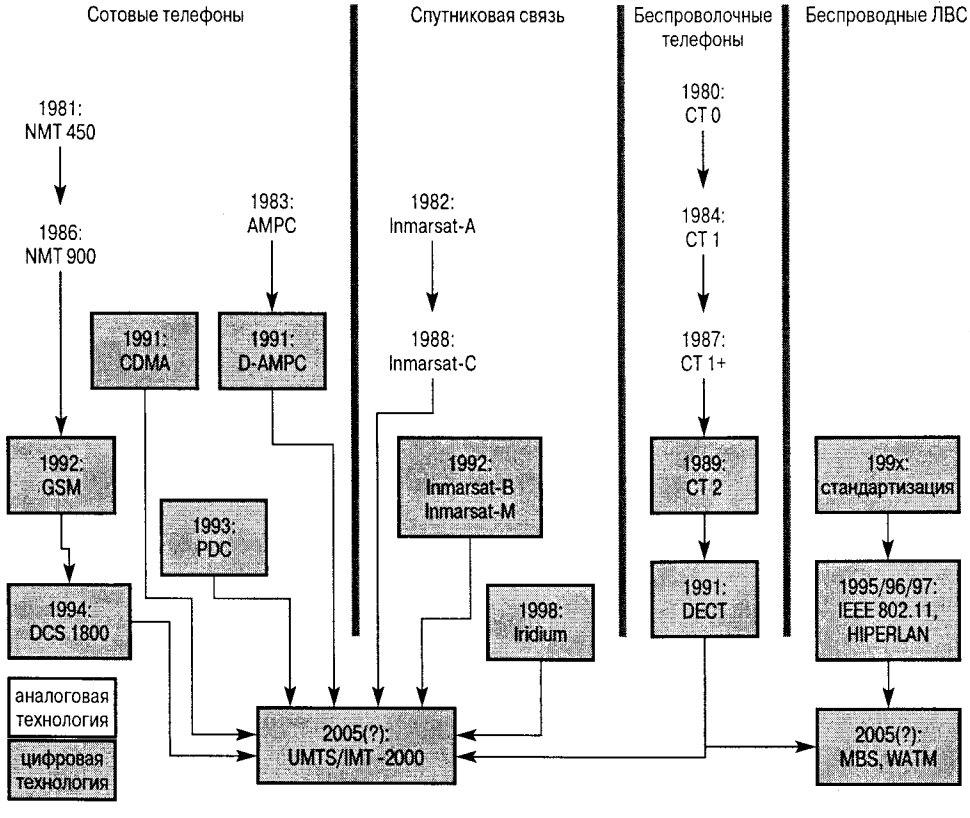

Рис. 1.3. Обзор некоторых беспроводных систем коммуникации

Эта диаграмма иллюстрирует развитие европейских сотовых и беспроводных телефонных систем, а также спутниковых коммуникаций и ЛВС. В настоящее время наблюдается тенденция к объединению, по крайней мере, некоторых технологий в единую **Мобильную широкополосную систему (MBS)** локальных беспроводных коммуникаций, регламентируемую стандартом IMT-2000. Система MBS работает в более широкой полосе частот и обеспечивает более высокое качество обслуживания. Даты на диаграмме обозначают начало работы различных систем (имеется в виду, что разработка и испытания проводились раньше). Некоторые системы будут описаны в следующих главах (сотовые и беспроводные телефоны — в главе 4, спутниковые коммуникации — в главе 5, беспроводные ЛВС — в главах 7 и 8).

1.3. Рынок мобильных коммуникаций

Сегодняшние темпы развития беспроводных коммуникаций подчеркивают огромный потенциал рынка этих технологий. Все больше и больше людей пользуется мобильными телефонами, приборы беспроводных технологий встраиваются во многие автомобили, во множестве мест используются беспроводные локальные сети.

На рис. 1.4 представлена диаграмма роста количества абонентов мобильной телефонной связи (GSM Data, 1998). Приведенные темпы роста просто невероятны, особенно в странах, находящихся вне трех больших регионов — Америки, Европы и Японии. Этот эффект в основном обусловлен возникновением в Азии большого количества служб мобильной телефонной связи.

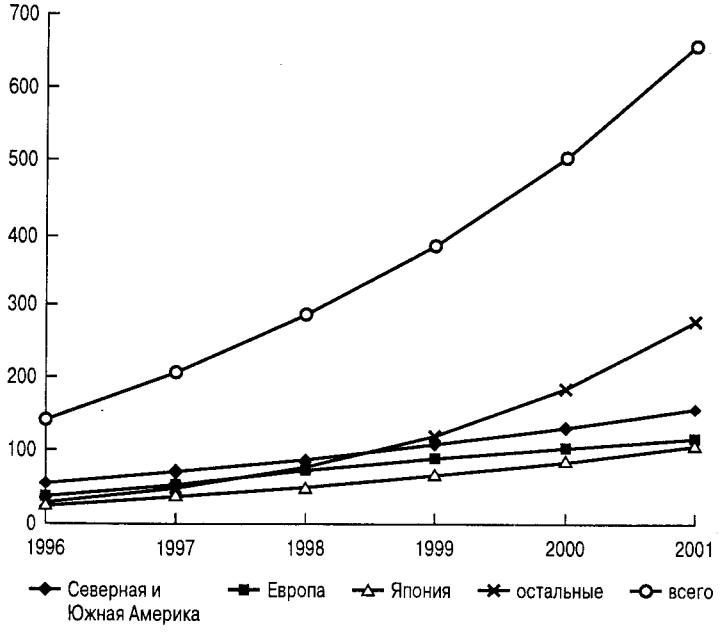

Рис. 1.4. Диаграмма роста количества абонентов мобильной телефонной связи (в миллионах)

На рис. 1.5 представлены данные о проникновении мобильных телефонов на рынок в 1998 году. После анализа этой диаграммы потенциал увеличения продаж во многих странах и общая тенденция к увеличению количества мобильных телефонов во всем мире становятся совершенно очевидными.

1.4. Некоторые еще нерешенные задачи

Несмотря на то, что в этой книге рассмотрено много мобильных систем и уже существующих решений организации беспроводного доступа, в данной области все еще остается много несделанного. Мы живем в самом начале эры мобильных и беспроводных сетей. Различия между стационарными и беспроводными сетями поднимают следующие вопросы.

- **Помехи**. Передачу с использованием радиоволн невозможно защитить от помех с помощью экранирования, как это делается в случае коаксиального кабеля или витой пары. Поэтому работа электромоторов или вспышки молнии способны вызвать сильные помехи и, в конечном счете, привести к ошибкам или потере информации.

- **Использование частот**. К сожалению, в нашем распоряжении имеется очень ограниченное количество частот. По этой причине их использование необходимо как-то координировать. Нужно также научиться использовать частоты более эффективно, например, с помощью новых схем модуляции (см. главу 2) или уплотнения по запросу (см. главу 3).

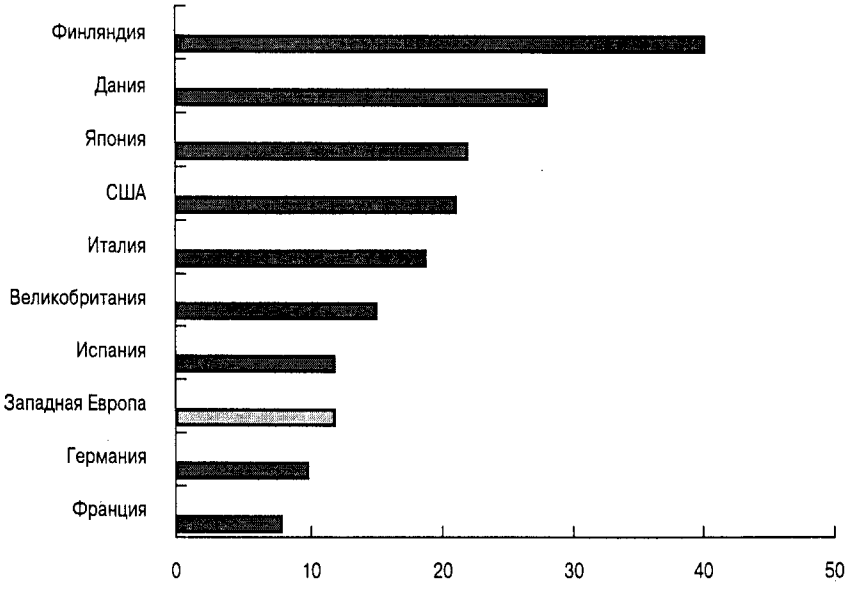

Рис. 1.5. Количество мобильных телефонов на 100 человек (1998 г.)

- **Низкая пропускная способность**. Хотя скорость передачи у беспроводных устройств постоянно растет, она все равно очень низкая в сравнении со стационарными системами. У локальных беспроводных систем она достигает нескольких Мбит/с, а глобальные системы обеспечивают всего лишь несколько десятков Кбит/с. Возникает необходимость приспособить существующие приложения для высокоскоростных соединений к такому новому для них окружению. Это нужно для того, чтобы пользователь мог работать с одним и тем же приложением как внутри, так и вне офиса.
- **Большие задержки и значительные колебания их длительности**. Серьезной проблемой для современных Internet-протоколов (TCP/IP) является большой разброс характеристик беспроводных соединений. В беспроводных системах задержки могут составлять до нескольких секунд, а соединение может быть очень асимметричным (т.е. качество связи существенно отличается в зависимости от направления передачи).
- **Низкая безопасность**. Очевидно, что портативные устройства украсть легче. Однако, более важно, что любой радиоинтерфейс подвержен опасности нежелательного прослушивания. Поэтому при организации беспроводного доступа обязательно использование кодирования, аутентификации и применение других механизмов обеспечения безопасности.
- **Общая среда передачи**. Любой радиодоступ всегда осуществляется через общую среду передачи. Поскольку проложить кабель между каждым передатчиком и приемником просто невозможно, различным участникам приходится "бороться" за право использования эфира. Несмотря на то, что разработано много разных схем доступа к среде, здесь все еще остается несколько нерешенных проблем, например, эффективное обеспечение требуемого качества обслуживания.

1.5. Упрощенная эталонная модель

Для описания структуры систем связи в этой книге используется **основная эталонная модель** [Tanenbaum, 1996]. Читателю, незнакомому с основами сетевых коммуникаций, рекомендуется предварительно ознакомиться с соответствующими главами в существующей литературе (например, [Halsall, 1996] или [Keshav, 1997]). В качестве примера беспроводного портативного устройства на рис. 1.6 приведена электронная записная книжка (PDA). Она поддерживает связь с базовой станцией, помещенной посредине рисунка. Базовая станция состоит из радиотрансивера (передатчика и приемника) и межсетевого устройства, связывающего беспроводное и стационарное звенья сети. Наконец, справа на рисунке показан компьютер, являющийся партнером PDA по сеансу связи.

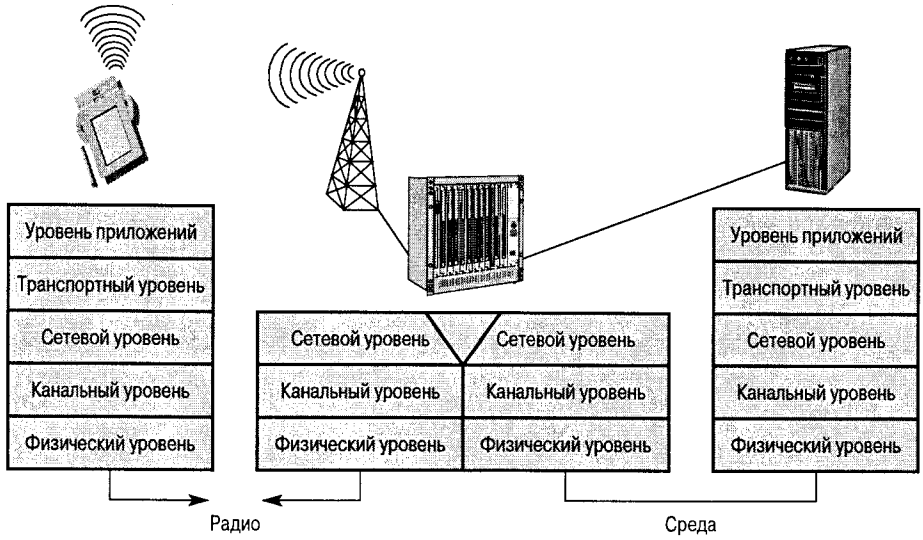

Рис. 1.6. Простая сеть, демонстрирующая схему работы используемой в книге эталонной модели

Под каждым элементом рассматриваемой сети (PDA, межсетевым устройством, компьютером) приведен используемый ими **набор протоколов**, реализованных в системе согласно требованиям эталонной модели. Для **конечных систем** (в нашем случае — это PDA и компьютер) необходим полный набор протоколов, включающий уровень приложений, а также транспортный, сетевой, канальный и физический уровни. Приложения, функционирующие на конечных системах, связаны между собой через службы нижних уровней. **Промежуточным системам** (межсетевое устройство) не обязательно иметь все уровни, — на рис. 1.6 приведены лишь сетевой, канальный и физический. В соответствии с основной эталонной моделью между собой взаимодействуют только одноранговые уровни набора протоколов (т.е. транспортный с транспортным, сетевой с сетевым и т.д.). Поэтому, согласно приведенной схеме, приложения конечных систем напрямую не взаимодействуют с какими-либо промежуточными системами. Ниже функции каждого уровня в беспроводной и мобильной среде будут объяснены более подробно.

- **Физический уровень.** Это самый нижний уровень любых систем коммуникации. Он отвечает за превращение потока битов в сигнал, который может быть послан передатчиком. На физическом уровне приемника сигнал вновь преобразуется в поток битов. В беспроводной связи на физическом уровне выполняется частотная селекция, генерация несущей частоты,

детектирование сигнала (даже искаженного сильными помехами), модуляция данных на несущей частоте и (в зависимости от схемы передачи) кодирование. Эти функции физического уровня подробно обсуждаются в главе 2, а позже они будут рассматриваться для каждой системы отдельно.

- **Канальный уровень**. Главными задачами этого уровня являются осуществление доступа к среде, уплотнение различных потоков данных, исправление ошибок передачи и синхронизация (т.е. обнаружение информационных кадров). Различные схемы доступа к среде обсуждаются в главе 3. Кроме того, в следующих главах есть разделы, посвященные особенностям канальных уровней систем отдельных типов. В целом можно сказать, что канальный уровень является ответственным за двухточечное соединение между двумя устройствами или соединение между одним передатчиком и несколькими приемниками.
- **Сетевой уровень**. Этот уровень отвечает за маршрутизацию пакетов и установку соединения через промежуточные системы. Здесь важнейшими понятиями являются адресация, маршрутизация, определение местонахождения устройства и переключение между различными сетями. В главе 9 приведено несколько возможных решений для сетевого уровня протокола Internet (IP). В других главах также содержатся разделы, посвященные работе сетевого уровня, поскольку маршрутизация необходима в большинстве случаев.
- **Транспортный уровень**. В эталонной модели этот уровень используется для установки сквозного соединения. Для транспортного уровня существенными являются вопросы качества обслуживания, управления потоком и насыщением, особенно в том случае, если для беспроводного соединения используются транспортные протоколы Internet (TCP и UDP).
- **Уровень приложений**. Наконец, поверх всех упомянутых ранее уровней, ориентированных на передачу, находятся приложения (вместе с уровнями, их поддерживающими). То, что вызывает для них интерес, — это местонахождение той или иной службы, поддержка мультимедийных приложений, адаптивные приложения, которые могут справляться с большими изменениями характеристик передачи, а также беспроводный доступ в Internet с портативных устройств. Одним из наиболее требовательных видов приложений является передача видеоизображения. Заинтересованного читателя мы отсылаем к работе [Bahl, 1998].

Еще одним очень важным вопросом является исследование воздействия беспроводных коммуникаций на здоровье людей. До сих пор до конца не понятно, в какой мере электромагнитные волны, испускаемые беспроводными устройствами, могут влиять на различные органы человека. Микроволновые печи и беспроводные ЛВС работают на одинаковой частоте 2,4 ГГц. Мощность излучения беспроводной ЛВС (скажем, 1 Вт) очень низка по сравнению с уровнем излучения микроволновой печи (800 Вт внутри печи). Однако, необходимы дальнейшие исследования, имеющие целью изучить влияние на человека долговременного маломощного излучения [Lin, 1997].

1.6. Общий обзор содержания книги

Данная книга построена по принципу обсуждения материала в направлении снизу вверх, как показано на рис. 1.7. В **главе 2** изложены основные концепции беспроводной передачи. Рассматриваемые в этой главе вопросы включают используемые для связи частоты, характеристики сигналов, антенны, особенности распространения сигналов и анализ нескольких основных схем уплотнения и модуляции. Материал не требует ни наличия глубоких электротехнических инженерных знаний, ни понимания всех физических основ функционирования беспроводных систем связи. Цель этой главы состоит скорее в том, чтобы помочь читателю понять некоторые решения, речь о которых пойдет при обсуждении высших уровней мобильных систем коммуникаций.

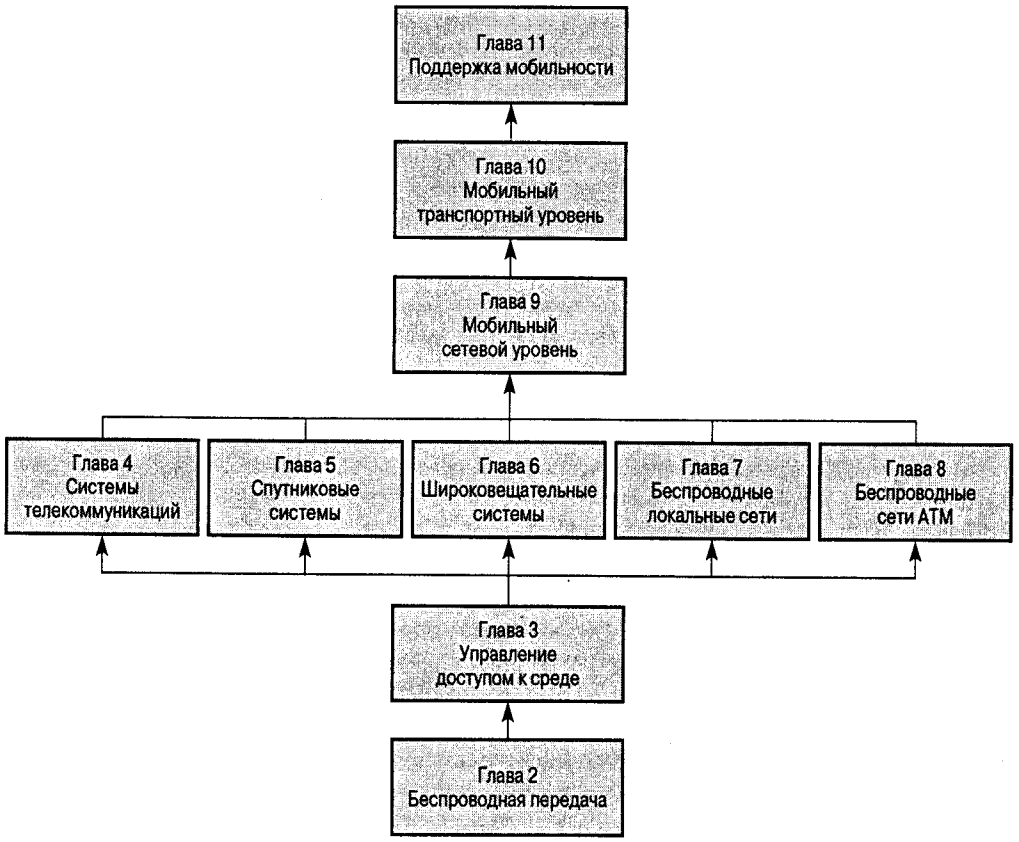

Рис. 1.7. Общая структура книги

В **главе 3** рассмотрен широкий спектр технологий доступа к среде. В начале главы объясняется, почему технологии доступа к среде для стационарных сетей нельзя перенести на беспроводные сети. Дальше обсуждаются некоторые проблемы беспроводных терминалов, использующих "пространство" в качестве общей среды. Рассматриваются методы доступа, которые применяются для различных целей: беспроводная мобильная телефонная связь, при которой доступ управляется центральной базовой станцией, и полностью децентрализованные эпизодические сети, где нет каких-либо выделенных станций. В этой главе также показано, как для доступа к среде можно использовать обсуждавшиеся в главе 2 схемы уплотнения. Особое внимание уделяется множественному доступу с кодовым разделением каналов (CDMA) как одному из перспективных методов будущего. Кроме того, рассматриваются варианты системы Aloha и схемы резервирования, пришедшие из спутниковых сетей.

После изучения материала главы 3 читатель может перейти к чтению любой из глав 4–8. В этих главах рассматриваются отдельные беспроводные системы передачи, использующие базовые технологии, обсуждавшиеся в главах 2 и 3, которые обеспечивают передачу данных на высшие уровни. **Глава 4** охватывает беспроводные системы, которые берут свое начало в индустрии телекоммуникаций и имеют соответствующие стандарты. Это, с одной стороны, система GSM — одна из наиболее успешных цифровых мобильных телефонных систем, и с другой — цифровые расширенные беспроводные телекоммуникации DECT — пример локальной беспроводной телекоммуникационной системы. В главе опи-

саны стандарты, частоты, службы, методы доступа и архитектура этих систем. Кроме того, обсуждается переход с передачи речи на пакетную передачу. Завершает главу обсуждение системы UTMS — кандидата в мобильные системы будущего.

Глава 5 представляет собой введение в спутниковые коммуникации. Опять же, предметом обсуждения здесь являются не технические детали, а потенциал спутниковых систем как средств обмена данными во всем мире. **Глава 6** посвящена относительно молодой области беспроводных коммуникаций — цифровым широковещательным системам. Широковещательные системы передают данные только в одном направлении, но зато обеспечивают новые услуги: цифровое радио (DAB), цифровое телевидение и видео (DVB). Очень важной особенностью этих систем является возможность передачи произвольных мультимедийных данных в форматах, известных нам из среды Internet. Система DAB — одна из немногих систем, способных передавать потоки мультимедийных данных на транспортные средства, которые двигаются со скоростью до 900 км/ч. Большинство других мобильных систем перестают работать уже при скорости 100 км/ч.

В **главе 7** рассматриваются три варианта беспроводных ЛВС — стандарты IEEE 802.11 и HIPERLAN, а также уже устаревшая система Bluetooth. После общего введения в проблемы проектирования беспроводных ЛВС подробно обсуждаются физический уровень, уровень управления доступом к среде и службы, предлагаемые в каждом из трех типов ЛВС. Особый интерес представляет сравнение различных подходов, положенных в основу стандартов IEEE 802.11 и HIPERLAN. Средняя скорость надежной передачи данных в первой системе около 2 Мбит/с, тогда как другая система обеспечивает гарантированное качество обслуживания (QoS) и скорость передачи данных до 23 Мбит/с. Тем не менее, все выпускаемое оборудование совместимо именно со стандартом IEEE 802.11, а для стандарта HIPERLAN пока не объявлено ни об одном серийно выпускаемом продукте. В некоторых случаях использование специальных частных решений и определенных отступлений от стандарта устройства IEEE 802.11 позволяет достичь скорости обмена до 10 Мбит/с. Система Bluetooth отличается от обеих этих систем, так как преимущественно она предназначена только для ближней связи, использующей очень дешевое оборудование. В этом случае обеспечивается скорость передачи данных лишь в несколько Кбит/с.

Глава 8 посвящена более сложным вопросам, а именно: беспроводному асинхронному режиму передачи (WATM). Режим ATM, как технология для стационарных сетей, предоставляет целый ряд гарантий качества обслуживания (QoS). Именно это послужило важнейшей причиной разработки беспроводных систем ATM. В главе 8 описана эталонная модель системы ATM и ее отдельные уровни, взаимодействие этой системы со стационарными сетями и шаги, необходимые для создания широкополосных сетей радиодоступа (BRAN). Однако роль, которую в будущем будут играть беспроводные сети с режимом ATM, пока еще не совсем понятна специалистам.

В **главе 9** представлены решения для мобильного сетевого уровня. Этот уровень может использоваться поверх различных технологий передачи информации, обсуждавшихся в главах 4—8. Хотя главной темой главы 9 является протокол Mobile IP, в ней также обсуждаются такие механизмы, как протокол динамической настройки узлов (DHCP) и принципы маршрутизации в эпизодических сетях. Поскольку протокол IP доминирует в сетях передачи данных (он является основным протоколом Internet), естественно было бы расширить этот протокол и на случай мобильного домена. В главе рассмотрены проблемы, связанные с протоколом IP и мобильностью узлов, а также приведены некоторые из возможных решений. Начиная с главы 9, термин "мобильность" используется чаще чем "беспроводная передача", так как для протоколов высших уровней не имеет значения, поддерживается ли мобильность за счет технологий беспроводной передачи либо за счет подключения переносного компьютера пользователя к сети по его

текущему местонахождению. Проблемы прерванных соединений и изменения точек доступа к сети в обоих случаях остаются практически одинаковыми.

Для того, чтобы работать с Internet-приложениями, например, службой пересылки файлов (FTP) или дистанционным входом в систему (Telnet), пользователям иногда необходимо установить надежное соединение между двумя узлами сети. В Internet это требование обеспечивается протоколом TCP. В **главе 10** обсуждаются изменения в протоколе TCP, которые необходимо сделать, чтобы этот протокол можно было использовать в мобильной среде. Имеется несколько решений, у каждого из которых есть свои сильные и слабые стороны. Это одна из причин того, что стандарта "Mobile TCP" до сих пор не существует.

Мобильные сети бесполезны без приложений, их использующих. В **главе 11** обсуждается несколько проблем, возникающих при использовании в мобильной среде приложений, рассчитанных на стационарные сети, а также рассматривается несколько новых разработок. Кроме всего прочего, эти приложения имеют дело с файловыми системами. Сохранение согласованности данных в распределенной системе является важнейшей из проблем. Как обеспечить согласованность в случае разъединения? Если нельзя обеспечить согласованность данных, может, лучше вообще запретить доступ к ним? В главе 11 обсуждается несколько систем, сталкивающихся с подобными проблемами.

Значительную часть информации в Internet составляет подсистема мультимедийных документов и гиперссылок между ними (WWW). В настоящее время доступ к WWW предполагает наличие соединения с высокой пропускной способностью, персонального компьютера с большим объемом оперативной памяти и мощным процессором, цветного монитора с высоким разрешением, а также требует установки многочисленных программных надстроек. Мобильные устройства очень ограничены в потреблении энергии, соответственно они обладают меньшей памятью и менее мощными процессорами. Кроме того, эти устройства должны быть портативными, поэтому устройства ввода ограничены в размерах и неудобны, а дисплеи имеют невысокое разрешение. В прошлом было предпринято несколько попыток устранить такую очевидную несогласованность двух рассматриваемых технологий. Мы рассмотрим один из последних подходов, а именно: протокол беспроводных приложений (WAP), который поддерживается многими компаниями.

Во всех главах помимо собственно технологических вопросов обсуждаются и проблемы безопасности. Механизмы безопасности важны на всех уровнях коммуникационной системы. Разные пользователи и даже разные нации имеют различные понятия о безопасности. Однако, совершенно очевидно, что системы связи, передающие через эфир личную информацию, должны удовлетворять особым требованиям к защите. Компаниям необходимо обезопасить свои коммуникации от прослушивания конкурентами, да и многим людям не нравится, что соседи могут слышать их личные разговоры. Беспроводные системы особенно уязвимы в этом отношении, ведь для прослушивания даже не потребуется какой-либо специальной подслушивающей аппаратуры. Поэтому, чтобы гарантировать конфиденциальность, необходимо применять специальные методы шифрования. Есть и другие механизмы поддержки конфиденциальности — аутентификация, соблюдение анонимности, защита от воспроизведения и т. д.

В конце каждой главы имеется подборка упражнений для закрепления материала и приводится рекомендуемая дополнительная литература. Завершает книгу полный список всех использовавшихся в ней аббревиатур.

1.7. Упражнения

1. Выясните текущее количество абонентов в различных системах. Учтите, что из-за быстрого развития мобильных коммуникаций любое опубликованное число всегда будет устаревшим!

1.8. Рекомендуемая литература

Bahl, S. (1998) "Supporting digital audio in a managed wireless network", *IEEE Communication Magazine*, 36, (6).

Department of Trade and Industry (1987) *Performance Specification: Radio equipment for use at fixed and portable stations in the cordless telephone service*, MPT 1334.

Electronic Industries Association (1989) *Mobile land station compatibility specification*, ANSI/EIA/TIA Standard 553.

ETSI (1994) *Common air interface specification to be used for the interworking between cordless telephone apparatus in the frequency band 864.1 MHz to 868.1 MHz, including public access services*, European Telecommunications Standards Institute, I-ETS 300 131 (1994–11).

ETSI (1998) *Digital Enhanced Cordless Telecommunications (DECT), Common Interface (CI)*, European Telecommunications Standards Institute, EN 300 175, V1.4.1 (1998–02).

GSM Data, Intel Corporation (1998) `http://www.gsmdata.com/`.

Halsall, F. (1996) *Data communications, computer networks and open systems*. Addison-Wesley.

Iridium Corporation (1998) `http://www.iridium.com/`.

International Telecommunication Union (ITU) (1998) *International Mobile Telecommunications*, (сборник рекомендаций), `http://www.itu.int/imt/`.

Keshav, S. (1997) *An engineering approach to computer networking*, Addison-Wesley.

Lin, J.C. (1997) "Biological aspects of mobile communication fields", (сборник статей), *Wireless Networks*, J.C. Baltzer, 3 (6).

Tanenbaum, A. (1996) *Computer Networks*, Prentice-Hall.

Telecommunications Industries Association (TIA) (1993a) *Mobile station land station compatibility specification for dual-mode narrowband analog cellular technology*, Interim Standard 88.

Telecommunications Industries Association (1993б) *Mobile station base station compatibility standard for dual-mode wideband spread spectrum cellular systems*, Interim Standard 95.

Telecommunications Industries Association (1996) *800 MHz TDMA cellular radio interface mobile station base station compatibility*, Interim Standard 136A.

Беспроводная передача 2

В этой книге основное внимание уделяется высшим уровням протоколов мобильных коммуникаций и элементам компьютерных наук, а не электротехническим вопросам радиосвязи и передачи данных. Поэтому в данной главе представлены лишь те фундаментальные аспекты беспроводной передачи информации, которые необходимы для понимания проблем, возникающих на высших уровнях протоколов в связи с различными неполадками в ходе передачи. Для читателей, заинтересованных в более детальном рассмотрении какой-то проблемы, приводятся ссылки на соответствующую литературу. Реальные характеристики беспроводной передачи информации в этой главе иногда сознательно упрощаются, чтобы избежать размывания общей картины.

Передача информации через различные кабели обычно не приводит к взаимным помехам, однако, для беспроводной передачи данных ситуация совершенно иная. Поэтому распределение использующихся для передачи частот подлежит строгой регламентации. Первый раздел содержит общий обзор таких частот. В следующих разделах приводятся основные сведения об антеннах, сигналах и их распространении. Изменение характеристик распространения по ходу радиопередачи создает определенные осложнения и часто приводит к появлению ошибок передачи. Поскольку среда всегда является общей, особую важность имеют вопросы уплотнения. Различные схемы уплотнения используются для того, чтобы уменьшить взаимные помехи различных передатчиков.

Чтобы передавать данные на определенных (несущих) частотах, используется модуляция. Стандартным схемам модуляции посвящен отдельный раздел этой главы. Данные схемы применяются в различных беспроводных системах коммуникаций, рассмотренных в главах 4–8. В следующем за ним разделе обсуждается расширение спектра — специальная технология передачи, более устойчивая к ошибкам. Завершает главу краткое введение в сотовые системы связи.

2.1. Частоты радиопередачи

Радиопередача может производиться с использованием различных полос частот. На рис. 2.1 показана приблизительная схема распределения спектра электромагнитных волн, используемого для передачи данных. Эта схема охватывает частоты от 300 Гц до 300 ТГц.

Длина волны λ непосредственно связана с частотой уравнением

$$\lambda = c/f,$$

где $c \cong 3\times10^8$ м/с (скорость света в вакууме) и f — частота. В обычных проводных сетях для передачи частот порядка нескольких сотен кГц используется медная витая пара, а для частот порядка нескольких сотен МГц — коаксиальный кабель (новые схемы кодирования позволяют работать на частоте в несколько сотен МГц даже при использовании медной витой пары). Далее, для частот вплоть до нескольких сотен ТГц, используется волоконная оптика. В этом случае,

однако, чаще говорят не о частотах, а о длинах волн, которые равны, например, 1500 мкм, 1350 мкм и т.д. (инфракрасный диапазон).

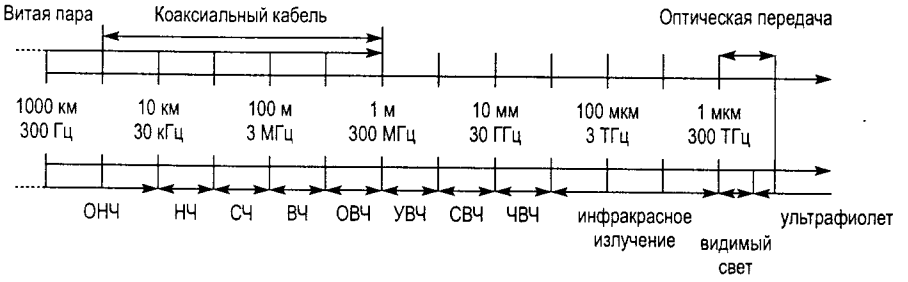

Рис. 2.1. Спектр частот

Радиопередача начинается с диапазона **очень низких частот** (ОНЧ) — порядка нескольких кГц. Это диапазон очень длинных волн. Волны низких частот могут проникать через поверхность воды и способны огибать поверхность земли, поэтому используются для связи с подводными лодками. На этих частотах в диапазоне длинных волн (ДВ) до сих пор работают несколько радиостанций, например, в Германии (между 148,5 КГц и 283,5 КГц). Сотни радиостанций используют диапазоны **средних частот** (СЧ) и **высоких частот** (ВЧ) для радиовещания с амплитудной модуляцией (**АМ**). Для этой цели выделяются диапазон средних волн (СВ — между частотами 520 кГц и 1605,5 кГц) и несколько диапазонов коротких волн (КВ — между частотами 5,9 МГц и 26,1 МГц). Кроме того, в диапазоне от 87,5 МГц до 108 МГц ведутся радиопередачи с частотной модуляцией (FM). Частоты, ограничивающие указанные выше диапазоны, обычно фиксируются национальными стандартами и могут различаться в разных странах. Для (любительских) радиопередач по всему миру чаще всего используются короткие волны. Это обусловлено тем, что они отражаются от ионосферы. Мощность передачи может достигать 500 кВт, что очень много в сравнении с 1 Вт для мобильного телефона.

Телевизионные станции работают на более высоких частотах. Для передачи типичных аналоговых телевизионных сигналов используются диапазоны 174–230 МГц и 470–790 МГц, т.е. полосы очень высоких частот (ОВЧ) и ультравысоких частот (УВЧ). В этой области также работает система DAB (223–230 МГц и 1452–1472 МГц). Планируется, что этот же диапазон будет использовать цифровое телевидение (470–862 МГц), для которого выделят некоторые из старых частот аналогового телевидения. Кроме того, частоты УВЧ используются для аналоговой мобильной телефонной связи (450–465 МГц), на них работают цифровая система GSM (890–960 МГц, 1710–1880 МГц), цифровые беспроводные телефонные системы стандарта DECT (1880–1900 МГц) и многие другие системы. Диапазоны ОВЧ и особенно УВЧ позволяют использовать в мобильных телефонах очень небольшие антенны и устанавливать при этом относительно надежные соединения.

Сверхвысокие частоты (СВЧ) обычно используются для создания направленных микроволновых каналов (приблизительно 2–40 ГГц) и стационарных спутниковых служб в С-полосе (4 и 6 ГГц), Q-полосе (11 и 14 ГГц) и К-полосе (19 и 29 ГГц). Разрабатывается несколько систем в области **чрезвычайно высоких частот** (ЧВЧ), близкой к инфракрасному диапазону. Во избежание взаимных помех, использование всех радиочастот регулируется на государственном уровне. В Германии, например, регулирование охватывает частоты от 9 кГц до 275 ГГц.

Еще более высокие частоты применяются для организации оптической передачи. Ее можно использовать не только в волоконных оптических каналах, но и для беспроводной

связи. **Инфракрасная** (IR) передача используется для установки направленных соединений (например, для организации связи между различными зданиями с помощью лазерных каналов). Самая распространенная IR-технология, IrDA работает с длинами волн около 850–900 нм и применяется для соединений переносных компьютеров, персональных цифровых помощников (PDA) и т.д. Наконец, видимый свет использовался человечеством для беспроводной передачи на протяжении тысячелетий. Хотя такая связь и не очень надежна из-за помех, она все же полезна, поскольку в этом случае "приемники" встроены непосредственно в человеческое тело.

2.1.1. Регулирование использования частот

Как показывают примеры из предыдущего раздела, ресурсы радиочастот весьма ограничены. Кроме того, многие национальные (экономические) интересы усложняют задачу общего всемирного регулирования использования частот. За всемирную координацию действий в области телекоммуникаций (проводных и беспроводных) отвечает Международный телекоммуникационный союз (ITU), базирующийся в Женеве. Союз ITU является подразделением ООН. Сектор радиосвязи ITU (ITU-R) (предыдущее название этой организации — Международный консультативный комитет по радиовещанию (CCIR)) занимается стандартизацией беспроводной связи, в частности, планированием использования частот.

Чтобы достичь хотя бы частичного успеха во всемирной координации и отразить национальные интересы, мир был разделен комитетом ITU-R на три региона. В **регион 1** входят Европа, Ближний и Средний Восток, постсоветские государства и Африка. **Регион 2** охватывает Северную и Южную Америку, а также Гренландию. Дальний Восток, Австралия и Новая Зеландия были включены в **регион 3**. Внутри этих регионов за дальнейшее регулирование использования частот ответственны национальные ведомства. Например, в США эти обязанности закреплены за Федеральной комиссией связи (FCC), а в Европе несколько стран имеют общий орган — Европейскую конференцию почтовых и телекоммуникационных ведомств (CEPT). Хотя CEPT и отвечает за общее планирование, многие задачи были переданы другим ведомствам (что делает процесс регулирования весьма запутанным). К примеру, Европейский институт стандартов по телекоммуникациям (ETSI) отвечает за стандартизацию и объединяет национальные органы стандартизации, государственных поставщиков, производителей, группы пользователей и исследовательские институты.

Комитет ITU-R периодически проводит конференции (типичный пример — Всемирная конференция по радиосвязи, WRC), на которых обсуждается и утверждается распределение частот во всех трех регионах. Достижение полной гармонии в использовании частот — довольно сложная задача. Поскольку во многих странах и регионах уже развернуты определенные технологии, такие страны изменяют сложившееся распределение частот очень неохотно. Но для использования спутниковых коммуникаций согласование использования частотных диапазонов просто необходимо. Спутники, особенно новое поколение спутников на низких орбитах (см. главу 5), не "считаются" с национальными нормами и действуют во всемирном масштабе.

В табл. 2.1 приведено несколько примеров частот, используемых для мобильной телефонной связи (аналоговой и цифровой), беспроводных телефонов и беспроводных ЛВС во всех трех регионах. Более старые системы, например, Скандинавская мобильная телефонная система (NMT), действуют не во всей Европе и могут работать на различных частотах в разных государствах. Более новые (цифровые) системы (стандартизованные институтом ETSI) являются совместимыми во всей Европе.

Таблица 2.1. Примеры распределения частот

	Европа	США	Япония
Мобильные телефоны	**NMT** 453–457 МГц, 463–467 МГц; **GSM** 890–915 МГц, 935–960 МГц; 1710–1785 МГц, 1805–1880 МГц	**AMPS, TDMA, CDMA** 824–849 МГц 869–894 МГц; **GSM, TDMA, CDMA** 1850–1910 МГц, 1930–1990 МГц	**PDC** 810–826 МГц, 940–956 МГц, 1429–1465 МГц, 1477–1513 МГц
Беспроводные телефоны	**CT1+** 885–887 МГц, 930–932 МГц; **CT2** 864–868 МГц; **DECT** 1880–1900 МГц	**PACS** 1850–1910 МГц, 1930–1990 МГц; **PACS-UB** 1910–1930 МГц	**PHS** 1895–1918 МГц; **JCT** 254–380 МГц
Беспроводные ЛВС	**IEEE 802.11** 2400–2483 МГц; **HIPERLAN 1** 5176–5270 МГц	**IEEE 802.11** 2400–2483 МГц	**IEEE 802.11** 2471–2497 МГц

Хотя более старые аналоговые **мобильные телефонные** системы (например, система NMT или ее производные, работающие на частоте 450 МГц) еще действуют, в Европе господствует полностью цифровая система **GSM** (см. раздел 4.1), работающая на частотах 900 МГц и 1800 МГц. Система GSM на частоте 1800 МГц также известна под названием DCS1800 — цифровой сотовой системы. В США, в отличие от Европы, комиссией FCC было разрешено использование нескольких сотовых технологий в одинаковых полосах частот около 850 МГц. Это привело к сосуществованию нескольких различных решений. Вначале на таких частотах действовала лишь аналоговая расширенная мобильная телефонная система AMPS. Позже возникли мобильные телефонные системы, работающие в дуальном режиме (стандарт IS-54), поддерживающие цифровые услуги множественного доступа с временным разделением (TDMA) и аналоговую систему AMPS. Кроме того, были разработаны стандарт IS-136 для цифровых систем TDMA (который также известен как NA-TDMA или Североамериканский TDMA) и стандарт IS-95 для телефонных систем множественного доступа с кодовым разделением каналов (CDMA). Таким образом, в США не приняли на государственном уровне общей мобильной телефонной системы, а положились на рыночные механизмы. Возникло много различных систем и поэтому в США, в отличие от Европы, до сих пор не обеспечивается полное покрытие. Долгое обсуждение различных достоинств и недостатков систем CDMA и TDMA в конце концов привело ко всемирному успеху системы GSM. Сейчас GSM используют более чем в 120 странах, причем пользователи могут путешествовать с одним и тем же мобильным телефоном в Зимбабве, Узбекистане, Швеции, Сингапуре, Тунисе, России, Италии, Греции, Германии, Китае, Бельгии, Австрии и многих других странах.

В Японии была установлена персональная цифровая сотовая система (PDC), известная раньше под названием японской цифровой сотовой системы (JDC). Первые мобильные телефоны, охватывающие системы по всему миру, появились в 1999 году.

В мире существует много различных стандартов **беспроводных телефонов**. Однако, это разнообразие не приводит к такому количеству проблем, как разные стандарты мобильной телефонной связи. Все еще используются некоторые старые аналоговые системы, например, **CT1+**. Кроме того, для беспроводных телефонов часто применяются и цифровые технологии. Здесь

можно привести несколько примеров. Первым цифровым стандартом беспроводных телефонов был стандарт **CT2**, появившийся в Великобритании. В Европе действует стандарт цифровых расширенных беспроводных телекоммуникаций **DECT** (см. раздел 4.2), а в США — система связи персонального доступа PACS (в нелицензируемой полосе частот — **PACS-UB**). В Японии персональная система переносных телефонов PHS заменила аналоговую японскую систему беспроводных телефонов **JCT**. В этих системах уже обеспечивается взаимодействие с мобильными телефонами, в частности, со стандартами DECT и GSM.

Стандарты для беспроводных ЛВС (**WLAN**) представляют особый интерес для организации беспроводной мобильной связи между компьютерами в университетских городках или зданиях. Сейчас разработано несколько систем в свободной от лицензирования полосе **ISM**. Самые привлекательные из них используют частоту 2,4 МГц и доступны для свободной от лицензирования деятельности во всем мире (хотя ограничивающие частоты, мощность передачи и т. п. могут быть разными для различных государств). Здесь самым распространенным является стандарт **IEEE 802.11** , который будет обсуждаться в главе 7. Европейский стандарт беспроводных ЛВС **HIPERLAN 1** работает на частоте 5,2 МГц, но пока непонятно, где будет доступна эта полоса частот.

Многие полосы частот были закреплены за радиомагистралями (например, трансевропейская радиомагистраль TETRA, в зависимости от государства, использует полосы 380–400 МГц, 410–430 МГц, 450–470 МГц), пейджинговыми и спутниковыми службами, полетной телефонной системой TFTS (1670–1675 МГц и 1800–1805 МГц) и т.д. Кроме того, планируется выделение новых частот в диапазонах 1885–2025 МГц и 2110–2200 МГц для организации универсальной системы мобильных телекоммуникаций UTMS. Несмотря на сильное экранирование более высоких частот различными препятствиями, такие частоты представляют особый интерес для высокоскоростной передачи. В связи с этим рассматриваются свободные от лицензирования полосы частот около 5,7 ГГц, 17,2 ГГц, 24 ГГц и даже 61 ГГц.

2.2. Сигналы

Сигналы являются физическим представлением данных. Обмен данными между пользователями системы связи возможен только посредством передачи сигналов. В основной эталонной модели ISO/OSI за превращение данных (т.е. битов) в сигналы отвечает уровень 1 [Halsall, 1996].

Сигналы являются функциями координат и времени. Значения данных выражаются через параметры сигнала. Периодические сигналы, и особенно синусоидальные волны (несущая), — наиболее интересные типы сигналов для радиопередачи. (Способы отображения данных на носители объясняются в разделе 2.6.) Общий вид синусоидальной волны следующий:

$$g(t) = A^t \sin(2\pi f_t t + \varphi_t)$$

Параметрами сигнала являются **амплитуда** A, **частота** f и **фазовый сдвиг** φ. Амплитуда, как множитель в функции g, может изменяться со временем (см. раздел 2.6.1), поэтому и введено обозначение A_t. Частота f выражает периодичность сигнала. Период связан с частотой соотношением $T = 1/f$. В уравнениях часто используется специальное обозначение $2\pi f = \omega$. Частота f также может зависеть от времени (см. раздел 2.6.2). Наконец, фазовый сдвиг определяет сдвиг сигнала относительно такого же сигнала без сдвига. Пример сдвига функции приведен на рис. 2.2. Здесь показана синусная функция без сдвига и такая же функция (т.е. с той же амплитудой и частотой), но с фазовым сдвигом φ. Представление данных с помощью сдвига фазы обсуждается в разделе 2.6.3.

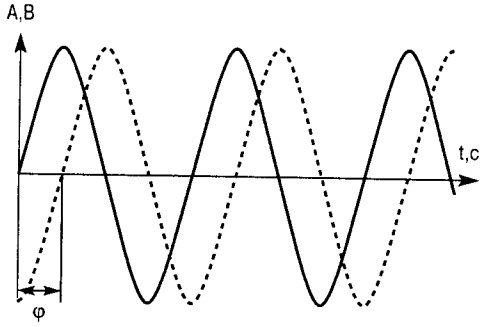

Рис. 2.2. Представление сигнала на временной диаграмме

Синусоидальные волны представляют для нас особый интерес, так как с помощью **разложения Фурье** любой периодический сигнал можно выразить через набор синусных и косинусных функций:

$$g(t) = \frac{1}{2}c + \sum_{n=1}^{\infty} a_n \sin(2\pi nft) + \sum_{n=1}^{\infty} b_n \cos(2\pi nft)$$

Здесь параметр c определяет **постоянную составляющую** сигнала, а коэффициенты a_n и b_n являются амплитудами n-х синусных и косинусных функций. Это уравнение показывает, что для представления произвольной периодической функции необходимо бесконечное число синусных и косинусных функций. Частоты этих функций (так называемые **гармоники**) возрастают с увеличением параметра n и являются кратными **основной частоте** f. Однако, поскольку полоса пропускания любой среды, воздуха, кабелей, передатчиков, и т.д. ограничена, для спектра передаваемых частот существует верхний предел. Поэтому в действительности для представления периодических функций достаточно рассмотреть конечное число синусных и косинусных функций — все системы передачи обладают ограниченными полосами пропускания и не могут передавать произвольные периодические сигналы. Для наших целей достаточно знать, что периодический сигнал состоит из одной или нескольких синусоидальных функций. Далее все объяснения будут проводиться на примере одной синусоидальной функции, т.е. для случая единственной частоты.

Обычно сигналы представляются на временной диаграмме (см. рис. 2.2). В этом случае изменение величины сигнала рассматривается во времени (время чаще всего измеряется в секундах (c), а амплитуда сигнала в других единицах — например, в вольтах (B)). Именно с таким представлением сигналов мы имеем дело при работе с осциллографом. Как уже упоминалось выше, это представление позволяет отобразить и фазовый сдвиг сигналов (φ).

Однако, когда сигнал состоит из множества различных частот (как в случае использования разложения Фурье), представление его на временной диаграмме является неудобным. В этом случае лучше работать с представлением сигнала на **частотной диаграмме** (см. рис. 2.3). Здесь рассматривается зависимость амплитуды отдельной частотной составляющей сигнала от частоты. На рис. 2.3 изображен лишь один пик, поэтому сигнал содержит только одну частотную составляющую (иначе говоря, единственную синусную функцию). Произвольные периодические функции могут иметь много различных пиков. Соответствующее изображение называют частотным спектром сигнала и наблюдают с помощью спектрального анализатора. Математическими инструментами для перехода от временного представления сигнала к частотному и обратно являются прямое и обратное преобразования Фурье.

Фазовая диаграмма, показанная на рис. 2.4, является третьим способом представления сигнала. В этом представлении, называемом также диаграммой фазового состояния, амплитуда

сигнала M и его фаза φ изображаются в полярных координатах. (Длина вектора обозначает амплитуду, а угол — фазовый сдвиг сигнала.) Ось X соответствует фазе 0 (**синфазная составляющая**). Фазовый сдвиг 90° изображается точкой на оси Y (**квадратурная составляющая**).

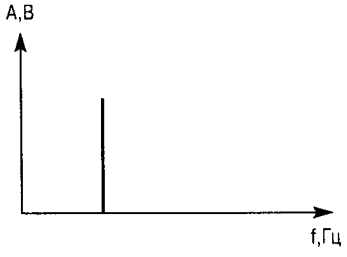

Рис. 2.3. Представление сигнала на частотной диаграмме

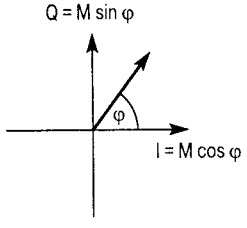

Рис. 2.4. Представление сигнала на диаграмме фазового состояния

2.3. Антенны

Само слово "беспроводный" указывает на то, что этот режим передачи обходится без проводов и передача сигналов осуществляется непосредственно через пространство. Для передачи электромагнитных волн не требуется никакой среды (например, эфира). Однако, все же необходимо как-то связать передатчик с внешним миром и, наоборот, внешний мир с приемником. Эта задача решается с помощью **антенн**. Антенны представляют собой устройства, которые принимают и излучают энергию в виде электромагнитных волн. Приемник и передатчик могут быть связаны со своими антеннами проводами, кабелями или любыми другими подходящими проводниками.

Теоретической эталонной антенной является **изотропный излучатель** — точка в пространстве, излучающая одинаковую мощность во всех направлениях. Иначе говоря, все точки передаваемой электромагнитной волны с одинаковой мощностью излучения находятся на сфере, в центре которой размещена антенна. Таким образом, **диаграмма направленности** изотропного излучателя симметрична по всем направлениям (рис. 2.5).

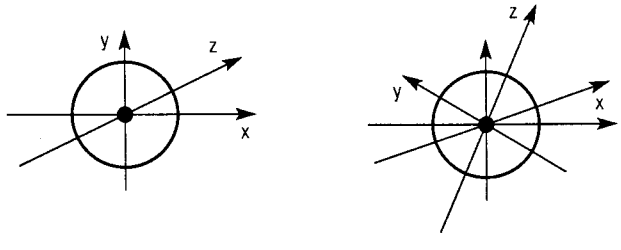

Рис. 2.5. Диаграмма направленности идеального изотропного излучателя

Однако, в действительности такой антенны не существует. В реальных антеннах всегда присутствуют **эффекты направленности**, т.е. интенсивность излучения неодинакова в разных направлениях. Простейшей антенной является тонкий симметричный **вибратор** (также называемый вибратором Герца), показанный на рис. 2.6 (справа). Для достижения эффективной передачи или приема длину вибратора нельзя выбирать произвольно. Например, она может равняться половине длины волны λ передаваемого сигнала. При этом получается очень эффективное распределение энергии излучения. В случае, когда антенна устанавливается на крыше автомобиля, целесообразно использовать длину вибратора, равную λ/4.

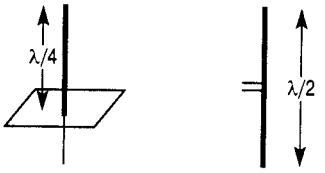

Рис. 2.6. Простые антенны

Вибратор длиной λ/2 обладает однородной диаграммой направленности (действует как **ненаправленная** антенна) в одной из плоскостей (рис. 2.7, слева). В оставшихся двух плоскостях диаграмма направленности имеет форму восьмерки (рис. 2.7, слева и в центре). При использовании антенны этого типа преодоление окружающих препятствий достигается за счет увеличения мощности сигнала. Препятствиями могут быть горы, долины, здания и т. п.

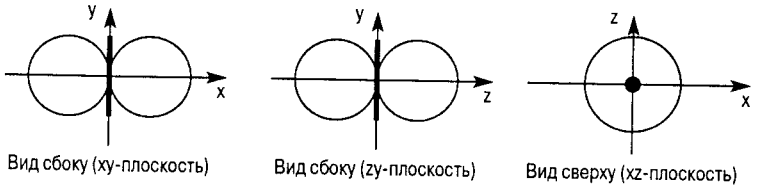

Рис. 2.7. Диаграмма направленности простого вибратора

Когда антенна размещена, например, в долине или между зданиями, однородная диаграмма направленности оказывается малоэффективной. В подобных случаях чаще используются **направленные антенны** с выделенными направлениями передачи и приема. На рис. 2.8 приведена диаграмма направленности направленной антенны с главным лепестком в направлении оси *X*. Ярким примером направленных антенн являются спутниковые антенны (тарелки).

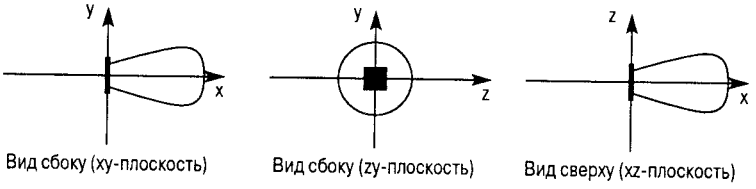

Рис. 2.8. Диаграмма направленности направленной антенны

В сотовых системах, которые будут обсуждаться в разделе 2.8, обычно применяются именно направленные антенны. Несколько направленных антенн, объединенных в одной точке пространства, образуют **секторную антенну**. Ячейку можно разбить, например, на три или шесть секторов. Подобным образом достигается возможность многократного использования частот (см. раздел 2.8). Диаграммы направленности таких секторных антенн приведены на рис. 2.9.

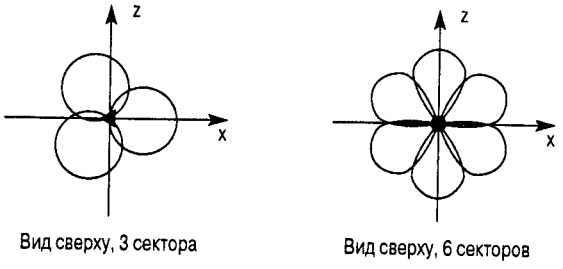

Рис. 2.9. Диаграммы направленности секторных антенн

Для улучшения приема и снижения влияния негативных эффектов многолучевого распространения (раздел 2.4.3) также можно объединить несколько антенн. Такие антенны называют **многоэлементными антенными решетками**. Этот подход позволяет использовать различные схемы разнесения. Одной из таких схем является **разнесение с коммутацией** или **разнесение с автовыбором**, когда приемник всегда использует антенный элемент с самой большой выходной мощностью. Для увеличения коэффициента направленного действия антенны применяется **сложение разнесенных сигналов**. При этом, чтобы избежать взаимного подавления, необходимо произвести коррекцию фаз сигналов. Существуют различные схемы разнесенных антенн. Две из них показаны на рис. 2.10. Первая схема (слева) состоит из отражателя и размещенных на нем двух антенн длиной λ/4. Расстояние между антеннами равно λ/2. Вторая схема (справа) объединяет три вибратора длиной λ/2, находящихся на расстояниях λ/2 друг от друга. В общем случае эти интервалы должны быть кратными λ/2.

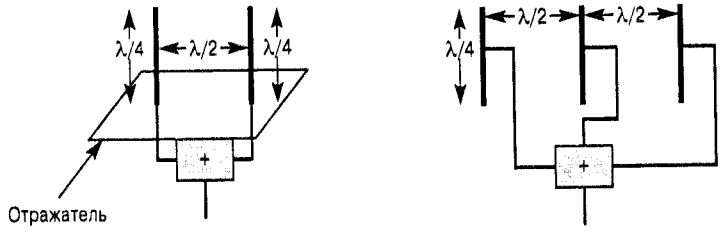

Рис. 2.10. Разнесенные антенные системы

Более эффективным решением являются **интеллектуальные антенны**, в которых использование нескольких антенных элементов комбинируется со специальной обработкой сигнала. Это позволяет оптимизировать диаграммы направленности излучения/приема в зависимости от реакции окружающей среды на сигнал. Такие антенны способны приспосабливаться к изменениям в мощности приема и условиях передачи, а также к различным условиям распространения сигналов, о чем речь пойдет в следующем разделе.

2.4. Распространение сигналов

Как и в обычных проводных цепях, в беспроводных коммуникационных сетях всегда имеются передатчики и приемники сигналов. Однако, между этими двумя типами сетей есть существенные различия в отношении распространения сигналов. В то время как в проводных сетях сигнал может распространяться только по проводам (это может быть медная витая пара, коаксиальный кабель, оптоволоконный кабель и т.д.), в беспроводных сетях нет каких-либо кабелей, определяющих направление распространения сигнала. Если провод не поврежден, он в большинстве случаев обладает одинаковыми характеристиками в любой своей точке. Поэтому поведение распространяющегося по этому проводу сигнала (например, зависимость принимаемой мощности от длины кабеля) можно точно предсказать. В случае беспроводной передачи такое предсказуемое поведение возможно лишь в вакууме, т.е. когда между передатчиком и приемником нет никакого вещества. В этом случае ситуация выглядит так, как показано на рис. 2.11.

- **Зона передачи**. В пределах некоторого радиуса возможна передача, т.е. прием сигнала осуществляется с достаточно малым количеством ошибок. Приемник способен установить связь с передатчиком, а также сам может работать как передатчик.
- **Зона детектирования**. В пределах второго радиуса возможно детектирование передачи, т.е. передаваемая мощность достаточно велика (выделяется из фонового шума). Однако, для установления связи количество ошибок оказывается слишком велико.

- **Зона помех**. Внутри третьего, еще большего, радиуса передатчик фактически препятствует другим передачам, создавая фоновый шум. Приемник не может детектировать посылаемые сигналы, поэтому они лишь создают помехи другим сигналам.

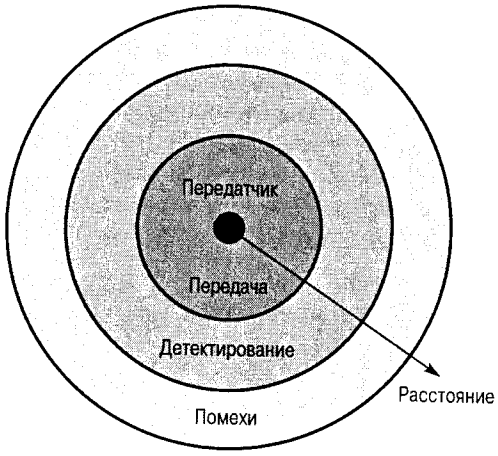

Рис. 2.11. Зоны передачи, детектирования и помех

Такая упрощенная идеальная схема иллюстрирует представление об отдельных **ячейках,** образующихся вокруг передатчика (эти вопросы кратко обсуждаются в разделе 2.8). Однако, реальная жизнь происходит не в вакууме, и при радиопередаче необходимо учитывать наличие атмосферы, гор, зданий, возможное перемещение передатчиков и приемников и т.д. Три описанных выше окружности на практике превращаются в многоугольники, форма которых изменяется со временем и зависит от частоты передачи. В следующих разделах рассматривается несколько возникающих в этом случае проблем, демонстрирующих различия между беспроводной и проводной передачами.

2.4.1. Потери на пути распространения

В свободном пространстве радиосигналы распространяются так же, как и свет (независимо от частоты), т.е. по прямой линии (без учета гравитационных эффектов). Если между передатчиком и приемником можно провести подобную линию, то ее называют **линией прямой видимости** (**ЛПР**). Однако даже когда между передатчиком и приемником нет никакого вещества (вакуум), мощность сигнала все равно подвержена **потерям в свободном пространстве**. Если обозначить расстояние между передатчиком и приемником через d, то принимаемая мощность P_n будет пропорциональна $1/d^2$ (**закон обратных квадратов**). Причина этого явления очень проста. Представим, что передатчик — это точка в пространстве. После излучения передатчиком сигнала с определенной энергией он распространяется в пространстве со скоростью света как волна сферической формы. Если нет никаких препятствий, то радиус сферы постоянно возрастает, и энергия сигнала равномерно распределяется по ее поверхности. Площадь же поверхности сферы S возрастает с увеличением радиуса d согласно уравнению $S = 4\pi d^2$.

Даже когда между передатчиком и приемником нет никакого вещества, все равно необходимо учитывать некоторые дополнительные параметры. В частности, принимаемая приемником мощность зависит от длины волны и коэффициентов направленного действия антенн передатчика и приемника. Когда же между передатчиком и приемником появляется вещество, ситуация еще более усложняется. Большинство радиопередач ведется через атмосферу, поэтому сигналы распространяются через воздух, дождь, снег, туман, пыль, смог и т.п. Хотя на неболь-

ших расстояниях (например, для ЛВС, речь о которых пойдет в главе 7) такие **потери на пути распространения** или **затухание** не приводят к значительным осложнениям, передача на большие расстояния (например, спутниковая передача, глава 5) сильно подвержена влиянию атмосферы. Погодные условия, например, сильный дождь, могут влиять даже на обычную мобильную телефонную связь. Дождь поглощает большую часть излучаемой антенной энергии (похожий эффект используется для приготовления пищи в микроволновых печах), и это может привести к фактическому прекращению функционирования каналов связи.

Радиоволны способны частично проходить сквозь различные объекты. Обычно, чем ниже частота передачи, тем выше степень проникновения. Длинные волны распространяются на большие расстояния даже в воде — этот эффект используется для связи с подводными лодками. В то же время для экранирования радиоволн высокой частоты может оказаться достаточно даже одного дерева. Чем выше частота, тем больше поведение радиоволн напоминает поведение света. Это явление легко понять с помощью спектра частот, изображенного на рис. 2.10.

2.4.2. Дополнительное влияние условий на распространение сигналов

В предыдущем разделе было сказано, что в свободном пространстве радиосигналы, как и свет, всегда распространяются по прямой линии. В действительности между передатчиком и приемником радиосигналов почти никогда не бывает линии прямой видимости. Мобильные телефоны используются в больших городах с небоскребами, в горах, внутри зданий, автомобилей, и т.д. Таким образом, кроме затухания из-за расстояния, на распространении сигналов между передатчиком и приемником сказывается еще несколько эффектов, сильно зависящих от частоты.

Крайней формой затухания является **блокирование** или **экранирование** радиосигналов крупными препятствиями (см. рис. 2.12, слева). Чем выше частота сигнала, тем больше его поведение подобно распространению света. Поэтому даже небольшие объекты, например, обычная стена, грузовик на улице или деревья, способны блокировать такие сигналы. Еще один важный эффект, а именно, **отражение** сигналов, показан на рис. 2.12 справа. Если объект достаточно крупный по сравнению с длиной волны передачи (сюда можно отнести, например, большие здания, горы или поверхность земли), сигнал будет от него отражаться. Так как при этом некоторая часть мощности сигнала всегда поглощается объектом, при отражении происходит ослабление сигнала. Отражение позволяет передавать сигналы даже при отсутствии линии прямой видимости. Это типичный случай для передачи в городах или горных районах. Сигналы передатчика могут несколько раз отразиться от стен зданий, прежде чем достигнут приемника. Чем больше отражений испытывает сигнал, тем слабее он становится.

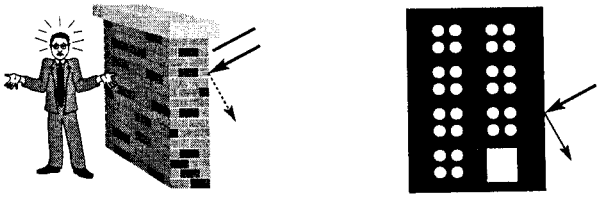

Рис. 2.12. Блокирование (экранирование) и отражение радиоволн

Экранирование и отражение обусловлены объектами, размеры которых намного превышают длину волны сигнала (в этом случае радиоволны ведут себя как "частицы"). В то же время существуют эффекты, подчеркивающие "волновой" характер радиосигналов. Если препятствие имеет размеры порядка длины волны или меньше, радиоволны могут **рассеиваться** на нем (см. рис. 2.13, слева). При рассеивании поступающий сигнал разделяется на несколько более слабых исходящих сигналов. Экспериментально это явление обычно наблюдают с помощью лазера и маленького отверстия или препятствия. В нашем случае, однако, нужно учесть, что длины волн радиопередачи, например, в

системах GSM и AMPS, составляют несколько десятков сантиметров. Поэтому источниками рассеяния сигналов может быть множество предметов нашего обычного окружения. Еще одним эффектом является **дифракция** волн. Этот эффект весьма похож на рассеяние (рис. 2.13, справа). На вершине горы радиоволны отклоняются от прямой линии и распространяются в различных направлениях. Общим результатом процессов рассеяния и дифракции является изменение мощности сигнала в зависимости от местонахождения приемника.

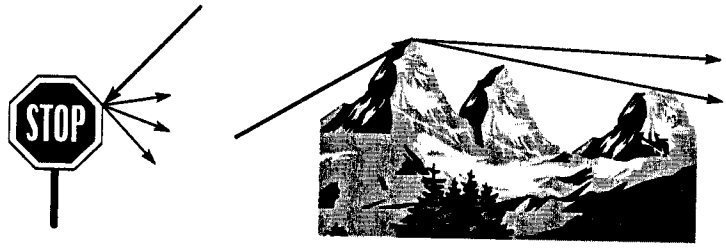

Рис. 2.13. Рассеяние и дифракция волн

2.4.3. Многолучевое распространение

Прямой передаче радиосигналов от передатчика к приемнику всегда сопутствуют эффекты распространения, рассматривавшиеся в предыдущем разделе, что является причиной одного из самых неприятных осложнений радиосвязи — возникновения **многолучевого распространения**. На рис. 2.14 показаны передатчик (слева) и возможный приемник (справа).

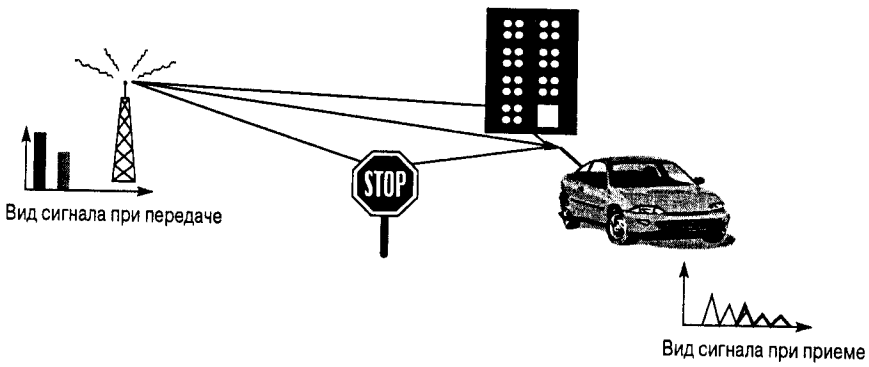

Рис. 2.14. Многолучевое распространение и межсимвольная интерференция

Радиоволны, испускаемые передатчиком, могут распространяться по прямой линии, отражаться от крупного здания или рассеиваться на меньших препятствиях. На этом упрощенном рисунке приведено лишь три возможных пути распространения сигнала. В действительности их может быть намного больше. Из-за конечности скорости света распространяющиеся по различным путям сигналы достигнут антенны приемника в разное время. Этот эффект (обуславливаемый многолучевым распространением) называется **расширением вследствие запаздывания**: начальный сигнал расширяется из-за запаздывания поступления отдельных его элементов. Расширение из-за запаздывания является обычным эффектом радиопередачи. Здесь нет проводов, направляющих волны по единственному пути. Тем не менее, похожее явление — дисперсия — происходит и при высокоскоростной оптической передаче сигналов по многомодовому оптоволокну [Halsall, 1996]. Заметим, что расширение сигнала вследствие запаздывания не зависит от движения передатчика или приемника. В городах типичное значение расширения сигнала вследствие запаздывания составляет порядка 3 мкс. Систе-

ма GSM, например, допускает расширение сигнала вследствие запаздывания до 16 мкс, т.е. более чем трехкилометровую разность хода отдельных лучей.

Как же влияет расширение задержки на сигналы, представляющие данные? Во-первых, короткий импульс может размываться. При этом он превращается в более широкий импульс или даже в несколько более слабых импульсов. На рис. 2.14 показано только три пути, поэтому импульс передатчика при приеме будет выглядеть как три меньших импульса. В реальной ситуации существуют сотни различных путей, соответственно, отдельный импульс при приеме выглядит как много более слабых импульсов. Каждому пути соответствует свое значение затухания, поэтому принимаемые импульсы имеют различную мощность. Некоторые из принимаемых импульсов слишком слабые даже для детектирования (т.е. проявляются в виде шума).

Рассмотрим теперь второй импульс, показанный на рис. 2.14. С точки зрения передатчика оба импульса разнесены. При приеме эти два импульса интерферируют, т.е. перекрываются во времени. Теперь представим, что каждый импульс соответствует символу, а один или несколько символов соответствуют биту. Тогда энергия, предназначенная для одного символа, может перетекать к соседнему символу. Этот эффект называется **межсимвольной интерференцией** (InterSimbol Interference, ISI). Чем выше требуемая скорость передачи, тем худшими будут последствия межсимвольной интерференции, так как начальные символы будут располагаться все ближе и ближе друг к другу. Межсимвольная интерференция ограничивает пропускную способность радиоканала с многолучевым распространением (а оно наблюдается в большинстве случаев). Из-за этих помех, представляющие различные символы, сигналы могут гасить друг друга, что ведет к неправильной интерпретации при приеме и, как следствие, ошибкам передачи.

В таких случаях большую пользу приносит знание характеристик канала. Если приемнику известны задержки различных путей (или, по крайней мере, главных путей распространения сигнала), он может компенсировать искажения, внесенные каналом. Можно, например, сначала передать **настроечную последовательность** сигналов, известную приемнику. Затем в приемнике происходит сравнение принятого сигнала с исходной настроечной последовательностью и коррекция искажений с помощью устройства компенсации — эквалайзера [Wesel, 1998].

Межсимвольная интерференция и расширение сигналов вследствие запаздывания имеют место даже в случае неподвижных радиопередатчиков и приемников. Ситуация еще более ухудшается, если передатчик и/или приемник находятся в движении. При этом постоянно изменяются и характеристики канала, и пути распространения сигналов. Подобный эффект часто наблюдается при прослушивании обычных аналоговых радиоприемников в автомобилях. Во время езды мощность принимаемого сигнала может произвольным образом резко изменяться. Такие быстрые изменения принимаемой мощности во времени называются **краткосрочным затуханием**. Причина в том, что из-за наличия разных путей распространения сигналы могут приходить с разным фазовым сдвигом и даже взаимно погашаться, как показано на рис. 2.15. Приемнику необходимо постоянно приспосабливаться к изменениям характеристик канала, например, изменяя параметры эквалайзера. Однако, если подобные изменения происходят слишком быстро (как, например, на автомагистрали в городе), приемник не успевает к ним приспосабливаться, и число ошибок передачи резко возрастает.

На рис. 2.15 представлен еще один эффект — **долгосрочное затухание** принимаемого сигнала. Долгосрочное затухание показано здесь на примере мощности, усредненной по времени. Оно обусловлено изменением расстояния между передатчиком и более удаленными препятствиями. В большинстве случаев передатчики могут компенсировать долгосрочное затухание, увеличивая или уменьшая мощность передачи так, чтобы принимаемый сигнал не выходил за определенные рамки.

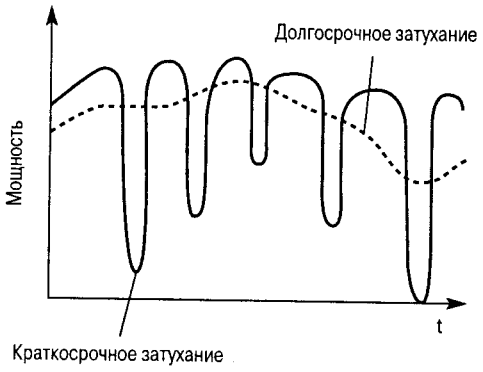

Рис. 2.15. Краткосрочное и долгосрочное затухание

Существует много других эффектов, оказывающих влияние на радиопередачу. К ним можно отнести, например, доплеровский сдвиг частоты, обусловленный движением приемника или передатчика. В случае звуковых волн его легко можно наблюдать даже при относительно низкой скорости, однако, в случае радиоволн для достижения этого эффекта передатчики или приемники должны двигаться очень быстро. Примерами таких устройств могут служить спутники (глава 5). Здесь доплеровский эффект приводит к случайным сдвигам частоты. Дополнительную информацию о характеристиках беспроводных каналов связи заинтересованный читатель сможет найти в работе [Anderson, 1995]. Нам же пока достаточно будет знать, что многолучевое распространение ограничивает пропускную способность каналов (из-за межсимвольной интерференции), а движение передатчиков и приемников приводит к дополнительным проблемам, обусловленным быстрыми изменениями характеристик каналов.

2.5. Уплотнение

Уплотнение является фундаментальным механизмом не только в системах связи, но и в практике повседневной жизни. Уплотнение позволяет нескольким пользователям совместно использовать некоторую среду, сводя взаимные помехи к минимуму. В качестве примера из повседневной жизни можно рассмотреть автомагистраль с несколькими полосами движения. Много пользователей (водителей автомобилей) используют одну и ту же среду (автомагистраль) почти без помех (аварий). Это оказывается возможным благодаря созданию нескольких полос движения (уплотнению с пространственным разделением). Кроме того, различные автомобили могут использовать одну и ту же среду (полосу) в разные моменты времени (уплотнение с временным разделением).

Эти простые примеры иллюстрируют повседневное использование методов уплотнения. Ниже рассматривается применение средств уплотнения в системах беспроводной связи. Механизмы, управляющие уплотнением и распределением среды между пользователями (правила дорожного движения), подробно обсуждаются в главе 3, посвященной управлению доступом к среде.

2.5.1. Уплотнение с пространственным разделением

В случае беспроводных коммуникаций уплотнение может проводиться в четырех измерениях: **в пространстве, во времени, по частоте** и **по коду**. Задача уплотнения — выделить каждому каналу связи пространство, время, частоту и код с минимумом взаимных помех и одновременно с максимальным использованием среды. Термин "канал связи" обозначает здесь лишь сочетание передатчика и приемника, которым необходимо обмениваться данными между собой. Характеристики каналов связи (например, пропускная способность, количество ошибок и т.д.) будут обсуждаться при рассмотрении конкретных технологий.

На рис. 2.16 показано шесть каналов k_i и трехмерная система координат. Оси этой системы соответствуют коду c, времени t и частоте f. Также приводятся (трехмерные) пространственные области s_i, соответствующие окружностям, ограничивающим зоны помех (рис. 2.11). Эта схема описывает первый тип уплотнения — **уплотнение с пространственным разделением** (Space Division Multiplexing, SDM). Как же в этом случае достигается разделение различных каналов? Каналы k_1, k_2 и k_3 отображаются на три "области" s_1, s_2 и s_3, которые очевидным образом разделяют каналы между собой и предотвращают перекрытие их зон помех. Пространство между зонами помех иногда называют **защитным интервалом**. Подобные защитные интервалы необходимы во всех четырех представленных здесь схемах уплотнения.

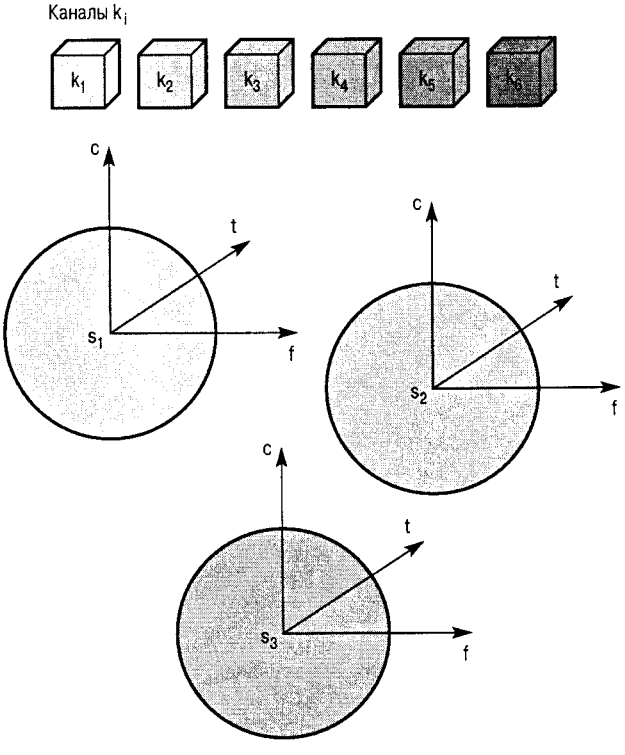

Рис. 2.16. Уплотнение с пространственным разделением (SDM)

Для оставшихся каналов (k_4, k_5, k_6) необходимо еще три дополнительных области. В примере с автомагистралью это означает, что каждый водитель получает свою полосу движения. Понятно, что подобная процедура приводит к неэкономному расходованию пространства. Однако, именно этот принцип использовался в старой аналоговой телефонной системе: каждый абонент был связан с местной телефонной станцией отдельной парой медных проводов. В случае беспроводной передачи схема SDM подразумевает использование отдельных передатчиков для каждого канала связи и достаточно большие расстояния между передатчиками. Данная схема уплотнения используется, например, для радиостанций FM-диапазона, у которых зона передачи ограничена каким-то одним регионом. Поэтому многие радиостанции во всем мире могут работать на одной и той же частоте без взаимных помех. Проблемы при использовании схемы SDM возникают при попытке установить в одной области пространства сразу несколько каналов (например, при вещании нескольких радиостанций в одном городе на одинаковой

частоте). В таких случаях необходимо использовать одну из других схем уплотнения (частотное уплотнение, временное уплотнение или уплотнение с кодовым разделением).

2.5.2. Уплотнение с частотным разделением

В схемах **уплотнения с частотным разделением** (Frequency Division Multiplexing, FDM) ось частот разбивается на несколько неперекрывающихся полос частот (рис. 2.17), после чего каждому каналу k_i выделяется собственная полоса. Передатчики, работающие в определенной полосе частот, могут использовать ее беспрерывно. Опять же, для избежания перекрытия полос частот (иначе называемого **помехами соседних каналов**) необходимы **защитные интервалы**. Подобная схема используется радиостанциями, работающими в одном и том же регионе. При этом каждая станция имеет собственную частоту. В этой очень простой схеме уплотнения не требуется сложной координации передатчиков и приемников — достаточно лишь настроить приемник на частоту нужного передатчика.

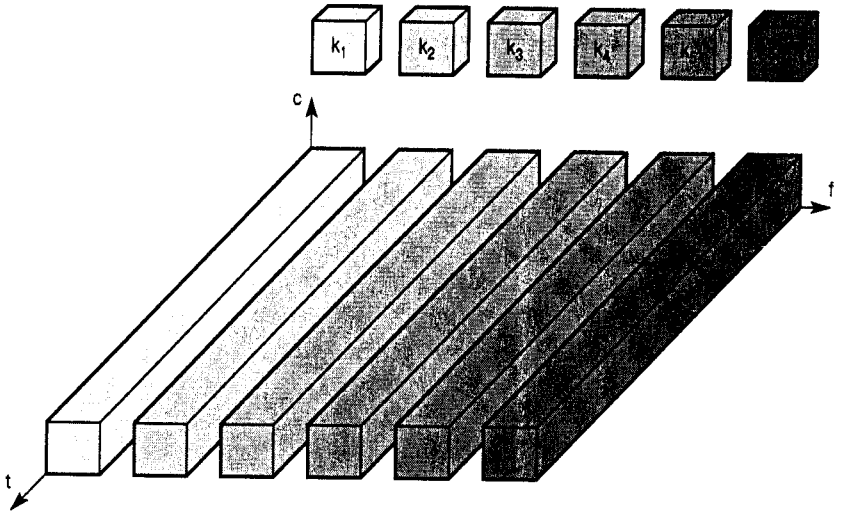

Рис. 2.17. Уплотнение с частотным разделением (FDM)

Однако и эта схема обладает недостатками. Радиостанции работают круглые сутки, а мобильная связь обычно используется лишь несколько минут в день. Поэтому выделение отдельной частоты для каждого из возможных направлений связи является совершенно неоправданной тратой (скудных) частотных ресурсов. Кроме того, назначение постоянных частот передатчикам ограничивает их число и делает схему разделения очень жесткой.

2.5.3. Уплотнение с временным разделением

Более гибкой и подходящей для обычных мобильных коммуникаций схемой является **уплотнение с временным разделением** (Time Division Multiplexing, TDM). Здесь каналу k_i на определенное время предоставляется вся полоса частот, т.е. все передатчики используют одинаковую частоту, но в различные промежутки времени (рис. 2.18). Эти промежутки времени, в течение которых среду используют различные передатчики, опять же разделены **защитными интервалами**. В примере с автомагистралью временные защитные интервалы соответствуют промежуткам между соседними автомобилями. Перекрытие двух передач во времени называется помехами совмещенного канала (для автомагистрали подобные "помехи" означают автомобильную ава-

рию). Чтобы исключить такие помехи, необходима точная синхронизация различных передатчиков. Это, конечно, является недостатком данной схемы, так как в данном случае нужно или обеспечить все передатчики точными часами, или найти способ передавать им сигнал синхронизации. Кроме того, теперь, помимо настройки частоты, приемнику необходимо поймать правильный промежуток времени. Достоинством этой схемы является возможность выделения большего количества времени передатчикам с большей нагрузкой и меньшего количества времени — менее загруженным передатчикам.

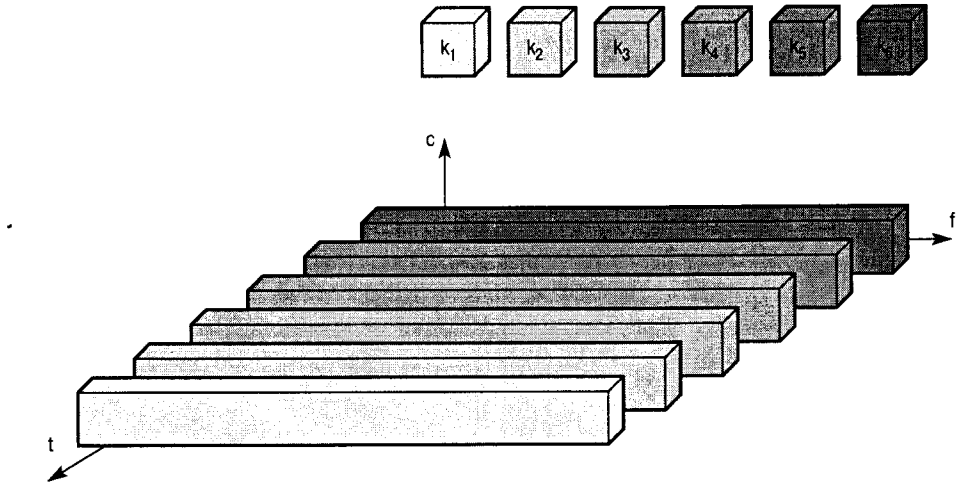

Рис. 2.18. Уплотнение с временным разделением (TDM)

Схемы уплотнения с частотным и временным разделением можно объединить. В этом случае каналу k_i выделяется определенная полоса частот на некоторый промежуток времени (рис. 2.19). Здесь защитные интервалы необходимы и во временном, и в частотном измерениях. Такая схема более устойчива к частотно-избирательным помехам, т.е. к помехам в некоторой узкой полосе частот. Эта полоса может использоваться каналом только на протяжении короткого времени. Кроме того, такая схема обеспечивает некоторую (слабую) защиту от перехвата, так как для прослушивания необходимо знать последовательность частот, используемую передатчиком. В мобильном телефонном стандарте GSM объединение частотного и временного уплотнения применяется для осуществления передач между мобильным телефоном и так называемой базовой станцией (см. раздел 4.1).

Недостатком этой схемы опять же является необходимость согласования работы различных передатчиков. Здесь нужно одновременно управлять и последовательностью частот, и моментами времени, в которые происходит изменение этих частот. Если для двух передатчиков одинаковая частота будет выбрана в одно и то же время, они будут мешать друг другу. Однако, если изменение частоты (иначе называемое перескоком частоты) происходит достаточно быстро, промежутки взаимных помех будут небольшими. При этом, в зависимости от используемой схемы кодирования данных в сигналы, приемник сможет восстановить исходные данные. (Эта технология обсуждается в разделе 2.7.2.)

2.5.4. Уплотнение с кодовым разделением

Схемы SDM и FDM хорошо известны с самого начала использования механизма радиопередачи. Схема TDM тоже имеет множество практических применений. Относительно новой схемой в коммерческих системах связи является **уплотнение с кодовым разделением**

(Code Division Multiplexing, CDM). Благодаря присущей этой схеме высокой безопасности, сначала она использовалась только в военных целях (вместе с технологией расширенного спектра, см. раздел 2.7). Однако, сейчас она начинает играть важную роль и во многих сценариях гражданских беспроводных коммуникаций (подробнее об этом рассказано в разделе 3.5). Как показано на рис. 2.20, все каналы k_i используют для передачи одинаковую частоту в одно и то же время. Разделение здесь достигается с помощью присвоения каждому каналу некоторого "кода". Для обеспечения **защитных интервалов** используются коды, находящиеся в кодовом пространстве на определенном "расстоянии" между собой, например, **ортогональные коды**. Техническая реализация схемы CDM и механизмы доступа к среде обсуждаются в разделе 2.7 и главе 3. Замечательная книга [Viterbi, 1995] охватывает все аспекты уплотнения с кодовым разделением.

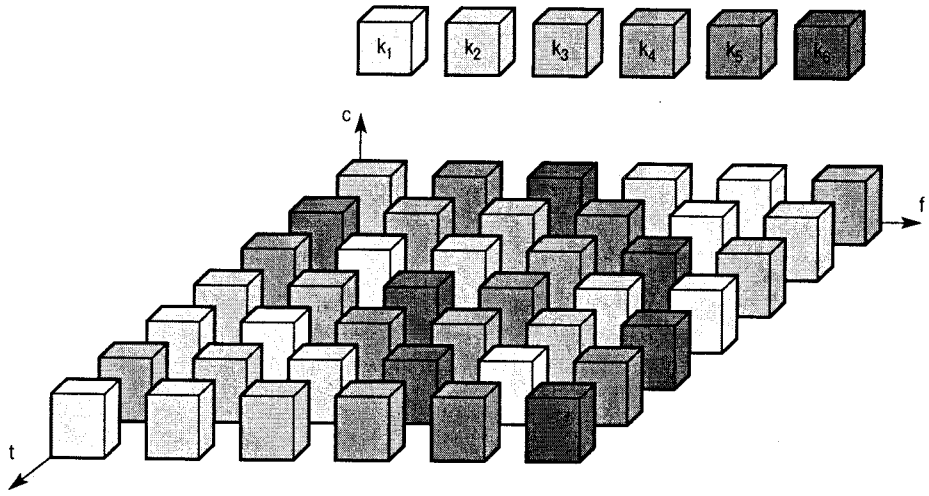

Рис. 2.19. Объединение частотного и временного уплотнения в единую схему

В качестве примера схемы CDM из повседневной жизни можно взять вечеринку, на которой собрались люди из разных стран. Участники вечеринки устанавливают каналы связи, т.е. разговаривают друг с другом, используя одинаковый диапазон частот (в зависимости от голоса, это может быть 300–6000 Гц) в одно и то же время. Когда все разговаривают на одном языке, необходимо пространственное разделение (т.е. разбиение на группы, разговаривающие с "ограниченной мощностью передачи"). Однако, если использовать другой код, т.е. другой язык, человек может "настроиться" на него и отличать разговоры на этом языке от всех остальных разговоров (которые будут проявляться в виде фонового шума). Этот пример также поясняет, почему защищенность передачи является неотъемлемой чертой технологии кодового разделения: сигналы можно получать, но они будут бесполезными, если неизвестен язык. С использованием секретного кода (языка) установление безопасного канала возможно и во "враждебном" окружении. (На обычной вечеринке это может привести к недоразумениям.) В нашем примере также могут потребоваться определенные защитные интервалы. В частности, одновременное использование шведского и норвежского языков может не сработать, так как они слишком близки между собой. Однако, например, шведский и финский языки настолько "ортогональны", что подобные каналы связи легко можно будет разделить.

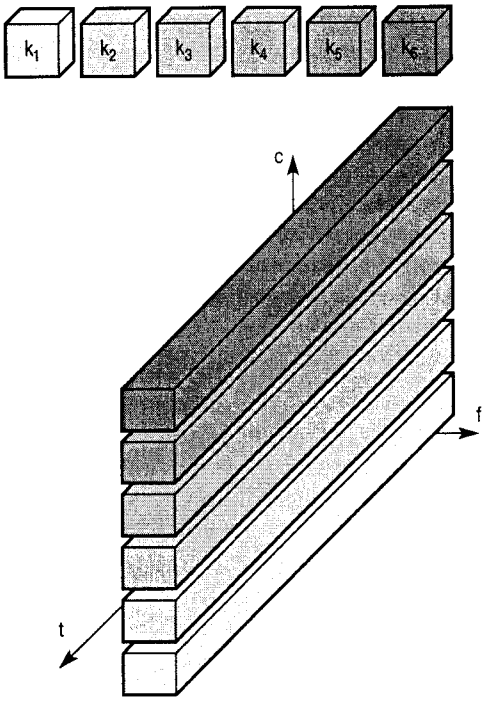

Рис. 2.20. Уплотнение с кодовым разделением (CDM)

Главным достоинством схемы CDM, важным для беспроводной передачи, является хорошая защита от помех и прослушивания. Хотя отдельным каналам все же необходимо присваивать различные коды, тем не менее, кодовое пространство просто огромно по сравнению с частотным пространством. Поэтому присвоение индивидуального кода каждому передатчику обычно не приводит к проблемам. Главный недостаток этой схемы — относительно высокая сложность приемников (см. раздел 3.5). Приемник должен знать код и отделять канал с нужными данными от фонового шума, состоящего из остальных сигналов и шума окружающей среды. Кроме того, для правильного декодирования необходима точная синхронизация передатчика и приемника.

2.6. Модуляция

Три основных схемы модуляции можно проиллюстрировать на примере синусоидальной волны, введенной в разделе 2.2:

$$g(t) = A_t \sin(2\pi f_t t + \varphi_t)$$

Эта функция содержит три параметра — **амплитуду** A_t, **частоту** f_t и **фазу** φ_t. Любой из этих параметров может изменяться в соответствии с поступающими данными или каким-нибудь другим сигналом. Главной темой данного раздела является **цифровая модуляция**, при которой цифровые данные (0 и 1) переводятся в аналоговый сигнал (модулирующий сигнал). Цифровая модуляция необходима в тех случаях, когда требуется передавать данные через среду, допускающую только аналоговую передачу. В качестве примера из области проводных сетей можно привести старую аналоговую телефонную систему — для подключения к этой системе компьютера требуется модем. Модем преобразовывает цифровые данные в аналоговые сигналы и наоборот, аналоговые сигналы в цифровые данные.

Цифровая передача используется, например, в проводных локальных сетях и внутри самих компьютеров [Halsall, 1996]. Однако, в беспроводных сетях непосредственное применение цифровой передачи невозможно. В этом случае поток битов вначале необходимо преобразовать в аналоговый сигнал. Тремя основными способами такого преобразования являются **амплитудная манипуляция (ASK)**, **частотная манипуляция (FSK)** и **фазовая манипуляция (PSK)**. Более подробно эти схемы будут обсуждаться в последующих разделах.

Кроме преобразования цифровых данных в аналоговые сигналы, беспроводная передача нуждается в дополнительной, **аналоговой модуляции**. Она переносит на несущую радиочастоту центральную частоту модулирующего аналогового сигнала, полученного после цифровой модуляции. К примеру, цифровая модуляция преобразовывает поток битов, передаваемых со скоростью 1 Мбит/с, в узкополосный сигнал частотой 1 МГц. В беспроводной системе этот аналоговый сигнал нельзя передавать непосредственно по нескольким причинам.

- **Антенны**. В разделе 2.3 было показано, что для эффективной работы антенны ее размеры должны иметь тот же порядок, что и длина волны сигнала. В примере с модулирующим сигналом частотой 1 МГц высота антенны составляла бы несколько сотен метров. Понятно, что для переносных устройств это не очень практично. В то же время для частоты 1 ГГц можно использовать антенны длиной всего лишь в несколько сантиметров.
- **Уплотнение с частотным разделением**. Схему частотного уплотнения нельзя применить к немодулированной передаче. Для частотного уплотнения (раздел 2.5.2) необходимо с помощью аналоговой модуляции перевести модулирующие сигналы на различные несущие частоты. Чем выше несущая частота, тем шире полоса частот, доступная для модулирующих сигналов.
- **Характеристики среды**. Все эффекты, обсуждавшиеся в разделе 2.4, — затухание, проникновение сквозь препятствия, отражение, рассеяние и дифракция — сильно зависят от длины волны сигнала. Поэтому в каждом случае необходимо выбирать такие несущие частоты, которые бы лучше подходили для данного применения: длинные волны — для подводных лодок, короткие волны — для переносных устройств, очень короткие волны — для направленной микроволновой передачи и т.д.

Для аналоговой модуляции, как и для цифровой, существуют три основные схемы: **амплитудная модуляция (AM)**, **частотная модуляция (FM)** и **фазовая модуляция (PM)**. Более детальную информацию относительно этих схем заинтересованный читатель может найти в книгах [Halsall, 1996] и [Stallings, 1997].

На рис. 2.21 показана (упрощенная) структурная схема радиопередатчика цифровых данных. На первом этапе выполняется цифровая модуляция. Здесь происходит преобразование цифровых данных в аналоговый модулирующий сигнал с использованием одной из схем, обсуждаемых в последующих разделах этой главы. Затем центральная частота полученного аналогового сигнала посредством механизма аналоговой модуляции переносится на несущую радиочастоту. Далее модулированный сигнал передается на антенну.

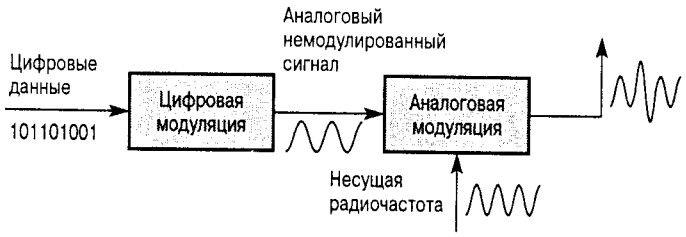

Рис. 2.21. Модуляция в передатчике

Приемник (рис. 2.22) с помощью своей антенны принимает сигнал и превращает (демодулирует) его в аналоговый узкополосный сигнал (для этого ему должна быть известна несущая частота). Для аналогового радиоприемника, настроенного на прослушивание радиостанции, этого было бы вполне достаточно — музыка состоит из аналоговых немодулированных сигналов. Для получения цифровых данных требуется еще один шаг — детектирование битов или кадров, и в этом случае приемник и передатчик необходимо синхронизировать между собой. Способ синхронизации зависит от используемой схемы цифровой модуляции. После выполнения синхронизации принимается решение о том, что же представляет собой данный сигнал — 1 или 0, после чего восстанавливаются исходные данные.

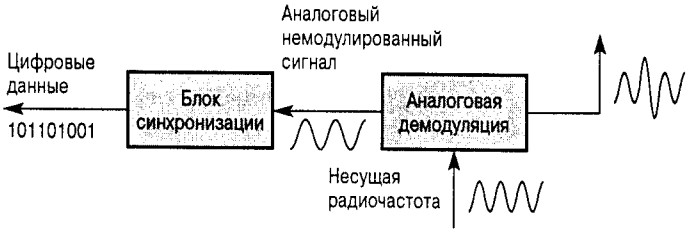

Рис. 2.22. Демодуляция и восстановление данных в приемнике

Схемы цифровой модуляции, представленные в последующих разделах, отличаются во многих аспектах. Среди них можно назвать **спектральную эффективность** (т.е. насколько эффективно используется доступный спектр частот), **энергетическую эффективность** (т.е. сколько энергии необходимо для передачи битов — этот вопрос очень важен для портативных устройств, работающих на батареях) и **устойчивость** к эффектам многолучевого распространения, шумам и помехам [Wesel, 1998].

2.6.1. Амплитудная манипуляция

Амплитудную манипуляцию (Amplitude Shift Keying, ASK), самую простую из существующих схем цифровой модуляции, иллюстрирует рис. 2.23. Два двоичных значения, 0 и 1, представляются различными уровнями амплитуды сигнала. В приведенном примере двоичному нулю соответствует нулевая амплитуда. Для этой простой схемы достаточно узкой полосы частот, но она весьма восприимчива к помехам. Эффекты многолучевого распространения, шум, затухание сильно влияют на амплитуду сигнала. Поэтому в беспроводной среде не удается обеспечить ее постоянство. По этой причине схему ASK обычно не используют для беспроводной радиопередачи. Однако, в системе проводной передачи, обладающей самыми высокими характеристиками, а именно, при оптической передаче, применяется как раз амплитудная манипуляция. В этом случае поступление светового импульса соответствует единице, а отсутствие света — нулю. Несущая частота оптических систем составляет несколько сотен ТГц. Схема ASK также применяется в беспроводной инфракрасной передаче, использующей направленный пучок лучей или рассеянный инфракрасный свет (см. главу 7, "Беспроводные локальные сети").

2.6.2. Частотная манипуляция

Для беспроводной передачи часто используется частотная манипуляция (Frequency Shift Keying, FSK) (см. рис. 2.24). В простейшей из схем FSK, так называемой бинарной FSK, двоичной единице присваивается частота f_1, а двоичному нулю — частота f_2. Бинарную частотную манипуляцию легко реализовать с помощью переключения между двумя генераторами с частотами f_1 и f_2. Во избежание резких изменений в фазе сигнала используются специальные частотные модуляторы с **модуляцией без разрыва фазы** (Continuous Phase Modulation, CPM).

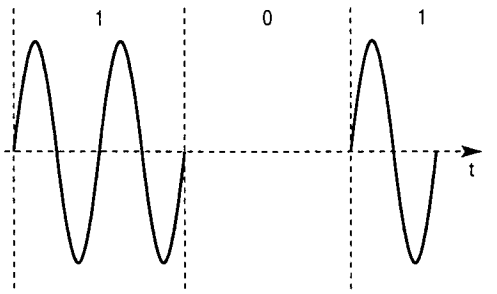

Рис. 2.23. Амплитудная манипуляция (ASK)

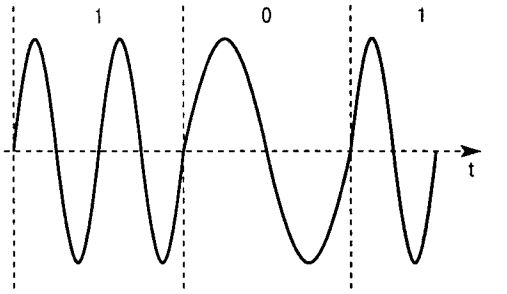

Рис. 2.24. Частотная манипуляция (FSK)

Демодуляцию проще всего осуществить с помощью двух полосовых фильтров, один из которых пропускает частоту f_1, а другой — f_2. Затем в компараторе сравниваются выходные сигналы фильтров и определяется более сильный из них. Для схемы FSK необходима более широкая полоса частот, чем для схемы ASK, но она менее восприимчива к помехам.

2.6.3. Фазовая манипуляция

Наконец, при **фазовой манипуляции** (Phase Shift Keying, PSK) для представления данных используются сдвиги фазы сигнала. Когда в цифровом сигнале за единицей следует нуль, в модулирующем аналоговом сигнале происходит сдвиг фазы на 180° или величину π (см. рис. 2.25). То же самое происходит, когда за нулем следует единица. Такая простая схема, в которой фаза сдвигается на 180° всякий раз, когда изменяются данные, называется **бинарной фазовой манипуляцией** (BPSK). В простейшем модуляторе BPSK частота f умножается на +1, если двоичными данными является единица, и на –1, если двоичные данные представляют собой нуль.

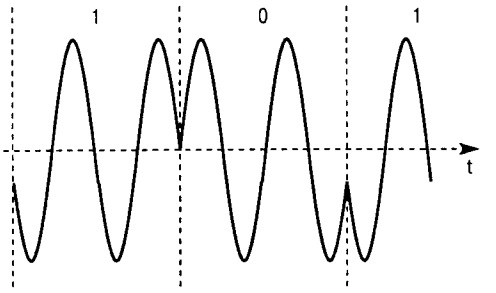

Рис. 2.25. Фазовая манипуляция (PSK)

Для правильного приема передаваемого сигнала передатчик и приемник необходимо синхронизировать по частоте и фазе. Это можно осуществить с помощью **контура фазовой автоподстройки частоты** (ФАПЧ). Фазовая манипуляция, по сравнению с частотной, менее подвержена помехам, но ее передатчики и приемники имеют более сложную конструкцию.

2.6.4. Улучшенная частотная манипуляция

Во многих беспроводных системах используется **манипуляция с минимальным сдвигом** (Minimum Shift Keying, MSK) — одна из наиболее известных схем BFSK. В основу схемы MSK положена двоичная частотная манипуляция без резких изменений фазы, т.е. MSK относится к схемам модуляции без разрыва фазы. На рис. 2.26 приведен пример реализации схемы MSK. На первом этапе информационные биты разделяются на четные и нечетные биты с удвоенной продолжительностью. Эта схема использует две частоты — нижнюю f_1 и верхнюю $f_2 = 2f_1$.

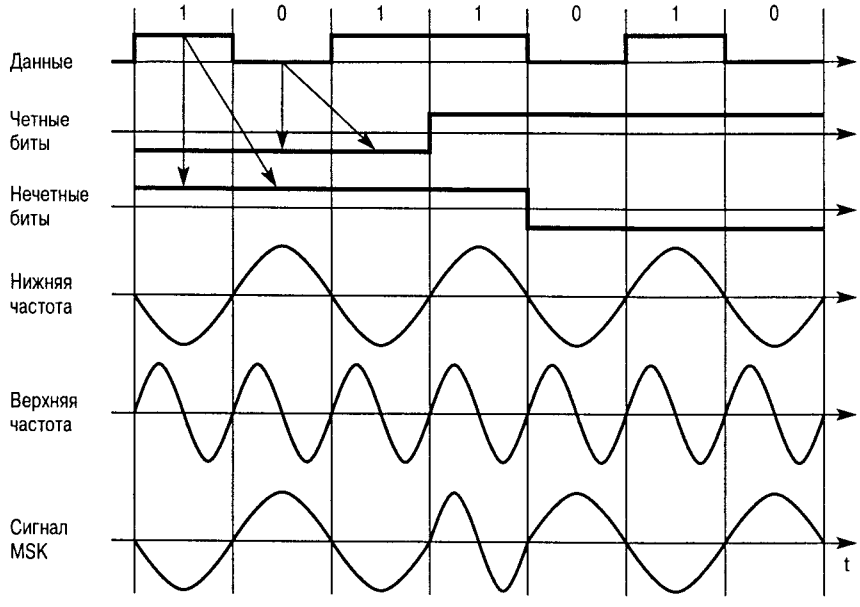

Рис. 2.26. Манипуляция с минимальным сдвигом (MSK)

При генерации сигнала в схеме MSK нижняя или верхняя (или же инвертированная или неинвертированная) частоты выбираются по следующим правилам.

- Если четный и нечетный биты равны нулю, то используется инвертированная верхняя частота f_2 (т.е. частота f_2 используется со сдвигом на 180°).
- Когда четный бит равен единице, а нечетный — нулю, то инвертируется нижняя частота f_1. На рис. 2.26 этот случай иллюстрируют 5-й и 7-й столбцы.
- Если четный бит равен нулю, а нечетный — единице, как для столбцов 1–3, то применяется частота f_1 без изменения фазы.
- Наконец, если оба бита равны единице, выбирается исходная частота f_2.

Итак, когда четный и нечетный биты равны, всегда выбирается верхняя частота. Если нечетный бит равен нулю, сигнал инвертируется. Это позволяет избежать любых сдвигов фазы в конечном сигнале схемы MSK.

При добавлении в схему так называемого гауссова фильтра нижних частот схема MSK превращается в **гауссову манипуляцию с минимальным сдвигом** (GMSK). Эта схема цифровой модуляции применяется во многих европейских беспроводных стандартах, в частности, в стандартах GSM, DECT (см. главу 4) и HIPERLAN (см. раздел 7.4). Фильтр сужает слишком широкий спектр частот, необходимый для работы схемы MSK.

2.6.5. Улучшенная фазовая манипуляция

Существует несколько способов улучшения простой схемы PSK. В основной схеме BPSK используется лишь одно возможное значение сдвига фазы — 180°. На рис. 2.27 слева схема BPSK показана на фазовой диаграмме (обычно такое представление оказывается более удобным, чем представление на временной диаграмме, приведенное на рис. 2.25). Справа на этом же рисунке представлена **квадратурная схема PSK** (QPSK) — одна из самых распространенных схем фазовой манипуляции. Она обеспечивает более высокую скорость передачи

при использовании той же полосы частот за счет кодирования двух битов в одном сдвиге фазы. С другой стороны, в этой схеме можно уменьшить полосу частот и получить такую же скорость передачи, как и для схемы BPSK.

Схему QPSK (как и другие схемы фазовой манипуляции) можно реализовать двумя способами. Сдвиг фазы производится относительно **опорного сигнала** (той же самой частоты). При использовании этой схемы нулевой сдвиг фазы означает, что сигнал находится в фазе с опорным сигналом. Комбинации битов `11` в схеме QPSK отвечает фазовый сдвиг 45°, комбинациям `10`, `00` и `01` — 135°, 225° и 315° соответственно. Все сдвиги фазы рассматриваются относительно опорного сигнала. Части сигнала "выбираются" и объединяются передатчиком, как показано на рис. 2.28. Для восстановления данных при приеме нужно сравнить поступающий сигнал с опорным. Одним из недостатков этой схемы является необходимость формирования опорного сигнала в приемнике. Передатчик и приемник нужно часто синхронизировать, применяя, например, специальные синхронизирующие комбинации перед приходом данных пользователя или используя в качестве опорной частоту пилотного сигнала.

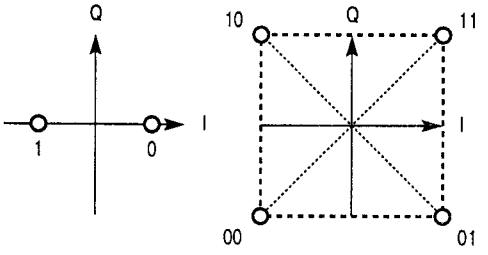

Рис. 2.27. Схемы BPSK и QPSK, представленные на фазовой диаграмме

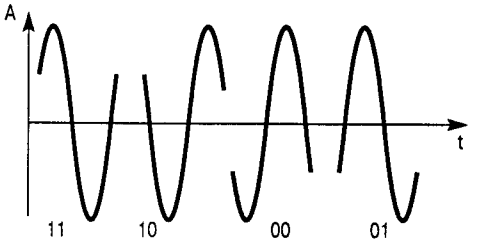

Рис. 2.28. Схема QPSK во временной области

Этих проблем можно избежать с помощью **дифференциальной схемы QPSK** (DQPSK). Здесь сдвиг происходит относительно фазы двух предыдущих битов, а не относительно опорного сигнала. В этом случае в приемнике для восстановления данных сравниваются два сигнала данных, а опорный сигнал не нужен. Схема DQPSK применяется в американских беспроводных технологиях IS-136 и PACS, а также в японской системе PHS.

Для улучшения предложенных схем можно использовать и много других углов сдвига фазы. Например, при использовании восьми фазовых углов в одном фазовом сдвиге можно закодировать три бита. Кроме того, можно объединить схемы фазовой и амплитудной манипуляции. Этот подход применяется, в частности, в схеме квадратурной амплитудной модуляции (QAM), используемой в стандартных модемах на 9 600 бит/с (рис. 2.29). Здесь объединены три различных амплитуды и 12 углов. При этом в одном изменении фазы/амплитуды кодируется 4 бита. Проблемы беспроводной связи обусловлены шумами и межсимвольной интерференцией. Чем больше используется "точек" фазовой области, тем труднее их разделить. Схема DQPSK оказалась одной из наиболее эффективных в этом отношении [Wesel, 1998].

2.6.6. Модуляция с несколькими несущими

Несколько в стороне от других находится схема **модуляции с несколькими несущими** (MultiCarrier Modulation, MCM). При рассмотрении европейской цифровой радиосистемы DAB (см. раздел 6.3) эту схему модуляции также называют **ортогональным уплотнением с частотным разделением** (OFDM) или **кодированным частотным уплотнением** (COFDM). Главным достоинством схемы MCM является высокая устойчивость к межсимвольной интерференции. Как было указано в разделе 2.4.3, межсимвольной интерференции более подвержены системы с высокой скоростью пе-

редачи данных. При модуляции с несколькими несущими высокоскоростной поток данных разбивается на несколько потоков с меньшей скоростью, каждый из них использует независимую несущую частоту. Пусть, к примеру, необходимо передать n символов в секунду. Если есть c поднесущих частот, то на каждой из них будет передаваться n/c символов в секунду. Кроме того, при использовании схемы QPSK каждый символ будет представлять два бита. В системе DAB, например, используется от 192 до 1 536 поднесущих частот.

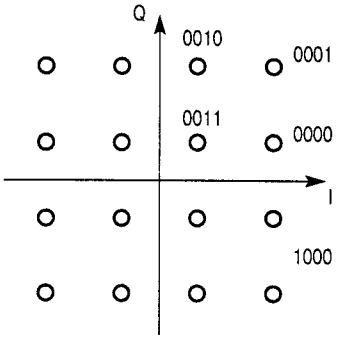

Рис. 2.29. Квадратурная амплитудная модуляция

Еще одним достоинством схемы MCM является то, что при ее использовании частотноизбирательное затухание влияет только на некоторые поднесущие частоты, а не на весь сигнал. В большинстве случаев при применении модуляции с несколькими несущими передача данных происходит с защитными интервалами между символами или группами символов. Это позволяет приемнику бороться с нежелательными эффектами многолучевого распространения. Более детальную информацию о схеме MCM заинтересованный читатель найдет в книге [Wesel, 1998], стандарте [ETSI, 1997] и в разделе 6.3.

2.7. Расширение спектра

Как подразумевает само название, в технологии **расширенного спектра** применяется расширение полосы частот, используемой для передачи данных. Расширение полосы частот имеет несколько достоинств. Главным преимуществом этой технологии является устойчивость к **узкополосным помехам**. На диаграмме *а* (рис. 2.30) показан идеализированный узкополосный сигнал исходных пользовательских данных (график изменения мощности передачи P в зависимости от частоты f). На диаграмме *б* передатчик расширяет спектр передаваемого сигнала, т.е. превращает узкополосный сигнал в широкополосный. Энергия, необходимая для передачи сигнала (площадь диаграммы), остается постоянной, но распространяется на больший диапазон частот. Поэтому уровень мощности передаваемого сигнала может быть значительно ниже уровня мощности исходного узкополосного сигнала. Данный уровень мощности бывает трудно отличить даже от фонового шума. Все это затрудняет отделение полезного сигнала от шума и, следовательно, осложняет его детектирование.

В процессе передачи сигнала на него накладываются узкополосные и широкополосные помехи, как показано на диаграмме *в*. В приемнике осуществляется сужение спектра, т.е. преобразование расширенного полезного сигнала обратно в узкополосный сигнал. При этом спектр узкополосных помех расширяется, а спектр широкополосных помех остается без изменений. Затем преобразованный сигнал пропускается через полосовой фильтр, обрезающий частоты слева и справа от исходного спектра узкополосного сигнала (диаграмма *д*). Теперь приемник легко может восстановить исходные данные, так как уровень мощности полезного сигнала дос-

таточно высок — иначе говоря, полезный сигнал значительно мощнее оставшихся помех. В последующих разделах будет показано, как осуществляется подобное расширение спектра.

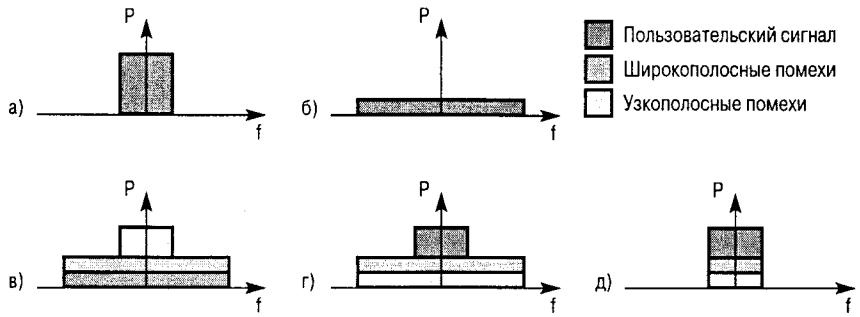

Рис. 2.30. Технология расширенного спектра: расширение и сужение спектра сигнала

Таким образом, применение технологии расширенного спектра позволяет ослабить узкополосные помехи в отдельном канале. Не менее успешно эту технологию можно использовать и для нескольких каналов. Рассмотрим ситуацию, изображенную на рис. 2.31. Шесть различных каналов используют частотное уплотнение. Это значит, что каждому каналу для передачи выделяется отдельная узкая полоса частот. Во избежание появления помех соседних каналов выделенные полосы частот разделены защитными интервалами. Как уже упоминалось в разделе 2.5.2, при использовании этого метода необходимо очень экономно распределять частоты. Дополнительно на рис. 2.31 показано качество передачи в каждом из каналов. Качество передачи в канале зависит от его частоты и является общей мерой помех, присутствующих на данной частоте. Качество канала зависит также от времени — на диаграмме показан лишь моментальный "снимок". В зависимости от характеристик приемника возможен случай, когда для устойчивого приема подходят лишь 1, 2, 5 и 6 каналы, а качество передачи в 3 и 4 каналах настолько низкое, что восстановление передаваемых данных оказывается невозможным. Передача в 3 и 4 каналах нарушается из-за наличия узкополосных помех. Однако, подобная ситуация наблюдается только в определенный момент времени, а в следующий момент ситуация может оказаться совершенно иной. В общем случае использование подобных узкополосных сигналов часто приводит к серьезным затруднениям в установлении связи.

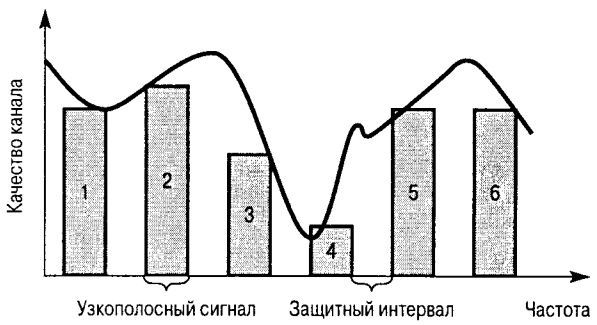

Рис. 2.31. Узкополосные помехи без использования технологии расширенного спектра

Как в этой ситуации может помочь применение технологии расширенного спектра? Выше было показано, что расширение спектра сигнала увеличивает его устойчивость к узкополосным помехам. Данную технологию можно применить одновременно ко всем узкополосным сигна-

лам. На рис. 2.32 показано, как все узкополосные сигналы подвергаются расширению с превращением в широкополосные сигналы, использующие один и тот же диапазон частот. Теперь уже не потребуется распределять частоты (при наших упрощенных допущениях), так как все передатчики используют одну и ту же полосу частот. Но как же приемникам восстановить нужный сигнал?

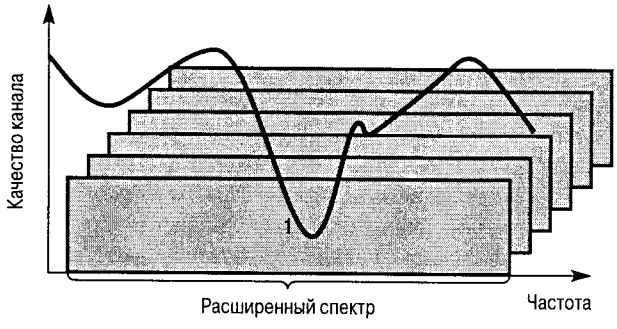

Рис. 2.32. Ослабление узкополосных помех с помощью расширения спектра сигнала

Для организации разделения каналов в этом случае вместо частотного уплотнения используется уплотнение с кодовым разделением. Подобный подход демонстрирует тесную связь между технологией расширения спектра и схемой CDM. В разделах 2.7.1 и 2.7.2 будет показано, как расширение узкополосного сигнала достигается за счет применения специального **кода**. Каждому каналу выделяется собственный код, который необходимо использовать при приеме для восстановления данных. Без знания кода сигнал восстановить нельзя, он ведет себя как фоновый шум. Использование секретного кода для расширения обеспечивает высокую защищенность передачи по технологии расширенного спектра. Сосуществование нескольких сигналов без координации (если не учитывать то, что коды должны обладать определенными свойствами), устойчивость к узкополосным помехам, относительно высокая безопасность и характеристики, подобные фоновому шуму, делают комбинацию расширенного спектра и схемы CDM весьма привлекательными для военного использования.

В настоящее время объединение технологии расширения спектра со схемой CDM становится все более целесообразным для использования не только в военных, но и в гражданских приложениях. Как уже упоминалось выше, мировые ресурсы частот весьма ограничены. Применение технологии расширенного спектра позволяет вводить новые технологии передачи на тех же полосах частот, где уже работают современные узкополосные системы. Этот подход используется в американских мобильных телефонных системах. Полоса частот около 850 МГц была уже занята системами, работающими по схемам TDM и FDM (системы AMPS и IS-54). Тем не менее, в том же диапазоне оказалось возможным внедрение новой системы с кодовым разделением (IS-95).

Технология расширенного спектра имеет и определенные недостатки. Одним из них является повышенная сложность приемников, которые должны выполнять сужение спектра и расшифровку сигнала. Тем не менее, в настоящее время, благодаря распространению цифровой обработки сигналов, сужение спектра можно осуществить и при высокоскоростной передаче данных. Другая проблема возникает из-за того, что в технологии расширения спектра сигнала используется более широкая полоса частот. Несмотря на то, что расширенные сигналы ведут себя почти как обычный шум, они все же превышают уровень фонового шума и способны влиять на другие передачи, если не принять специальных мер предосторожности.

Расширение спектра можно осуществить двумя различными способами, которые рассматриваются в двух последующих разделах.

2.7.1. Расширение спектра методом прямой последовательности

В системах **расширения спектра методом прямой последовательности** (Direct Sequence Spread Spectrum, DSSS) к потоку передаваемых данных применяется логическая операция исключающего ИЛИ (XOR) с сигналом так называемой **чиповой последовательности**. В примере, приведенном на рис. 2.33, в результате этой операции получается или последовательность (0110101), или двоичное дополнение к ней. В то время как каждый бит полезного сигнала имеет продолжительность t_b, чиповая последовательность состоит из более коротких импульсов, называемых **чипами**, с продолжительностью t_c. При корректной генерации чиповая последовательность должна выглядеть как случайный шум, поэтому иногда ее также называют **псевдошумовой** последовательностью. **Коэффициент расширения** $s = t_b/t_c$ определяет полосу частот, занимаемую результирующим сигналом. Если исходному сигналу необходима полоса шириной w, то после расширения ширина полосы частот результирующего сигнала будет равна $s * w$. Хотя в нашем простом примере коэффициент расширения равен всего лишь 7 (а чиповая последовательность 0110101 не совсем случайна), в гражданских приложениях коэффициенты расширения обычно составляют от 10 до 100, а для военных приложений эта цифра может достигать даже 10 000. В беспроводных ЛВС, совместимых со стандартом IEEE 802.11 и использующих технологию DSSS, применяется последовательность 10110111000 — так называемый код Баркера.

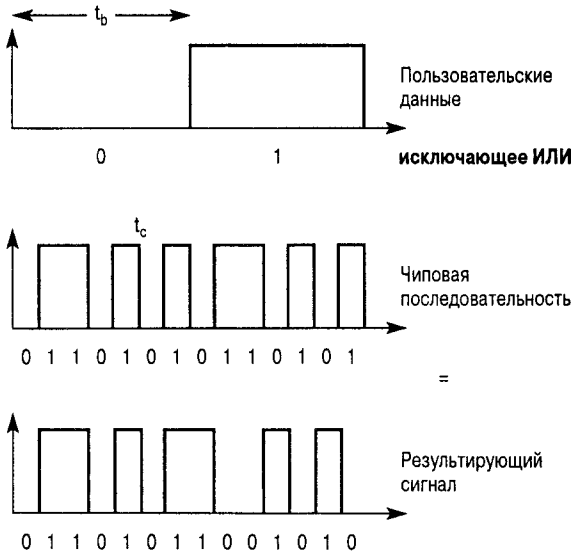

Рис. 2.33. Расширение спектра по технологии DSSS

До сих пор обсуждалось только расширение сигнала. Однако передатчикам и приемникам, использующим технологию DSSS, необходимы и некоторые дополнительные компоненты, показанные на рис. 2.34 и 2.35. На первом этапе в передатчике DSSS (рис. 2.34) происходит расширение передаваемых данных с помощью некоторой чиповой последовательности (**цифровая модуляция**). Затем, как объяснялось в разделе 2.6, расширенный сигнал модулирует несущую радиочастоту (**радиомодуляция**). Предположим, например, что полезный сигнал с частотой 1 МГц после расширения с помощью кода Баркера будет иметь частоту 11 МГц. Далее этот сигнал накладывается на несущую радиочастоту (например, 2,4 ГГц для полосы ISM) и передается.

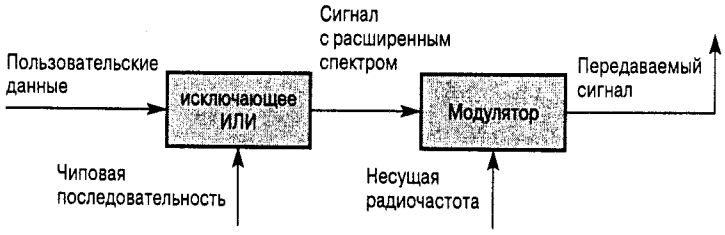

Рис. 2.34. Структурная схема передатчика, использующего технологию DSSS

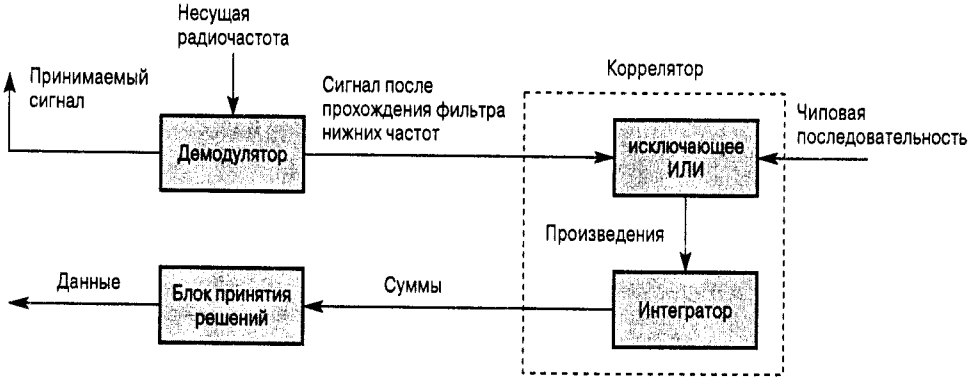

Рис. 2.35. Структурная схема приемника, использующего технологию DSSS

Приемник, использующий технологию DSSS, устроен сложнее, чем передатчик. Приемник должен выполнять функции, обратные двум этапам модуляции, выполняемым в передатчике. Однако, из-за шумов и многолучевого распространения для восстановления данных необходимы некоторые дополнительные механизмы. На первом этапе в приемнике осуществляется демодуляция принятого сигнала. Для этого используется такая же несущая частота, как у передатчика, подавляющая радиочастотную составляющую входного сигнала. В результате получается сигнал примерно с такой же полосой частот, как и исходный полезный сигнал с расширенным спектром. Для выделения этого сигнала иногда применяется дополнительная фильтрация.

Подобная схема демодуляции широко используется в обычных радиоприемниках. Однако, следующий этап обработки сигнала составляет серьезную проблему для приемников, работающих по технологии DSSS, и сильно усложняет их конструкцию. Во-первых, приемнику должна быть известна исходная чиповая последовательность, т.е. приемник должен генерировать точно такую же псевдослучайную чиповую последовательность, как и передатчик. Во-вторых, последовательности передатчика и приемника должны быть точно синхронизированы, так как в приемнике происходит логическое умножение импульсов чипового и поступающего из антенны сигналов. Это предусматривает выполнение еще одной операции исключающего ИЛИ (более подробные объяснения можно найти в разделе 3.5) и наличие механизма доступа к среде, построенного по этой схеме. На протяжении битового сигнала, длительность которого также определяется посредством синхронизации, **интегратор** складывает все вычисленные произведения. Умножение чиповых импульсов и поступивших сигналов и сложение произведений в интеграторе называется корреляцией, а соответствующее устройство называется **коррелятором**. Наконец, **блок принятия решений** обрабатывает сформированные интегратором суммы и определяет, что представляет каждая из этих сумм — двоичную единицу или двоичный нуль.

Если передатчик и приемник полностью синхронизированы, а сигнал не слишком искажен шумами и многолучевым распространением, то технология DSSS отлично работает, следуя описанной выше упрощенной схеме. При передаче полезного сигнала, равного 01, и применении кода Баркера формируется расширенный сигнал 10110111000010010000111. В приемнике, после демодуляции несущей, к этому сигналу побитово применяется операция исключающего ИЛИ, использующая в качестве чиповой последовательности тот же код Баркера. В результате получаются две суммы логических произведений: 0 для первого бита и 1 — для второго. В блоке принятия решений первой сумме ставится в соответствие двоичный нуль, второй сумме — двоичная единица, и, таким образом, исходный полезный сигнал восстанавливается.

В действительности, однако, ситуация часто оказывается существенно сложнее. Предположим, что демодулированный сигнал слегка искажен и имеет вид 10100101000011010000111. В результате для первого бита сумма произведений равна 2, для второго — 10. Тем не менее, блок принятия решений сможет выполнить корректное преобразование, если будет ставить в соответствие всем суммам, меньшим 4, двоичный нуль, и всем суммам, большим 7, — двоичную единицу. Однако, в данной ситуации очень важно поддерживать синхронизацию с передатчиком сигнала. Но что же происходит при наличии многолучевого распространения? Между передатчиком и приемником существуют несколько путей с различными задержками, причем разные пути могут характеризоваться различными коэффициентами затухания. В этом случае возможным решением являются так называемые **приемники Рейка** (Rake receiver). В таком приемнике для n самых главных путей распространения используется n корреляторов. Каждый коррелятор синхронизирован с передатчиком и учитывает задержку на отдельном пути. Как только приемник обнаруживает новый, более "мощный" путь, он присваивает его корреляторy с самым "слабым" путем распространения. Затем выходные сигналы корреляторов объединяются и направляются в блок принятия решений. Таким образом, приемники Рейка используют преимущества многолучевого распространения, конструктивно объединяя различные пути [Viterbi, 1995].

2.7.2. Расширенный спектр с перескоком частоты

В системах **расширения спектра методом перескока частоты** (Frequency Hopping Spread Spectrum, FHSS) доступная полоса частот разбивается на много каналов с меньшими полосами, разделенными защитными интервалами. Передатчик и приемник используют один из этих каналов на протяжении некоторого времени, а затем перескакивают на другой канал. Таким образом, в этой системе применяется комбинация схем FDM и TDM. Схема использования отдельных каналов называется **последовательностью перескоков**. Время, на протяжении которого используется канал с определенной частотой, называется **временем пребывания**. Системы FHSS бывают двух типов — с медленным и быстрым перескоком (см. рис. 2.36).

При **медленном перескоке** передатчик использует одну и ту же частоту для передачи нескольких битов. На рис. 2.36 показано пять битов полезных данных, каждый продолжительностью t_b. Работая по схеме с медленным перескоком, передатчик использует частоту f_2 на протяжении времени пребывания t_d, необходимого для передачи первых трех битов. Затем передатчик перескакивает на следующую частоту f_3. Системы с медленным перескоком обычно дешевле и менее требовательны, чем системы с быстрым перескоком, но они более подвержены узкополосным помехам. Медленный перескок частоты может применяться, в частности, в системе GSM (см. раздел 4.1).

В системах с **быстрым перескоком** передатчик несколько раз изменяет частоту на протяжении передачи одного бита. В приведенном выше примере перескок происходит три раза на протяжении одного бита. Системы с быстрым перескоком реализовать сложнее, так как передатчик и приемник должны быть синхронизированы более точно. Это необходимо для осуще-

ствления перескока в примерно один и тот же момент времени. Однако, такие системы значительно лучше преодолевают узкополосные помехи и частотноизбирательное затухание, поскольку привязаны к определенной частоте на протяжении очень короткого времени. Примером системы FHSS является технология Bluetooth, рассматриваемая в разделе 7.5. Она предусматривает выполнение 1 600 перескоков в секунду. В системе Bluetooth для перескоков используется 79 несущих частот, равномерно распределенных в интервале 1 МГц диапазона ISM около 2,4 ГГц.

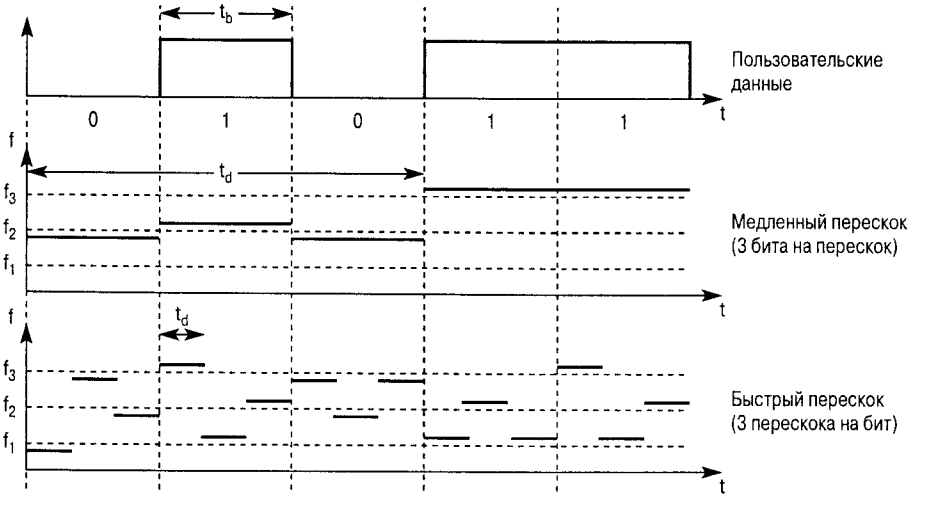

Рис. 2.36. Системы FHSS с медленным и быстрым перескоком частоты

На рис. 2.37 и 2.38 приведены упрощенные структурные схемы передатчиков и приемников, работающих по технологии FHSS. На первом этапе передатчик схемы FHSS осуществляет модуляцию передаваемых данных согласно одной из схем цифро-аналоговой модуляции (см. раздел 2.6). В качестве примера можно указать схемы FSK или BPSK. Пусть по схеме FSK для двоичного нуля используется частота f_0, а для двоичной единицы — частота f_1. В результате формируется узкополосный модулирующий сигнал. На следующем этапе выполняется перескок частоты в соответствии с установленной последовательностью перескоков. Сведения о последовательности перескоков направляются в синтезатор частот, генерирующий несущие частоты f_i. При второй модуляции модулирующий узкополосный сигнал накладывается на несущую частоту и формируется новый расширенный сигнал. При этом нулю соответствует частота $f_i + f_0$, а единице — частота $f_i + f_1$. Если различные передатчики схемы FHSS используют только неперекрывающиеся последовательности перескоков, т.е. никогда не работают на одинаковой частоте f_i в одно и то же время, то их передачи не мешают друг другу. Такой подход требует координации работы всех передатчиков и согласования их последовательностей перескоков. Псевдослучайные последовательности перескоков можно использовать и без координации, как в системах DSSS, причем вполне достаточно, чтобы эти последовательности отвечали определенным требованиям, сводящим помехи к минимуму. При перескоке несколько передатчиков могут выбрать одинаковую частоту, но для систем с быстрым перескоком время пребывания на этой частоте будет мало, поэтому и взаимные помехи окажутся минимальными.

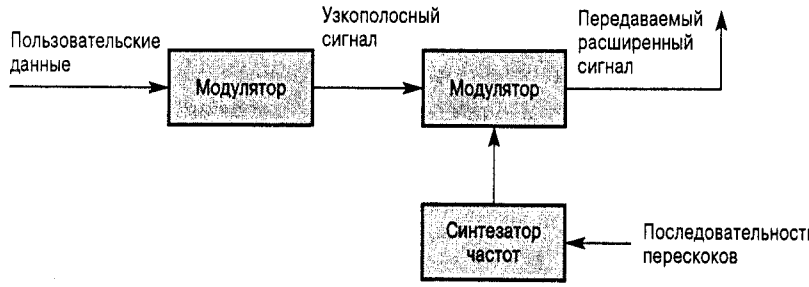

Рис. 2.37. Схема устройства передатчика, использующего технологию FHSS

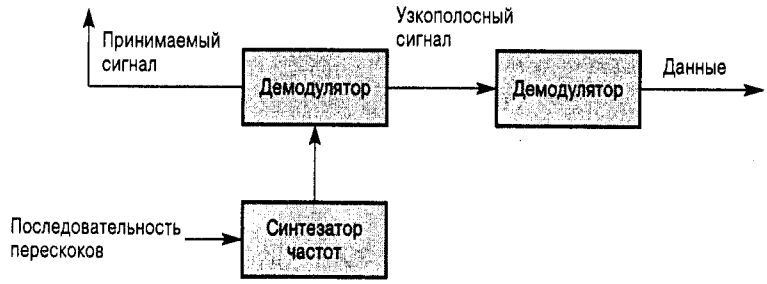

Рис. 2.38. Схема устройства приемника, работающего по технологии FHSS

Приемник системы FHSS должен знать используемую последовательность перескоков. Кроме того, он должен постоянно поддерживать синхронизацию с передатчиком. При приеме выполняются действия, обратные процедуре модуляции, в результате чего передаваемые данные восстанавливаются. Кроме того, необходимо использовать несколько фильтров (на упрощенной диаграмме на рис. 2.38 они не показаны).

В системах FHSS расширение достигается проще, чем в системах DSSS. В то время как системы DSSS всегда используют всю доступную полосу частот, в системах FHSS используется только ее часть. С другой стороны, системы DSSS более устойчивы к затуханию и эффектам многолучевого распространения. Кроме того, сигналы DSSS гораздо труднее детектировать — если код расширения неизвестен, детектирование практически невозможно. Когда каждый передатчик использует для расширения сигнала (DSSS или FHSS) свою собственную псевдослучайную числовую последовательность, в системе реализуется уплотнение с кодовым разделением. Более подробную информацию относительно применения технологии расширенного спектра и ее теоретического обоснования можно найти в книгах [Viterbi, 1995], [Peterson, 1995] и [Dixon, 1994].

2.8. Сотовые системы

В сотовых системах мобильной связи для уплотнения применяется схема пространственного разделения (SDM). Каждый передатчик, обычно называемый **базовой станцией**, покрывает некоторую площадь — **ячейку**. Размеры ячейки могут составлять от нескольких десятков метров в зданиях и сотен метров в городах до десятков километров в сельской местности. Ячейки никогда не имеют форму идеальных кругов или шестиугольников (как показано на рис. 2.39). В действительности их форма зависит от реального окружения передатчика (зданий, гор, долин и т. п.), погодных условий, а иногда даже и от загрузки системы. Типичными системами, использующими этот подход, являются системы мобильных телекоммуникаций (см. главу 4). В

этом случае мобильная станция, находящаяся внутри ячейки, окружающей некоторую базовую станцию, всегда устанавливает связь с этой базовой станцией и наоборот.

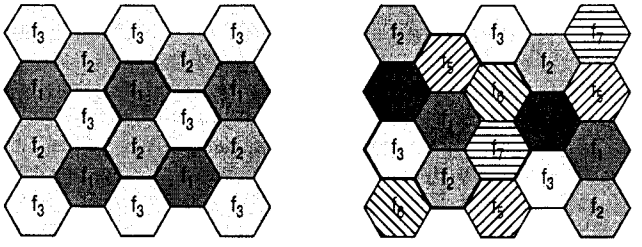

Рис. 2.39. Сотовые системы с тремя и семью сотовыми кластерами

В обсуждаемом контексте возникает вопрос: почему поставщики мобильных сетей устанавливают несколько тысяч базовых станций в каждой стране (что очень дорого) вместо того, чтобы использовать мощные передатчики с огромными ячейками (как действуют, например, радиостанции).

Причина в том, что сотовые системы с маленькими ячейками обладают определенными **преимуществами**.

- **Более высокая пропускная способность**. Применение пространственного уплотнения позволяет повторно использовать частоты. Если два передатчика находятся далеко друг от друга, т.е. их зоны помех не перекрываются, они могут работать на одинаковой частоте. В большинстве систем мобильной телефонной связи частоты (или комбинации перескоков) присваиваются отдельным пользователям, соответственно для других пользователей они недоступны. Так как ресурсы частот небольшие, число пользователей в одной ячейке оказывается ограниченным. Поэтому большие ячейки не позволяют увеличивать количество пользователей — напротив, число пользователей на единицу площади в них получается меньше. Это является одной из причин использования очень маленьких ячеек в городах, где много людей пользуются мобильными телефонами.
- **Меньшая мощность передачи**. Хотя для базовых станций вопросы мощности не представляют большой проблемы, для мобильных станций (телефонов) они являются крайне важными. Приемнику, находящемуся далеко от базовой станции, необходима мощность передачи, намного превышающая несколько ватт. Однако, соответствующее энергообеспечение является серьезной проблемой для мобильных переносных устройств.
- **Только местные помехи**. Большие расстояния между приемником и передатчиком приводят к появлению дополнительных помех. При использовании маленьких ячеек мобильные и базовые станции имеют дело только с локальными помехами.
- **Устойчивость**. Сотовые системы децентрализованы и поэтому более устойчивы к повреждениям отдельных составляющих. Поломка одной из антенн влияет на связь лишь на небольшом участке сети.

Маленькие ячейки имеют и свои **недостатки**.

- **Требуется развитая инфраструктура**. Для сотовых систем необходима сложная инфраструктура, связывающая все базовые станции в единое целое. Эта инфраструктура должна включать множество антенн, коммутаторов для перенаправления звонков, локальных регистраторов для определения местонахождения мобильной станции и т.п. Все это делает систему в целом очень дорогой.

- **Переключение базовых станций.** При смене ячейки мобильная станция должна переключаться с одной базовой станции на другую. В зависимости от размеров ячеек и скорости движения такое переключение может происходить достаточно часто.
- **Распределение частот.** Для исключения взаимных помех передатчиков, использующих одну и ту же частоту, необходимо организовать эффективное распределение частот. С одной стороны, требуется избавиться от помех, а с другой — учитывать, что системе доступно лишь ограниченное количество частот.

Во избежание взаимного влияния различные передатчики, находящиеся в перекрывающихся диапазонах помех, используют схему частотного разделения (FDM). При объединении схем FDM и TDM (см. рис. 2.19) необходимо координировать используемые отдельными передатчиками комбинации перескоков. Общая цель — никогда не использовать одинаковые частоты в перекрывающихся диапазонах помех в одно и то же время. На рис. 2.39 показаны две возможные модели конфигурации ячеек, в которых помехи сведены к минимуму. Здесь ячейки объединены в **кластеры**. На схеме слева кластер состоит из трех ячеек, а на схеме справа — из семи. На схеме слева одна из ячеек кластера использует набор частот f_1, другая — набор частот f_2, а третья — набор частот f_3. В реальной жизни подобная конструкция выглядит несколько иначе. Конфигурация из правильных шестиугольников выбрана только для того, чтобы проиллюстрировать модель самым простым способом. При внимательном рассмотрении схемы этой конфигурации можно заметить, что одинаковые наборы частот периодически повторяются. Поэтому для того, чтобы исключить взаимные помехи между ячейками, использующими одинаковые наборы частот, необходимо ограничивать мощность передатчиков.

Для еще большего ослабления помех (при некоторых условиях трафика, т.е. при определенном количестве пользователей на единице площади) применяются **секторные антенны**. На рис. 2.40 показано использование трех секторов в одной ячейке для кластера с тремя ячейками. Обычно применение секторных антенн вместо ненаправленных имеет смысл при относительно больших размерах ячеек.

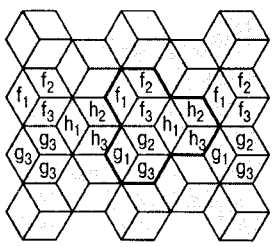

Рис. 2.40. Сотовая система с тремя сотовыми кластерами и тремя секторами в одной ячейке

Фиксированное распределение частот между сотовыми кластерами и ячейками не очень эффективно, если загрузка системы может динамически изменяться. К примеру, если одна из ячеек загружена сильно, а другая — слабо, имеет смысл организовать "заимствование" частот. Ячейкам с более интенсивным информационным обменом динамически выделяется большее количество рабочих частот. Подобная схема называется **распределением каналов с заимствованием** (Borrowing Channel Allocation, BCA), а первая схема носит название **стационарного распределения каналов** (Fixed Channel Allocation, FCA). Схема FCA из-за простоты использования применяется, в частности, в системе GSM. В то же время перед установкой такой системы требуется тщательно проанализировать потенциальную интенсивность информационного обмена.

Схема **динамического распределения каналов** (Dynamic Channel Allocation, DCA) используется в стандарте DECT (см. раздел 4.2). В этой схеме также возможно заимствование частот. Кроме того, можно свободно присваивать частоты различным ячейкам. При таком динамическом присваивании существует опасность взаимных помех между ячейками, использующими одинаковую частоту. Поэтому "заимствованная" частота может блокироваться окружающими ячейками.

Сотовым системам, использующим уплотнение с кодовым разделением вместо частотного уплотнения, не нужны такие замысловатые схемы распределения каналов. Здесь пользователи разделяются не частотами, а кодами. Однако, при планировании ячеек появляются другие

проблемы — размер ячейки зависит от текущей загрузки. В таких случаях говорят, что ячейки схемы CDM "дышат". При слабой загрузке ячейка покрывает большую площадь, которая уменьшается при увеличении загрузки. Причиной этого является возрастание уровня шумов при увеличении количества пользователей в ячейке. (Вспомним примеры из раздела 2.5: если неизвестен код, то каждый сигнал проявляется в виде шума, т.е. все больше людей участвуют в общем хоре.) Чем выше уровень шума, тем больше затухание сигнала и тем больше ошибок будет иметь место при передаче. В конце концов, мобильные станции, находящиеся дальше остальных от базовой станции, просто выпадут из ячейки. (Это похоже на попытку заговорить в переполненном помещении с человеком, находящимся достаточно далеко.)

2.9. Резюме

В этом разделе обсуждались основы беспроводной связи. Из рассмотрения было изъято большинство формул, которые можно найти в других книгах, посвященных беспроводной передаче и эффектам распространения радиоволн. Однако, мы надеемся, что приведенные примеры, обсуждение технологий и проблем, помогли читателю понять, в чем состоят фундаментальные различия между проводной и беспроводной связью и почему при отсутствии проводов протоколы и приложения на высших уровнях должны следовать совсем другим принципам.

Регулирование и гармонизация использования частот являются вопросами всемирного масштаба. В данной главе обсуждаются многие системы, использующие и одинаковые, и различные частоты. В этом отношении хотелось бы, чтобы в будущем появились частоты, доступные во всем мире, что позволит избежать применения дорогих многорежимных устройств.

Поскольку в основу беспроводной связи положены электромагнитные волны, для передачи и приема этих волн необходимы антенны. В то время как в базовых станциях мобильных телефонных систем часто используются направленные антенны, у самих мобильных устройств они ненаправленные. В данной главе рассмотрены стандартные эффекты экранирования, затухания, отражения, дифракции и рассеяния электромагнитных волн. Все эти эффекты имеют следствием самую главную проблему беспроводной связи — многолучевое распространение сигналов. При многолучевом распространении пропускная способность канала ограничивается из-за возникновения межсимвольной интерференции, при которой символы "размазываются" по другим символам из-за расширения времени задержки.

Так как существует лишь одна "среда" для беспроводной передачи сигналов, для увеличения общей пропускной способности системы могут применяться различные схемы уплотнения. Стандартными схемами являются SDM, FDM, TDM и CDM. В схеме FDM данные необходимо преобразовать в сигнал с определенной несущей частотой. Поэтому существуют два этапа модуляции. При цифровой модуляции данные преобразовываются в модулирующий сигнал, а при аналоговой модуляции центральная частота модулирующего сигнала накладывается на несущую радиочастоту. Выше было представлено несколько улучшенных схем модуляции, позволяющих в отдельном сдвиге фазы кодировать сразу несколько битов данных, что способствует повышению эффективности передачи.

Технология расширения спектра сигнала применяется для достижения сразу нескольких целей. Одной из них является достижение (по крайней мере частичной) защищенности передачи — если код расширения неизвестен, то сигнал будет выглядеть как обычный шум. Более того, как будет показано в следующем разделе, технология расширения спектра положена в основу специальных схем доступа к среде, использующих кодовое пространство. Применение сигналов с расширенным спектром повышает устойчивость системы к узкополосным помехам, поскольку сигнал расширяется на более широкую полосу частот, и узкополосные помехи оказывают влияние только на малую часть всего спектра сигнала.

В конце этой главы рассматривались сотовые системы. В сотовых системах для увеличения общей пропускной способности мобильных телефонных систем используется схема пространственного разделения (SDM). Хотя для этих систем необходимо очень тщательное планирование (т.е. приведение размеров ячейки в соответствие к ожидаемому информационному обмену), они являются одним из главных способов эффективного использования скудных частотных ресурсов.

2.10. Упражнения

1. Частотные стандарты разных стран могут отличаться. Выясните, какие стандарты действуют в вашей стране (для Европы можно обратиться на Web-узел Европейской радиослужбы с адресом `http://www.ero.dk/`, а для США — на Web-узел комиссии FCC с адресом `http://www.fcc.gov/`).
2. Почему волны с очень низкой частотой могут огибать поверхность земли? Почему они не используются для передачи данных в компьютерных сетях?
3. Почему институт ITU-R регулирует только "нижние" частоты (до нескольких сотен ГГц), а не "верхние" (например, в диапазоне ТГц)?
4. Чем отличаются подход к регулированию разработок в области мобильных телефонных систем в Европе и США? Какие последствия имеют эти различия?
5. Почему открытый международный доступ к некоторым полосам частот ISM является важным вопросом?
6. Возможна ли передача с помощью радиоволн и без каких-либо потерь цифрового сигнала, закодированного в виде прямоугольных импульсов (как внутри компьютеров)? Почему?
7. Можно ли использовать направленные антенны в мобильных телефонах? Почему? Как можно повысить коэффициент направленного действия антенн?
8. Какие основные проблемы возникают при распространении сигналов? Почему радиоволны не всегда распространяются по прямой линии? Почему отражение радиоволн с одной стороны полезно, а с другой пагубно?
9. Назовите несколько способов ослабления межсимвольной интерференции. Как уровень межсимвольной интерференции зависит от несущей частоты, скорости передачи и собственного движения передатчика/приемника? К каким последствиям приводит межсимвольная интерференция в схемах временного уплотнения?
10. Какими способами можно ослабить влияние узкополосных помех? В чем основная сложность различных предлагаемых решений?
11. Почему в большинстве случаев для радиопередачи данных недостаточно лишь цифровой модуляции? Какие общие цели достигаются с помощью цифровой модуляции? Какие существуют типичные схемы подобной модуляции?
12. Представьте фазовую диаграмму и точки, представляющие комбинации битов в схеме PSK (рис. 2.29). Как при приеме определить, какая из комбинаций была послана, если полученная "точка" находится между другими точками диаграммы? В чем сложность кодирования все большего количества битов в одном фазовом сдвиге?
13. Какими основными преимуществами обладают системы, использующие технологию расширения спектра сигнала? Как достигается расширение спектра? Чем заменяются защитные интервалы, присутствующие на рис. 2.31, в схеме с расширением спектра, представленной на рис. 2.32? Как можно использовать многолучевое распространение сигналов в системах с технологией DSSS?

14. Какие основные причины использования сотовых систем? Как обычно реализуется схема SDM и как она объединяется со схемой FDM? Как схема DCA влияет на частоты, используемые другими ячейками?
15. Что ограничивает число одновременных пользователей в системах TDM/FDM по сравнению с системой CDM? Что происходит с качеством передачи при увеличении загрузки ячейки, т.е. как дополнительный пользователь влияет на других пользователей, присутствующих в ячейке?

2.11. Рекомендуемая литература

Anderson, J.B.; Rappaport, T.S.; Yoshida, S. [1995] "Propagation measurements and models for wireless communications channels", *IEEE Communications Magazine*, 33, (1).

Dixon, R. [1994] *Spread spectrum systems with commercial applications*. John Wiley.

ETSI [1997] *Digital Audio Broadcasting (DAB) to mobile, portable, and fixed receivers*, European Telecommunications Standards Institute, ETS 300 401.

Goodman, D. [1997] *Wireless personal communications systems*, Addison Wesley Longman.

Halsall, F. [1996] *Data communications, computer networks and open systems*, Addison Wesley Longman.

Peterson, R.; Ziemer, R.; Borth, D. [1995] *Introduction to spread spectrum communications*. Prentice Hall.

Stallings, W. [1997] *Data and computer communications*. Prentice Hall. (Русский перевод шестого издания этой книги готовится издательским домом "Вильямс" к выпуску в 2002 г.)

Viterbi, A. [1995] *CDMA: Principles of spread spectrum communication*. Addison Wesley Longman.

Wesel, E. [1997] *Wireless multimedia communications: networking video, voice and data*. Addison Wesley Longman.

Управление доступом к среде 3

В этой главе представлено несколько алгоритмов **управления доступом к среде** (Medium Access Control, MAC), приспособленных к беспроводной связи. Управление доступом к среде включает все механизмы, регулирующие доступ пользователя к среде при применении схем SDM, TDM, FDM или CDM. В примере с уплотнением движения по автомагистрали, рассмотренном в начале главы 2, роль механизма MAC играют правила дорожного движения. Например, для управления пересечением перекрестка несколькими машинами по схеме TDM необходимы определенные правила; которые реализуются с помощью светофора. В предыдущей главе рассматривались главным образом механизмы физического уровня — уровня 1 эталонной модели ISO/OSI. Схема MAC относится к уровню 2 — уровню **управления каналом данных** (Data Link Control, DLC). Он подразделяется на уровень 2b — управление логическим каналом (Logical Link Control, LLC), и уровень 2a — управление доступом к среде (MAC) [Halsall, 1996]. Задачей схемы DLC является установка надежного двухточечного или многоточечного соединения между различными устройствами через проводную или беспроводную среду. Основные механизмы схемы MAC обсуждаются в следующих разделах, а схема LLC и высшие уровни вместе с соответствующими специфическими технологиями будут рассматриваться в последующих главах.

В начале главы объясняется, почему для беспроводной связи необходимы специальные схемы MAC и почему стандартные схемы MAC для проводных сетей в этом случае часто не срабатывают. (Скрытые и открытые, а также ближние и дальние терминалы, в отличие от проводных сетей, здесь являются серьезной проблемой.) Затем обсуждается несколько механизмов MAC для различных схем уплотнения, представленных в главе 2. Схемы SDM и FDM обычно статичны, т.е. определенное пространство или частота (или последовательность перескоков частот) присваиваются им на продолжительное время. Поэтому в этой главе основное внимание сосредоточено на механизмах TDM. Работать по схеме TDM очень удобно, поскольку настройка на определенную фиксированную частоту не представляет проблемы, а интервалы времени можно выделять по требованию, распределяя их между потребителями тем или иным образом. Здесь самыми известными алгоритмами являются Aloha (несколько версий), различные схемы резервирования и простой последовательный опрос.

В конце главы вновь обсуждается применение схемы CDM. В частности, показано, как в схеме MAC, использующей уплотнение с кодовым разделением, происходит присваивание кодов, обеспечивающих разделение различных пользователей в кодовом пространстве. Обычно используются не отдельные схемы в чистом виде, а смешанные схемы. Это делается для того, чтобы воспользоваться преимуществами различных схем. Завершает главу сравнение четырех основных схем доступа.

3.1. Специализированные схемы MAC как необходимость

Главный вопрос, возникающий при рассмотрении схем MAC для беспроводной связи, — можно ли в этих целях использовать схемы управления доступом к среде, разработанные для проводных сетей. В частности, можно ли использовать схему CSMA/CD так, как она описана в исходной спецификации для сетей IEEE 802.3 (известных также как сети Ethernet).

Поэтому мы начнем с рассмотрения схемы **множественного доступа к среде с обнаружением конфликтов и детектированием несущей** (Carrier Sense Multiple Access with Collision Detection, CSMA/CD), работающей следующим образом. Передатчик проверяет состояние среды (провода или коаксиального кабеля) с целью определения, является ли она свободной. Если среда занята, передатчик ожидает, пока она освободится. Если среда свободна, передатчик начинает передавать данные, продолжая следить за ее состоянием. При обнаружении конфликта при передаче передатчик останавливается и посылает сигнал о наличии помехи.

Почему эта схема не срабатывает в беспроводных сетях? По сути, в схеме CSMA/CD целью является избежать возникновения конфликтов не в передатчике, а в приемнике, т.е. сигнал должен достичь приемника без конфликтов. Однако, задача выявления возникающих конфликтов возлагается на передатчик. При использовании проводов это не является проблемой, так как интенсивность сигнала примерно одинакова по всему проводу. Если в каком-то месте провода возникает конфликт, он легко наблюдается всеми передатчиками и приемниками.

В беспроводных сетях ситуация совершенно иная. В главе 2 было показано, что интенсивность сигнала убывает пропорционально квадрату расстояния до передатчика. Передатчик при проверке наличия несущей может определить, что среда свободна и, соответственно, начать передачу данных. Но в приемнике может возникнуть конфликт из-за близкого соседства другого передатчика. Эта проблема скрытых терминалов подробно обсуждается в разделе 3.1.1. То же самое может происходить и при детектировании возникновения конфликтов. Передатчик не обнаруживает конфликтов и полагает, что передача данных прошла без них. На самом же деле наличие конфликтов могло вызвать разрушение данных при приеме. Таким образом, эта очень популярная схема MAC совершенно непригодна для беспроводной связи. В следующих разделах приведено еще несколько примеров недееспособности схем, применяющихся в стационарных сетях.

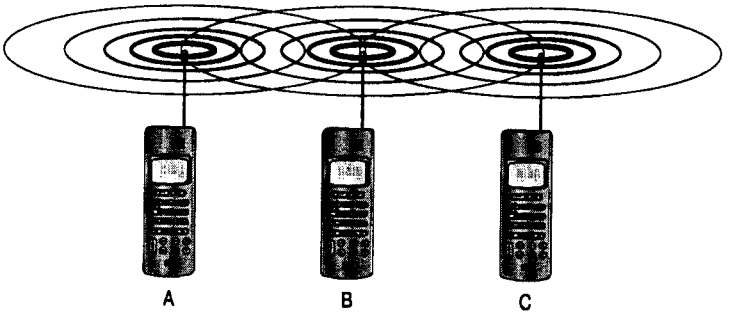

Рис. 3.1. Скрытые и открытые терминалы

3.1.1. Скрытые и открытые терминалы

Рассмотрим схему из трех мобильных телефонов, показанную на рис. 3.1. Зона передачи телефона A достигает телефона B, но не достигает телефона C (а значит, телефон C не попадает и в зону детектирования). Зона передачи телефона C достигает телефона B, но не достигает телефона A. Наконец, телефоны A и C попадают в зону передачи телефона B, но телефон A не может обнаружить сигнал телефона C и наоборот.

Начинается передача данных с телефона A на телефон B. Телефон C не принимает эту передачу. Допустим, что ему тоже необходимо передать что-то на телефон B, поэтому телефон C определяет состояние среды. Несущая частота не детектируется, и среда выглядит свободной. В результате телефон C также начинает передачу, что вызывает конфликт в телефоне B. Но телефон A не может обнаружить этот конфликт и продолжает свою передачу. Телефон A является **скрытым** для телефона C и наоборот.

В то время как скрытые терминалы вызывают конфликты, эффект, описанный ниже, приводит лишь к ненужным задержкам. Рассмотрим ситуацию, когда телефон B посылает что-то на телефон A, а телефону C необходимо передать данные на другой мобильный телефон, находящийся вне зон помех от телефонов A и B. Телефон C определяет наличие несущей частоты и приходит к заключению, что среда занята. В результате передача откладывается. Однако, поскольку телефон A находится вне зоны помех от телефона C, такое ожидание в действительности излишне. "Конфликт" в телефоне B не имеет значения, поскольку конфликтующий сигнал слишком слаб для того, чтобы достичь телефона A. В этой ситуации телефон C является для телефона B **открытым**.

3.1.2. Ближние и дальние терминалы

Рассмотрим ситуацию, изображенную на рис. 3.2. Пусть телефоны A и B передают данные с одинаковой мощностью передачи. Поскольку интенсивность сигнала убывает пропорционально квадрату расстояния, сигнал телефона B заглушает сигнал телефона A. В результате телефон C не может принять передачу телефона A.

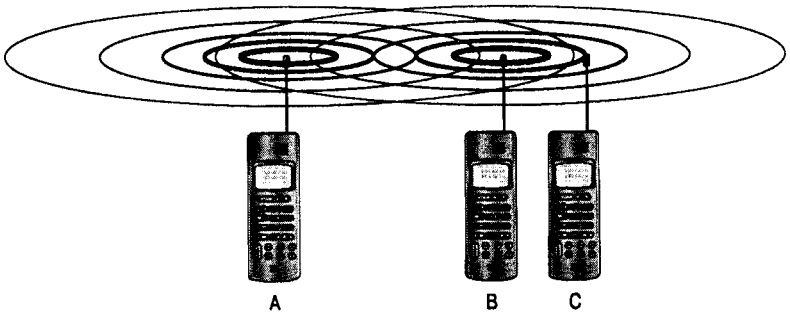

Рис. 3.2. Ближние и дальние терминалы

Теперь представим, что телефон C распределяет права на проведение передачи. Но в рассматриваемом случае терминал B заглушает терминал A на физическом уровне. В результате телефон C не может корректно применить схему управления, поскольку принимает только сигнал терминала B.

Эффект приближенности/удаленности является серьезной проблемой в беспроводных сетях, использующих схему CDM. Необходимо, чтобы все принимаемые сигналы имели примерно одинаковую интенсивность. Однако, возвращаясь к примеру разговоров на вечеринке из главы 2, человек, находящийся ближе к кому-то, всегда кажется ему говорящим громче, чем человек, находящийся дальше от него. В результате, даже когда передатчики разделены кодами, самый близкий из них может просто заглушить всех остальных. Поэтому для достижения примерно одинаковой интенсивности принимаемого сигнала от всех передатчиков необходим строгий контроль мощности их передачи.

3.2. SDMA

В схеме **множественного доступа с пространственным разделением** (Space Division Multiple Access, SDMA) между пользователями беспроводных сетей распределяются отдельные участки пространства. В качестве примера можно указать присвоение пользователю мобильного телефона оптимальной базовой станции. Мобильный телефон может принимать несколько базовых станций, но качество связи с каждой из них будет различным. Алгоритм MAC в каждом случае должен обеспечивать выбор, какая из базовых станций является наилучшей, принимая во внимание доступные частоты (FDM), временны*е* интервалы (TDM) или коды (CDM), в зависимости от используемой технологии. Обычно схема SDMA используется не сама по себе, а в комбинации с еще одной или несколькими другими схемами. В основу реализации алгоритма SDMA положен механизм ячеек и секторных антенн. Они образуют инфраструктуру, позволяющую реализовать **уплотнение с пространственным разделением** (SDM) (см. раздел 2.5.1).

3.3. FDMA

Схема **множественного доступа с частотным разделением** (Frequency Division Multiple Access, FDMA) включает все алгоритмы, реализующие распределение частот для каналов передачи согласно схеме **уплотнения с частотным разделением** (FDM), представленной в разделе 2.5.2. Распределение может быть фиксированным (например, для радиостанций или при общем планировании и регулировании частот) или динамическим (т.е. осуществляемым по запросам).

В "чистой" схеме FDMA частота присваивается некоторому каналу на все время его существования. Однако, можно скомбинировать схемы FDMA и TDMA, изменяя частоты по некоторому установленному образцу. В этом случае говорят о схеме с перескоком частот, которая широко применяется в беспроводных системах с целью исключения появления узкополосных помех. Передатчик и приемник должны согласовать используемую последовательность перескоков, иначе приемник не сможет настроиться на нужную частоту. Поэтому последовательности перескоков фиксированы, по крайней мере, на протяжении некоторого времени. В частотном пространстве невозможно резко перескочить на произвольную частоту (т.е. приемнику необходимо время, чтобы настроиться на нужную частоту). Это одно из главных различий между схемами FDM и TDM.

Кроме того, схема FDM часто применяется в сотовых сетях для организации одновременного доступа к среде базовой и мобильной станций. В этом случае между партнерами по связи устанавливается **дуплексный канал**. Для каждого из двух направлений — от мобильной станции к базовой и наоборот — используются различные частоты. Данная схема называется **дуплексом с частотным разделением** (Frequency Division Duplex, FDD). Опять же, обоим партнерам необходимо заблаговременно знать частоты, а не просто прослушивать среду. Одна из двух частот используется для **исходящей** линии связи (от мобильной станции до базовой станции или от наземного управления до спутника), другая — для **нисходящей** линии (от базовой до мобильной станции или от спутника до наземного управления).

В качестве примера схем FDM и FDD на рис. 3.3 приводится ситуация, возникающая в мобильной телефонной сети, организованной по стандарту GSM для частоты 900 МГц. Основная схема распределения частот в стандарте GSM фиксирована. Все исходящие линии связи используют частоты в полосе между 890,2 МГц и 915 МГц, а для нисходящих линий выделяется диапазон частот от 935,2 МГц до 960 МГц. Согласно схеме FDMA для установления дуплексного канала между базовой станцией (показана справа) и мобильным телефоном выделяются две частоты — для исходящей и нисходящей линии связи. Эти две частоты связаны между собой. Если исходящая частота для некоторого канала n равна $f_u = 890$ МГц $+ n \cdot 0{,}2$ МГц, то нисходящая частота связана с ней соотношением $f_d = f_u + 45$ МГц, т.е. $f_d = 935$ МГц $+ n \cdot 0{,}2$ МГц. Этот

пример иллюстрирует использование схемы FDM для множественного доступа и организации дуплекса согласно заранее предопределенной схеме. Похожие схемы FDM для организации дуплекса с частотным разделением применяются в системах AMPS, IS-54, IS-95, IS-136 и PACS. В главе 4 подробно обсуждается объединение этой схемы со схемой TDM, использующееся в системе GSM.

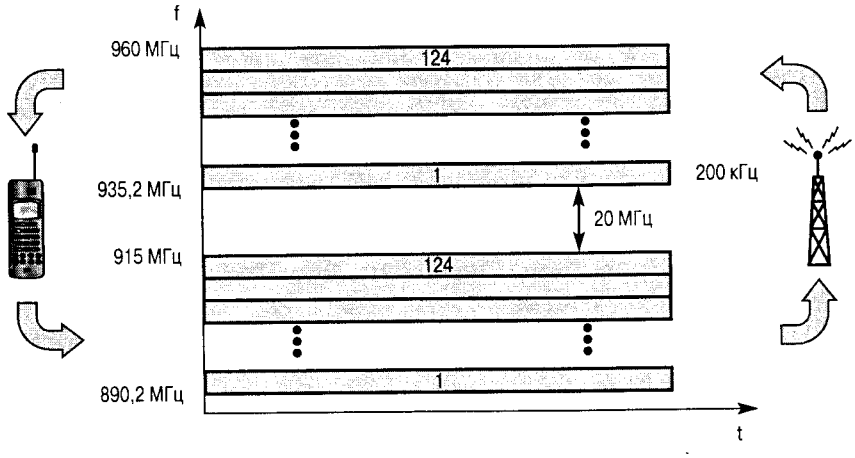

Рис. 3.3. Использование частотного уплотнения для множественного доступа и организации дуплекса

3.4. TDMA

Схема **множественного доступа с временным разделением** (Time Division Multiple Access, TDMA) в сравнении со схемой FDMA является более гибкой. Она включает все технологии, предусматривающие распределение интервалов времени, т.е. уплотнение по принципу TDM. В этом случае перестройка на различные частоты не требуется, поскольку приемник все время может оставаться на одной и той же частоте. При использовании только одной частоты передачи конструкция передатчиков и приемников упрощается, при этом сохраняется возможность реализации многих различных алгоритмов управления доступом к среде. Как уже упоминалось выше, прослушивание различных частот в одно и то же время — довольно трудная задача, тогда как прослушивать на одной и той же частоте разделенные во времени каналы гораздо проще. Почти все схемы MAC для проводных сетей работают именно по этому принципу. В качестве примера можно привести технологии Ethernet, Token Ring, ATM и многие другие [Halsall, 1996].

В данной схеме необходимо обеспечить синхронизацию передатчика и приемника во временной области. Опять же, для этой цели можно использовать фиксированную схему (т.е. выделить каждому каналу определенные временные интервалы) или схему с динамическим распределением интервалов времени. В случае динамического распределения, как и в обычных проводных схемах MAC, для каждой передачи потребуется использовать идентификацию (например, указывать адрес передатчика). В фиксированных схемах распределения доступа идентификация не нужна, но эти схемы являются менее гибкими в отношении изменения требований к пропускной способности системы. В последующих разделах представлено несколько примеров фиксированных и динамических схем, использующихся для беспроводной передачи. Чаще всего для достижения дополнительной гибкости эти схемы объединяются со схемой FDMA.

3.4.1. Фиксированная схема TDM

Фиксированное распределение временных интервалов между каналами является простейшим алгоритмом использования схемы TDM. Этот алгоритм обеспечивает фиксированную пропускную способность системы и часто используется в беспроводных телефонных системах. Схема MAC в этом случае совсем проста, так как единственный важный фактор здесь — использование зарезервированного временного интервала в нужный момент. Если это удается организовать, то каждой мобильной станции устанавливается ее очередь, чем исключаются любые взаимные помехи. Фиксированное распределение временных интервалов может осуществляться базовой станцией, разрешающей "споры" между различными мобильными станциями, нуждающимися в доступе к среде.

Фиксированные (по крайней мере, в некотором промежутке времени) схемы доступа идеально подходят для соединений с фиксированной пропускной способностью. Кроме того, эти схемы гарантируют фиксированную задержку. Например, в системах стандарта DECT передача данных гарантированно происходит каждые 10 мс. Схемы TDMA с фиксированным распределением доступа используются во многих цифровых мобильных телефонных системах, в частности, в системах IS-54, IS-136, GSM, DECT, PHS и PACS.

На рис. 3.4 показано, как фиксированные схемы TDM применяются для организации множественного доступа и установления дуплексного канала между базовой и мобильной станциями. Присвоение различных интервалов исходящей и нисходящей линиям связи, использующим одинаковую несущую частоту, называется **дуплексом с временным разделением** (TDD). Базовая станция, изображенная на рисунке, использует для нисходящей линии один из двенадцати доступных временных интервалов. Мобильная станция для исходящей линии связи использует один из следующих двенадцати временных интервалов. Таким образом, работа исходящей и нисходящей линий разделена во времени. Более того, в этой схеме одинаковую частоту без взаимных помех могут использовать до 12 различных мобильных станций. В приведенном ниже примере, который является стандартным для беспроводной телефонной системы DECT, схема распределения сигналов повторяется каждые 10 мс, т.е. каждый выделенный интервал имеет продолжительность 417 мкс. Подобный механизм обеспечивает каждой мобильной станции доступ к среде каждые 10 мс, причем независимо от других соединений.

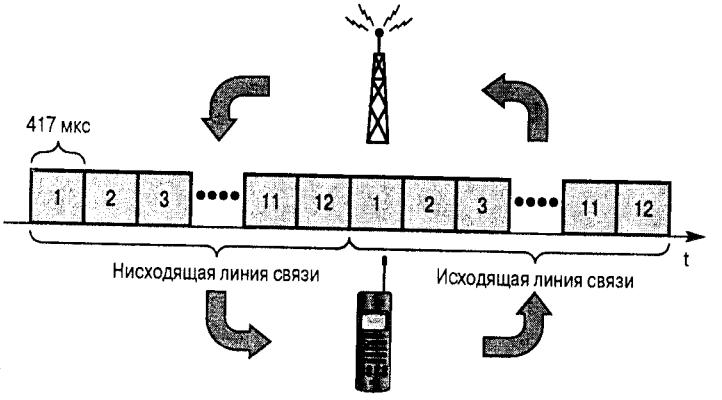

Рис. 3.4. Использование временного уплотнения для организации множественного доступа при дуплексных соединениях

Фиксированные схемы доступа, как показывает пример системы DECT, идеально подходят для соединений с постоянной интенсивностью передачи данных. В то же время для пульсирующего информационного обмена или асимметричных соединений их использование оказы-

вается неэффективным, поскольку в этом случае поток данных от базовой станции к мобильной или наоборот оказывается неравномерным. В качестве примера можно привести работу пользователя с Web-броузером. Во время просмотра уже загруженной страницы никакой передачи данных не происходит, но щелчок мышью на некоторой гиперссылке вызовет передачу данных от мобильной станции к базовой. Далее может последовать поступление некоторого, возможно, большого количества данных от Web-сервера. Хотя в системе DECT и можно сделать пропускную способность асимметричной (см. раздел 4.2), эта общая схема все равно слишком статична и неудобна для передачи данных. Довольно большая часть общей пропускной способности системы при этом теряется впустую. В рассмотренных выше случаях можно использовать схемы TDMA, управляемые по запросам и работающие без установки соединения. Эти схемы обсуждаются в последующих разделах.

3.4.2. Классическая схема Aloha

Как уже упоминалось, схема TDMA охватывает все механизмы, управляющие доступом к среде по принципу TDM. Но что произойдет, если схему TDM применить без управления доступом? Именно этот механизм реализован в классической схеме управления **Aloha**. Данная схема была изобретена в Гавайском университете и применялась в сети ALOHANET для беспроводного соединения нескольких станций. Схема Aloha не только не координирует доступ к среде, но и не решает конфликтов на уровне MAC. Вместо этого каждой станции доступ к среде предоставляется в любое время, как показано на рис. 3.5. Это схема случайного доступа без централизованного разрешения конфликтов и координации действий между отдельными станциями. Когда доступ к среде осуществляют одновременно две или больше станций, возникает **конфликт**, и передаваемые данные разрушаются. Данная проблема решается на высших уровнях системы (например, с помощью механизма повторной передачи данных).

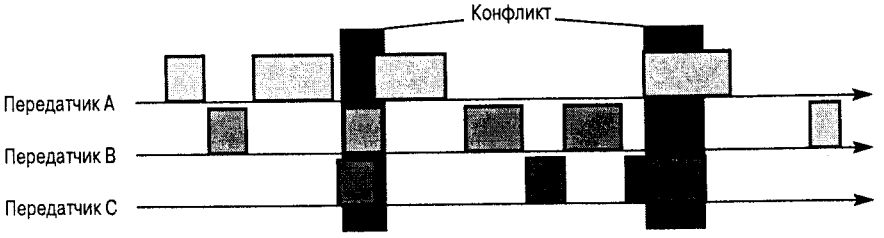

Рис. 3.5. Множественный доступ в классической схеме Aloha

Простая схема Aloha хорошо работает при слабой загрузке и не требует никаких сложных механизмов управления доступом к среде. При пуассоновском распределении частот поступления пакетов данных максимальная пропускная способность системы достигается при 18-процентной загрузке [Abramson, 1977], [Halsall, 1996].

3.4.3. Aloha с разделением

Первым усовершенствованием классической схемы Aloha было введение временных интервалов (**Aloha с разделением**). В этом случае все передатчики должны быть **синхронизированы**, а передача может начинаться только в начале **временного интервала**, как показано на рис. 3.6. Доступ к среде по-прежнему не координируется. При сделанных выше допущениях введение интервалов увеличивает уровень загрузки, при которой достигается максимальная пропускная способность системы, с 18 до 36 процентов.

В последующих разделах будет показано, что оба основных принципа схемы Aloha присутствуют во многих системах, реализующих распределенный доступ к среде. Системы Aloha от-

лично работают при относительно слабой загрузке (как и большинство прочих систем). В то же время они не предоставляют твердых гарантий качества передачи, в частности, не оговаривается максимальная задержка доступа к среде и минимальная пропускная способность всей системы. Здесь необходимы дополнительные механизмы, например, объединение схемы Aloha с фиксированными схемами.

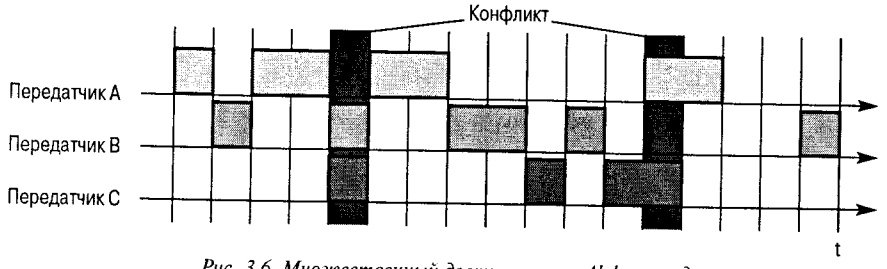

Рис. 3.6. Множественный доступ в схеме Aloha с разделением

3.4.4. Множественный доступ с детектированием несущей

Детектирование наличия несущей частоты перед доступом к среде является еще одним возможным усовершенствованием основной схемы Aloha. Оно обычно используется в схемах **множественного доступа с детектированием несущей** (Carrier Sense Multiple Access, CSMA) [Kleinrock, 1975]; [Halsall, 1996]. Детектирование несущей частоты и разрешение доступа к среде только тогда, когда несущая отсутствует, уменьшает вероятность возникновения конфликтов. Однако, как уже упоминалось в начале главы, скрытые терминалы обнаружить невозможно. Поэтому, если скрытый терминал передает данные одновременно с другим передатчиком, в приемнике может возникнуть конфликт. Данная простая схема управления доступом используется в большинстве беспроводных ЛВС (эти вопросы подробнее освещаются в главе 7).

Существует несколько версий схемы CSMA. При использовании **неустойчивой схемы CSMA** станции детектируют несущую и, если среда свободна, немедленно начинают передачу. Если среда занята, станция перед определением состояния среды выжидает случайный промежуток времени и повторяет свои действия. В ***p*-устойчивых системах CSMA** узлы тоже определяют состояние среды, но данные передаются с вероятностью p. Станция может отложить передачу до следующего временного интервала с вероятностью $1 - p$, т.е. доступ дополнительно разделен. В **1-устойчивых системах CSMA** все станции, которым необходимо передавать данные, одновременно получают доступ к среде, как только она освобождается. Для того, чтобы станции, ожидающие длительное время, имели некоторое преимущество в доступе к среде, можно использовать алгоритмы с обратной связью, чувствительные ко времени ожидания. Подобные механизмы применяются, в частности, в стандартной системе Ethernet [Halsall, 1996].

Одной из схем доступа, использующихся в беспроводных ЛВС стандарта IEEE 802.11, является схема CSMA с исключением конфликтов (CSMA/CA). Здесь детектирование несущей объединяется с механизмом обратной связи для того, чтобы в случае занятости среды обеспечить хотя бы некоторое равноправие среди конкурирующих станций. В спецификации HIPERLAN 1 используется еще одна усовершенствованная схема — бесприоритетный множественный доступ с исключением (Elimination Yield — Non-Preemptive Multiple Access, EY-NMPA). Здесь для разрешения конфликтов поочередно применяются несколько этапов определения состояния среды и доступа к ней, пока в конце концов один "победитель" не получит право передачи данных. Чтобы обеспечить преимущество для отдельных станций с более важными данными, в эту схему можно включить определенную систему приоритетов.

3.4.5. Множественный доступ с распределением по запросу

Общего усовершенствования систем доступа Aloha можно также добиться с помощью механизмов **резервирования** и объединения с некоторыми (фиксированными) схемами TDM. В этих схемах после периода передачи обычно следует период резервирования. На протяжении периода резервирования станции могут забронировать за собой будущие временные интервалы в периоде передачи. В течение периода резервирования могут происходить конфликты. В то же время, в зависимости от схемы, на протяжении периода передачи конфликты или вообще не возникают, или период передачи разбивается на меньшие периоды — с конфликтами и без них. Вообще говоря, эти схемы приводят к большей задержке при слабой загрузке системы, но обеспечивают ей более высокую пропускную способность.

Одной из основных схем TDM является **множественный доступ с распределением по запросу** (Demand Assigned Multiple Access, DAMA), называемый также схемой **Aloha с резервированием**. Эта схема часто применяется в спутниковых системах и имеет два режима (рис. 3.7). На этапе борьбы за доступ к среде, проводящейся в соответствии со схемой Aloha с разделением, все станции пытаются зарезервировать для себя будущие временные интервалы. Например, различные наземные станции пытаются зарезервировать за собой время доступа для спутниковой передачи. Поэтому конфликты на стадии резервирования не приводят к повреждению данных, а влияют только на относительно короткие запросы на передачу данных. Если будущий временной интервал удалось зарезервировать, то ни одна другая станция не может в это время осуществлять передачу. Таким образом, спутник собирает все успешные запросы (остальные разрушаются) и посылает назад список с указанием прав доступа к последующим интервалам. Этому списку подчиняются все наземные станции. Для поддержания работы фиксированной схемы резервирования и передачи (TDM) все станции время от времени синхронизируются. Схема DAMA относится к схемам с **явным резервированием**. Каждый интервал для передачи резервируется явно.

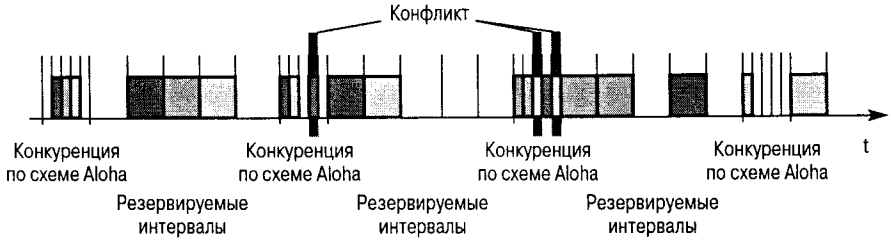

Рис. 3.7. Множественный доступ с распределением по запросу и явным резервированием

3.4.6. Множественный доступ PRMA с резервированием пакетов

Множественный доступ с резервированием пакетов (Packet Reservation Multiple Access, PRMA) является примером схемы со **скрытым резервированием**. Здесь интервалы резервируются неявно согласно следующей схеме. Определенное число интервалов образует кадр (на рис. 3.8 кадр содержит восемь интервалов). Кадры повторяются во времени (на рисунке показаны кадры с первого по пятый), т.е. применяется фиксированная схема TDM.

Базовая станция распространяет информацию о состоянии каждого интервала между всеми мобильными станциями (как показано на рисунке слева). После получения этого вектора всем станциям будет известно, какой из интервалов занят, а какой в данное время свободен. В приведенном на рис. 3.8 примере базовая станция рассылает для всех станций от A до F состояние схемы резервирования, имеющее вид "ACDABA-F". Это означает, что в следующем цикле пе-

редачи интервалы с первого по шестой и восьмой интервал заняты, а седьмой интервал пока свободен. Все станции, которым необходимо передавать данные, состязаются между собой согласно схеме Aloha за право использовать этот интервал. В описанном примере на него претендует более чем одна станция, поэтому возникает конфликт. Базовая станция возвращает состояние "ACDABA-F", указывая на то, что резервирование седьмого интервала оказалось неудачным и он пока не зарезервирован, а для других интервалов тоже ничего не изменилось. Станции опять могут состязаться за седьмой интервал. Кроме того, станции D и F на третьем и восьмом интервалах приостановили передачу. Все это отмечается базовой станцией после завершения второго кадра.

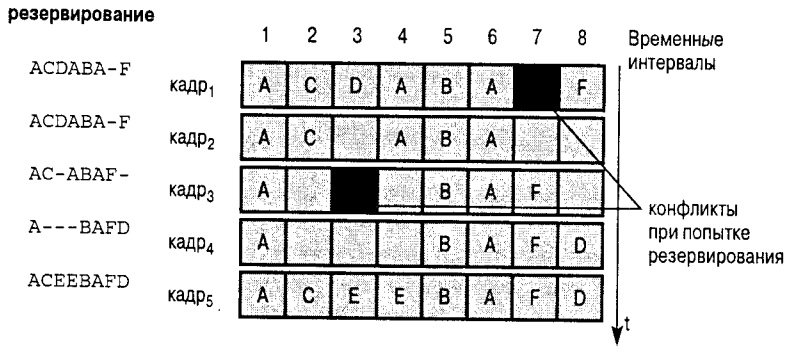

Рис. 3.8. Множественный доступ с распределением по запросу и скрытым резервированием

В начале третьего кадра базовая станция указывает, что теперь третий и восьмой интервалы освободились. Кроме того, базовой станцией отмечается, что станции F удалось зарезервировать седьмой интервал. Технология PRMA включает еще одну комбинацию фиксированных и случайных схем TDM с резервированием, отличную от рассмотренной выше. Как только какой-то станции удается выполнить резервирование, за ней неявно закрепляются все будущие интервалы данной серии. Этот механизм обеспечивает постоянную скорость передачи данных в системе. Схема Aloha с разделением используется только для свободных интервалов, а начатая передача данных не разрушается конфликтами.

3.4.7. Схема TDMA с резервированием

Схема TDMA с резервированием (рис. 3.9) имеет еще более фиксированную структуру, также обеспечивающую определенный случайный доступ. В фиксированной схеме TDM кадр образуется из N мини-интервалов, за которыми следуют $N \cdot k$ интервалов для данных. Каждой станции присваивается ее собственный мини-интервал. С его помощью она может зарезервировать до k интервалов для данных. Это обеспечивает каждой станции некоторую пропускную способность и фиксированную задержку. Остальные станции могут пересылать данные через свободные интервалы. Использование этих свободных интервалов можно регулировать простой круговой схемой, а можно и не координировать, применив схему Aloha. Подобная схема позволяет использовать, например, комбинацию изохронного информационно обмена с фиксированной скоростью передачи данных и максимально интенсивного информационного обмена без каких-либо гарантий.

3.4.8. Множественный доступ с исключением конфликтов

Вернемся к одной из важнейших исходных проблем — скрытым терминалам. Как рассматривавшиеся выше схемы доступа решают этот вопрос? Системы с центральными базовыми стан-

циями, использующими схему TDM, не испытывают проблем со скрытыми терминалами. Если для базовой станции терминал скрыт, она никак не может с ним связаться. Однако, как уже упоминалось, более или менее фиксированные схемы доступа не такие гибкие, как схемы Aloha. Что происходит, если базовых станций вообще нет, как в специализированных сетях? Эти вопросы подробнее обсуждаются в главе 7.

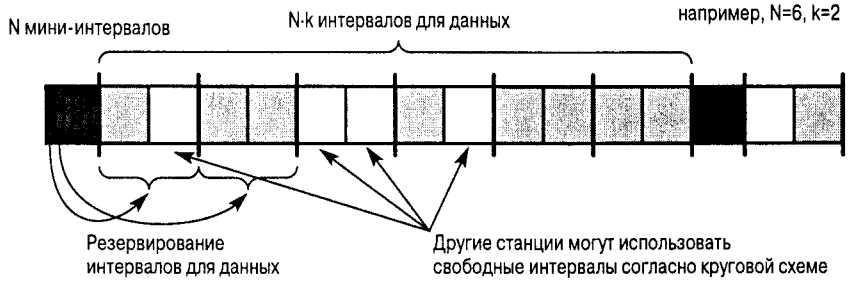

Рис. 3.9. Схема доступа TDMA с резервированием

Множественный доступ с избежанием конфликтов (Multiple Access with Collision Avoidance, MACA) является простой схемой, решающей проблему скрытых терминалов. Эта схема не требует базовых станций и относится к схемам Aloha со случайным доступом, но с динамическим резервированием. На рис. 3.10 изображена та же ситуация, что и на рис. 3.1. Напомним, что телефонам A и C необходимо передать данные на телефон B. Телефон A начинает передачу, но для телефона C он является скрытым. Поэтому, когда телефон C начинает свою передачу, в телефоне B возникает конфликт.

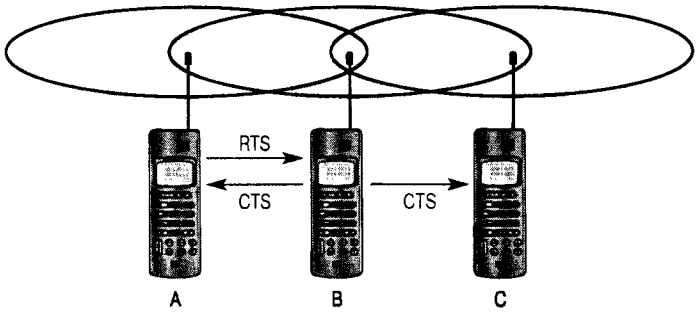

Рис. 3.10. Решение проблемы скрытых терминалов в схеме MACA

В схеме MACA телефон A, перед тем как начать передачу, посылает сигнал **"готов к передаче"** (RTS). Телефон B принимает сигнал RTS, содержащий названия передатчика и приемника, а также продолжительность будущей передачи. Сигнал RTS не достигает телефон C. Далее инициируется отправка телефоном B сигнала подтверждения **"готов к приему"** (CTS). Сигнал CTS тоже содержит названия передатчика и приемника и продолжительность будущей передачи. Этот сигнал достигает телефон C и среда на время передачи резервируется для будущего ее использования телефоном A. После приема сигнала CTS телефону C не разрешается посылать что-либо на телефон B на протяжении всего времени, указанного в сигнале. Таким образом, во время передачи данных конфликт в телефоне B возникнуть не может, и проблема скрытых терминалов решается — условия передачи остаются неизменными.

Конфликты все еще могут возникать при передаче сигнала RTS. Этот сигнал могут одновременно послать и телефон A, и телефон C. Однако, в сравнении с временем передачи дан-

ных, сигнал RTS является очень коротким, поэтому вероятность конфликта значительно ниже. Телефон B разрешает проблему состязания, указывая в сигнале CTS только одну станцию (если ему вообще удается восстановить сигнал RTS). Без соответствующего сигнала CTS передача не разрешается. Описанная схема является одной из дополнительных схем, принятых в стандарте IEEE 802.11 (более подробно эти вопросы рассматриваются в разделе 7.3).

Может ли схема MACA помочь и в решении проблемы "открытых терминалов"? Напомним, что в этом случае телефону B нужно передать данные на телефон A, а телефону C — какому-то другому телефону. Телефон C определяет состояние среды. Она занята передачей с телефона B. Поэтому телефон C откладывает свою передачу, несмотря на то, что она не может вызвать конфликт в телефоне A.

При использовании схемы MACA телефон B сперва передает сигнал RTS (рис. 3.11), содержащий название приемника (A) и передатчика (B). Телефон C на это сообщение не реагирует, так как не является приемником. В то же время телефон A посылает сигнал CTS, в котором телефон B определяется как передатчик, а телефон A — как приемник будущей передачи. Телефон C не принимает этого сигнала CTS и заключает, что телефон A находится вне зоны детектирования. Поэтому телефон C может начинать свою передачу, не вызывая конфликта в телефоне A. Проблема открытых терминалов решается без фиксированных схем доступа и базовых станций. Одним из недостатков схемы MACA являются дополнительные затраты, связанные с передачей сигналов RTS и CTS; если пакеты данных относительно небольшие, этими затратами нельзя пренебречь.

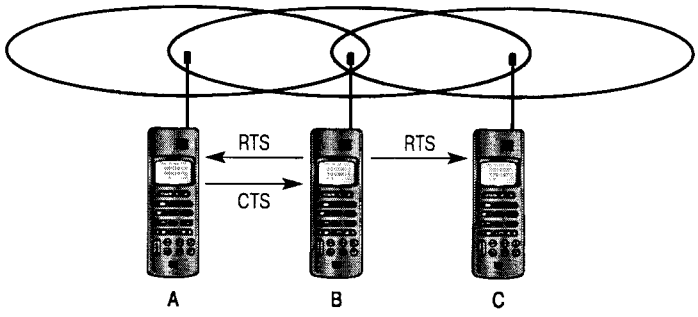

Рис. 3.11. Решение проблемы открытых терминалов в схеме MACA

На рис. 3.12 показаны упрощенные диаграммы состояний передатчика и приемника, работающих по схеме MACA. Передатчик бездействует, пока пользователь не потребует передачи пакета данных. Затем передатчик посылает сигнал RTS и ожидает получения права на передачу. Если приемник бездействует и получает сигнал RTS, он посылает обратно сигнал CTS и ожидает поступления данных. Передатчик принимает сигнал CTS и посылает данные. В другом случае (например, когда сигнал RTS потерян или вступил в конфликт) передатчик должен после некоторого перерыва опять послать сигнал RTS. После передачи данных передатчик ожидает подтверждения, чтобы вернуться в состояние бездействия. Если получены правильные данные, приемник посылает положительное подтверждение. В ином случае или по прошествии определенного времени ожидания приемник возвращается в бездействующее состояние. Если передатчик не получает никакого подтверждения приема или отрицательное подтверждение, он опять посылает сигнал RTS и ожидает получения права на передачу. Кроме того, приемник может указать, что он занят, с помощью отдельного сигнала RxBusy. В действительности, требуется еще много состояний и переходов, например, для ограничения числа повторных передач.

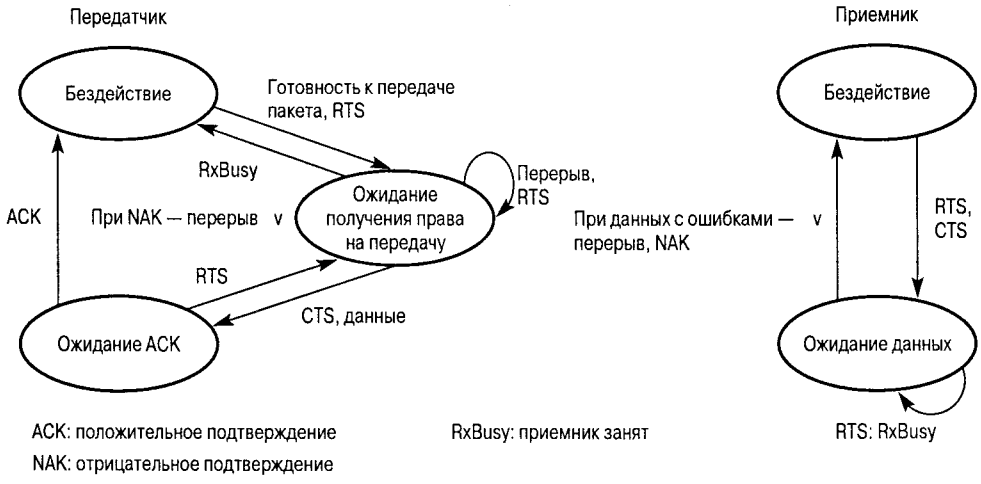

Рис. 3.12. Протокольные механизмы для организации множественного доступа с избежанием конфликтов

3.4.9. Опрос

Рассмотрим случай, когда передача одной из станций должна приниматься всеми остальными (примером может быть базовая станция в мобильной телефонной сети или любая другая выделенная станция). В этой ситуации можно применить схему **опроса**, широко применявшуюся ранее в терминальных сетях мейнфреймов. Опрос — это строго централизованная схема с одной главной и несколькими подчиненными станциями. Главная станция может опрашивать подчиненные станции несколькими способами: циклически (эта схема эффективна только в том случае, когда интенсивность информационного обмена у всех станций примерно одинакова), случайно, согласно резервированию (в качестве примера можно рассмотреть занятие с подготовленными студентами) и т. д. Кроме того, главная станция на протяжении фазы состязания за доступ может составить список станций, которым необходимо передавать данные. После этого проводится опрос среди всех станций из списка. Похожие схемы используются, в частности, в беспроводных ЛВС системы Bluetooth и в качестве одной из возможных функций доступа в системах стандарта IEEE 802.11, описанных в разделе 7.5.

3.4.10. Множественный доступ с детектированием подавления

Множественный доступ с детектированием подавления (Inhibit Sense Multiple Access, ISMA) является еще одной комбинацией нескольких различных схем. Он используется службой сотовой передачи цифровых пакетов данных (CDPD) в мобильной телефонной системе AMPS (эту схему также называют **множественным доступом с цифровым детектированием** — DSMA, Digital Sense Multiple Access). Здесь базовая станция извещает мобильные станции о занятости среды с помощью сигнала "занято" на нисходящей линии связи (рис. 3.13). После снятия сигнала "занято" доступ к исходящей линии связи не координируется. В случае успешной передачи базовая станция посылает подтверждение. Мобильные станции могут обнаружить конфликт только через отсутствие положительного подтверждения приема. При конфликтах применяются дополнительные механизмы отката и повторной передачи.

3.5. CDMA

Наконец, при использовании **уплотнения с кодовым разделением** (CDM) при передаче данных применяются коды с определенными характеристиками. В системах **множественного доступа с**

кодовым разделением (Code Division Multiple Access, CDMA) эти коды используются для разделения различных пользователей в кодовом пространстве. Именно они позволяют получать доступ к общей среде без помех. Главная задача — найти "хорошие" коды и отделить нужный сигнал от шума, создаваемого окружением и остальными сигналами.

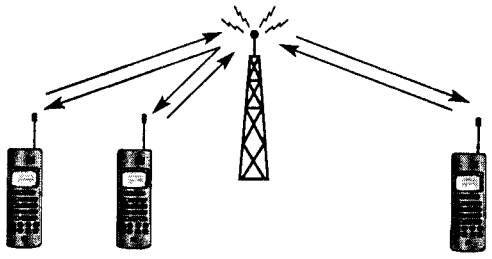

Рис. 3.13. Использование сигнала "занято" при множественном доступе с детектированием подавления

В главе 2 было показано, как коды применяются для расширения сигнала (например, с помощью схемы DSSS). В этом случае коды непосредственно управляют чиповой последовательностью. Но что такое "хороший код" в отношении схемы CDMA? Код каждого пользователя должен быть **ортогональным** ко всем остальным кодам и обладать хорошей автокорреляцией. Ортогональность в кодовом пространстве имеет тот же смысл, что и в обычном трехмерном пространстве. Представим себе трехмерную систему координат и векторы, исходящие из ее начала, т.е. из точки (0, 0, 0). Два вектора называются ортогональными, если их скалярное произведение равно нулю, как в случае двух векторов (2, 5, 0)*(0, 0, 17) = 0 + 0 + 0 = 0. Векторы (3, –2, 4) и (–2, 3, 3) тоже являются ортогональными: (3, -2, 4)*(–2, 3, 3) = –6 – 6 + 12 = 0. Однако, векторы (1, 2, 3) и (4, 2, –6) неортогональны друг другу (скалярное произведение равно –10), а векторы (1, 2, 3) и (4, 2, –3) "почти" ортогональны (скалярное произведение равно –1).

Теперь переведем эти результаты в кодовое пространство и объясним, что подразумевается под хорошей **автокорреляцией**. К примеру, код Баркера (+1, –1, +1, +1, –1, +1, +1, +1, –1, –1, –1) обладает хорошей автокорреляцией, т.е. его скалярный квадрат имеет большое значение, равное 11. При сдвигах кода Баркера на один или несколько чипов (представьте сдвиг 11-чипового кода Баркера относительно самого себя, повторенный несколько раз) абсолютная величина корреляции не превышает 1. Ее значение будет оставаться таким низким до тех пор, пока код опять точно не совместится сам с собой. Такое пиковое поведение обеспечивает точное восстановление данных при приеме.

Основы функционирования схемы CDMA можно объяснить с помощью следующих (теоретических) примеров.

- Двум передатчикам, A и B, необходимо передать данные. Схема CDMA присваивает им следующие ключевые последовательности: ключ $A_k = 010011$ и ключ $B_k = 110101$. Передатчику A нужно послать бит $A_d = 1$, передатчику B — бит $B_d = 0$. В этом примере удобно кодировать бинарный ноль как –1, а бинарную единицу как +1. Тогда к ним можно применять обычные правила сложения и умножения.
- Оба передатчика расширяют свой сигнал, используя ключи в качестве чиповых последовательностей (здесь термин "расширение" означает обычное умножение бита данных на всю чиповую последовательность). В действительности для расширения к отдельным битам данных применяются части намного большей чиповой последовательности. Передатчик A

посылает сигнал $A_s = A_d * A_k = +1*(-1, +1, -1, -1, +1, +1) = (-1, +1, -1, -1, +1, +1)$. Передатчик B для расширения своего сигнала производит те же действия, но со своим кодом: $B_s = B_d * B_k = -1*(+1, +1, -1, +1, -1, +1) = (-1, -1, +1, -1, +1, -1)$.

- Затем оба сигнала одновременно передаются на одинаковой частоте, поэтому в пространстве они накладываются друг на друга. Если интенсивность сигналов при приеме одинакова, а влиянием остальных передатчиков и шумом окружающей среды можно пренебречь, то приемник получит следующий сигнал: $C = A_s + B_s = (-2, 0, 0, -2, +2, 0)$.
- Допустим, что приемнику необходимы данные передатчика A и он настраивается на код A. Иначе говоря, для сужения применяется код передатчика A: $C*A_k = (-2, 0, 0, -2, +2, 0)*(-1, +1, -1, -1, +1, +1) = 2 + 0 + 0 + 2 + 2 + 0 = 6$. Поскольку результат намного превышает 0, приемник детектирует бинарную единицу. При настройке на передатчик B, т.е., при использовании кода B, приемник получит сигнал $C*B_k = (-2, 0, 0, -2, +2, 0)*(+1, +1, -1, +1, -1, +1) = -2 + 0 + 0 - 2 - 2 + 0 = -6$. Результат отрицательный, поэтому детектируется бинарный ноль.

В этом примере было сделано несколько упрощений. Коды были очень простыми, но, по крайней мере, ортогональными. Что более важно, шумы считались пренебрежимо малыми, хотя в действительности шумы могут существенно влиять на передаваемый сигнал C. Соответственно, полученные результаты (–6 и +6) могут оказаться не столь хорошими, а, например, близкими к нулю. В этом случае решить, что представляет собой сигнал — 0 или 1, — будет намного сложнее. Кроме того, интенсивность обоих принимаемых сигналов полагалась одинаковой. Однако, что произойдет, если интенсивность одного сигнала будет намного больше другого? Пусть, например, интенсивность сигнала B в пять раз больше, чем интенсивность сигнала A. Тогда $C' = A_s + 5*B_s = (-1, +1, -1, -1, +1, +1) + (-5, -5, +5, -5, +5, -5) = (-6, -4, +4, -6, +6, -4)$. Допустим, что приемнику нужно получить данные передатчика B: $C'*B_k = -6 - 4 - 4 - 6 - 6 - 4 = -30$. Бинарный ноль, посланный передатчиком B, легко детектируется. При приеме данных передатчика A сигнал будет выглядеть так: $C'*A_k = 6 - 4 - 4 + 6 + 6 - 4 = 6$. Понятно, что по (абсолютной) величине более сильный сигнал превосходит слабый (30 и 6). В то время как –30 детектируется как бинарный ноль, для 6 принять решение не так легко: в сравнении с 30, 6 находится близко к нулю и поэтому может интерпретироваться как шум. Вспомним пример с вечеринкой. Если кто-то разговаривает на своем языке слишком громко, то использовать другой язык в качестве ортогонального кода бессмысленно — все равно вас никто не поймет, а слова лишь увеличат шум. При всей своей крайней упрощенности описанный пример указывает на то, что в системах CDMA необходим контроль мощности передачи. Это одна из крупнейших проблем, возникающих при использовании систем CDMA.

В следующих примерах обсуждается работа схемы CDMA с расширением по схеме DSSS, использующей ортогональные коды. Здесь применяются более длинные коды и ключевые последовательности (т.е. более длинные, чем отдельные биты). Кодовые последовательности в системе IS-95 (мобильная телефонная система, использующая схему CDMA), например, состоят из $2^{42}-1$ чипов. Скорость передачи чипов равна 1228800 чипов/с (т.е. код повторяется через 41,425 суток). Более подробно схема CDMA рассматривается в книге [Viterbi, 1995].

На рис. 3.14 показана передача битовой последовательности 101 с передатчика A. Ключ A приведен в виде сигнала и в виде бинарной ключевой последовательности A_k. После расширения, т.е. после применения операции исключающего ИЛИ к последовательностям A_d и A_k получается сигнал A_s.

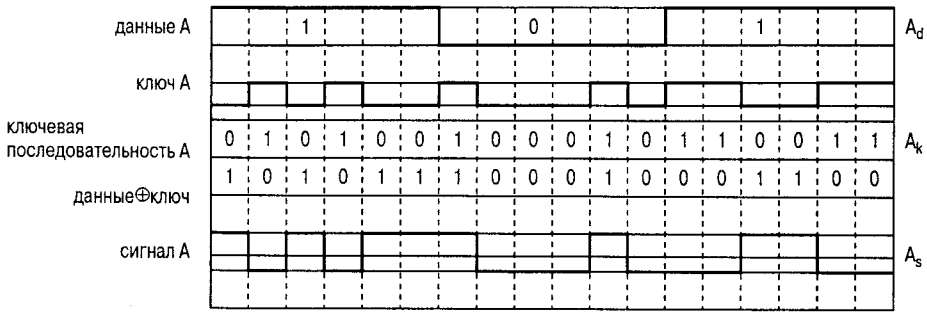

Рис. 3.14. Кодирование и расширение данных в передатчике A

То же самое происходит с данными передатчика B — битовой последовательностью 100. Результатом расширения с использованием кода является сигнал B_s. Сигналы A_s и B_s при передаче накладываются друг на друга (вновь предполагается, что интенсивности сигналов одинаковы, а шумами можно пренебречь). Таким образом, результирующий сигнал — это суммарный сигнал $A_s + B_s$, показанный на рис. 3.15.

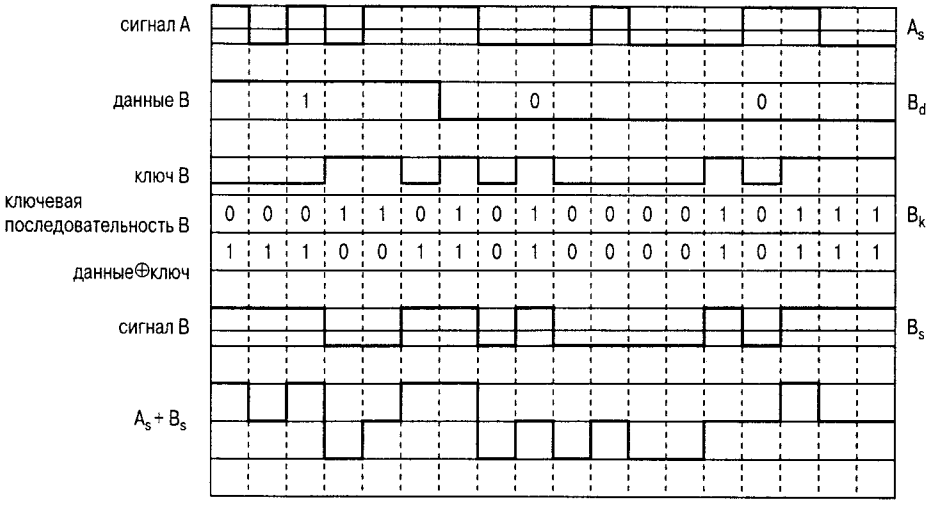

Рис. 3.15. Кодирование и расширение данных в передатчике B

Далее приемником производится попытка восстановления исходных данных передатчика A — сигнала A_d. Для этого приемник применяет ключ A (последовательность A_k) к принимаемому сигналу и подает полученный сигнал на интегратор (см. раздел 2.7.1). В интеграторе все произведения складываются (т.е. вычисляется их скалярное произведение). Затем в компараторе (рис. 3.16) принимается решение, каким данным — 0 или 1 — соответствует результирующий сигнал (здесь он инвертирован по отношению к исходным данным). Заметим, что, несмотря на искажение исходного сигнала сигналом B, результат совершенно очевиден.

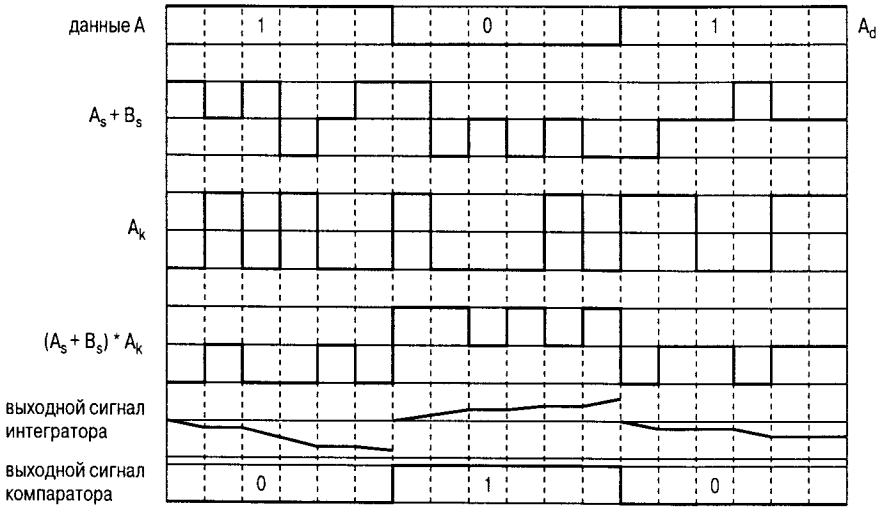

Рис. 3.16. Восстановление исходных данных передатчика A

То же самое происходит, когда приемнику нужно получить данные передатчика B (рис. 3.17). Компаратор легко детектирует исходные данные. Если рассмотреть сигнал $(A_s + B_s) * B_k$, становится понятно, что могло бы случиться, будь сигнал A сильнее. Небольшие пики, вызванные им, могли бы увеличиться и привести к неверному результату интегратора.

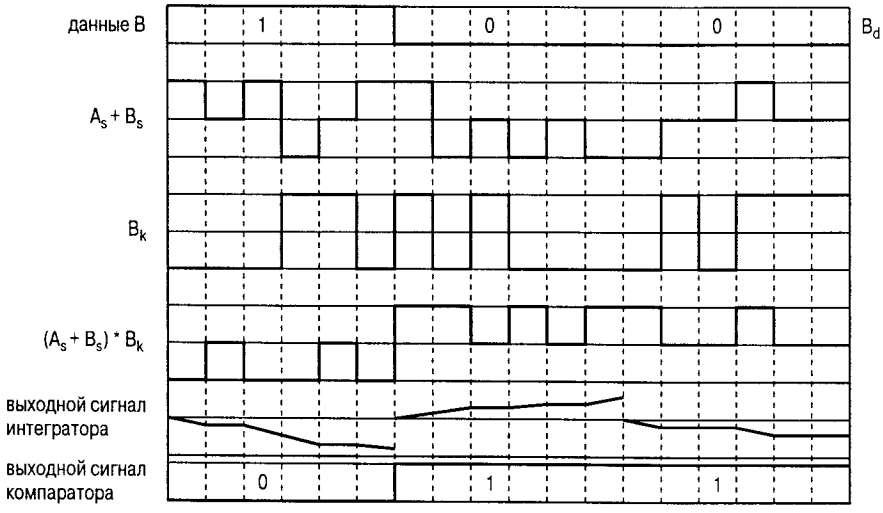

Рис. 3.17. Восстановление исходных данных передатчика B

Наконец, на рис. 3.18 представлена ситуация, когда приемник не синхронизирован с чиповой последовательностью передатчика или когда он использует неправильный ключ. Выходной сигнал интегратора весьма похож на шум, и компаратор не всегда может решить, что он представляет — нуль или единицу. Даже если компаратор совершенно точно детектирует отдельные единицы, передаваемую последовательность битов все же нельзя будет восстановить полностью.

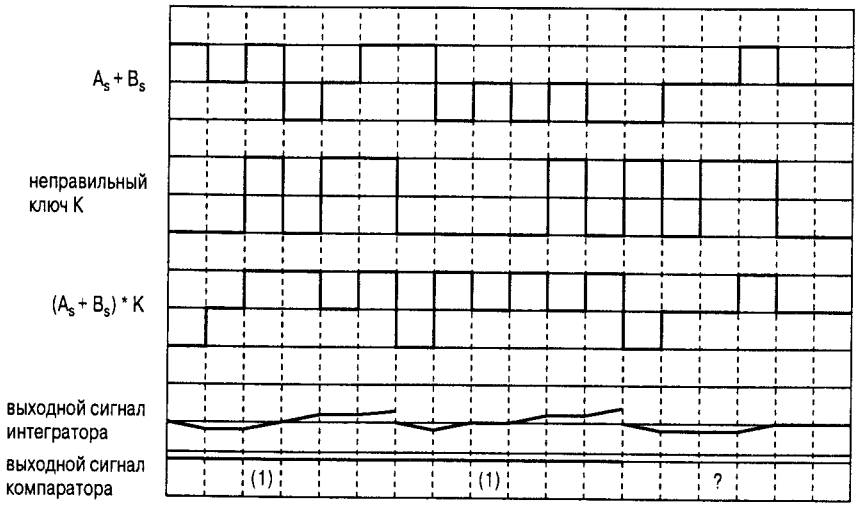

Рис. 3.18. Прием сигнала при использовании неправильного ключа

3.5.1. Множественный доступ с расширением по схеме Aloha

В предыдущих разделах было показано, как использование для расширения сигналов различных кодов с определенными характеристиками позволяет получить хорошую схему организации множественного доступа, получившую название CDMA. Однако, приемники и передатчики в схеме CDMA в действительности оказываются очень непростыми устройствами. Для обеспечения связи с n устройствами приемнику надо уметь декодировать n различных кодов. Эта схема хорошо подходит для различных служб, построенных на установке соединений, например, для мобильных телефонных служб. В то же время она слишком сложна для типичных систем с пульсирующим трафиком передачи данных, работающих без установки соединений. Компьютеры посылают данные лишь время от времени, поэтому использование кодов в этом случае может оказаться излишним. Схема Aloha очень проста, работает без установки соединений, но из-за частых конфликтов обеспечивает сравнительно низкую пропускную способность.

Что произойдет, если объединить механизм расширения по схеме CDMA и с механизмом доступа к среде по схеме Aloha? Другими словами, можно ли использовать схему CDMA лишь с одним кодом, т.е. без кодового разделения? В результате получается схема **множественного доступа с расширением по схеме Aloha** (Spread Aloha Multiple Access, SAMA), представляющая собой комбинацию схем CDMA и TDMA [Abramson, 1996].

Схема SAMA работает следующим образом. Во всех передатчиках для расширения используется один и тот же код (в приведенном на рис. 3.19 примере это код 110101). Стандартная ситуация, возникающая при доступе по схеме Aloha, показана на верхней части рисунка. Если рассмотреть узкополосный спектр, то можно заметить, что передатчики A и B одновременно получают доступ к среде. Поэтому при передаче для всех трех битов возникают конфликты.

Те же данные можно также передать с большей мощностью за более короткие промежутки времени, как показано в центральной части рисунка. Здесь расширение спектра применяется с целью расширения более коротких сигналов, т.е. для увеличения полосы пропускания (в приведенном примере коэффициент расширения $s = 6$). Расширяются оба сигнала, но фаза чиповой последовательности у них слегка отличается. Поэтому, когда один приемник синхронизирован с передатчиком A, а другой — с передатчиком B, разделение данных все еще будет возможно. Если число одновременно работающих передатчиков не превышает $0{,}1 - 0{,}2s$ (эта

цифра зависит также от уровня шума окружающей среды), то вероятность "конфликта" очень низкая [Abramson, 1996]. Главная проблема в этом подходе — найти хорошие чиповые последовательности. Очевидно, что код не может быть ортогонален сам к себе — он должен обладать хорошей автокорреляцией. В то же время корреляция должна быть низкой даже при совсем небольших отклонениях фазы. Максимум пропускной способности в этой схеме, как и в простой схеме Aloha, достигается при 18-процентной загрузке, однако, этот подход одновременно позволяет воспользоваться всеми преимуществами технологии расширенного спектра.

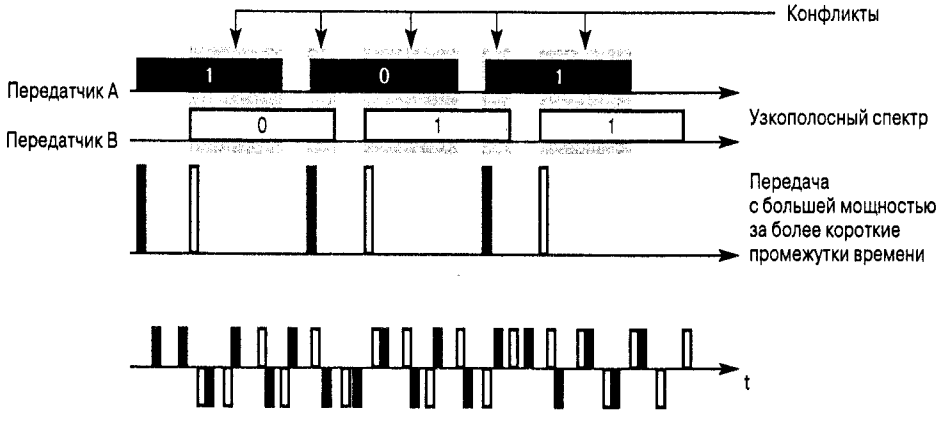

Рис. 3.19. Множественный доступ с расширением по схеме Aloha

3.6. Сравнение схем S/T/F/CDMA

В завершение главы в табл. 3.1 приведены сравнительные данные для четырех основных схем множественного доступа. Здесь описаны отдельные схемы MAC, а не их комбинации. Однако, в реальных системах схемы MAC всегда используются в комбинациях. Одна из типичных комбинаций составляется из схем SDMA/TDMA/FDMA и применяется в телефонных системах IS-54, GSM, DECT, PHS и PACS, а также в спутниковых системах Iridium и ICO. Объединение схем CDMA и SDMA используется в мобильной телефонной системе IS-95 и спутниковой системе Globalstar (см. главы 4 и 5).

Ожидания многих производителей и сетевых провайдеров относительно уровня производительности систем CDMA снизились в сравнении с началом 80-х годов в основном из-за трудностей, имевших место в мобильной телефонной системе IS-95. Тем не менее, согласно предлагаемой европейской системе UMTS схема CDMA будет интегрироваться в будущую мобильную телефонную систему. Схему CDMA можно использовать в комбинации со схемами FDMA/TDMA для увеличения емкости ячейки. В отличие от других, схема CDMA обладает преимуществами простоты переключения и плавно изменяющейся емкости ячеек. Переключение (более подробно оно рассматривается в главе 4) связано с коммутацией между двумя ячейками, т.е., изменением той базовой станции, к которой подсоединена мобильная станция. Простота переключения означает, что мобильная станция может плавно переключаться между соседними ячейками. Это достигается с помощью установки соединения с двумя базовыми станциями одновременно. В схеме CDMA данный процесс реализуется с использованием одинаковых кодов. В системах TDMA/FDMA осуществляется жесткое переключение, т.е. коммутация базовых станций и последовательностей перескоков (времени/частоты) происходит точно в момент переключения. Решение о переключении принимается в зависимости от интенсивности сигнала, поэтому возможны колебания между базовыми станциями.

Таблица 3.1. Сравнение механизмов SDMA, TDMA, FDMA и CDMA

Подход	SDMA	TDMA	FDMA	CDMA
Основная идея	Разбиение пространства на ячейки/сектора	Разбиение времени передачи на непересекающиеся временные интервалы, фиксированные или выделяемые по запросам	Разбиение полосы частот на непересекающиеся подполосы	Расширение спектра с использованием ортогональных кодов
Терминалы	В одной ячейке/секторе может быть активным лишь один терминал	Все терминалы используют одинаковую частоту и активны в течение коротких промежутков времени	Каждый терминал имеет собственную частоту, его работа не влияет на другие терминалы	Все терминалы могут непрерывно работать в одно и то же время в одном месте, не оказывая влияния друг на друга
Разделение сигналов	Структура из отдельных ячеек и направленные антенны	Синхронизация во временной области	Фильтрование частот	Коды и применение специальных приемников
Преимущества	Схема очень простая, возможно увеличение числа пользователей на единице площади	Схема устоявшаяся, полностью цифровая, очень гибкая	Схема простая, устоявшаяся, устойчивая к помехам	Гибкая схема, не требует тщательного планирования, упрощенное переключение базовых станций
Недостатки	Схема недостаточно гибкая, антенны обычно фиксированные	Необходимы защитные интервалы (многолучевое распространение), трудно добиться синхронизации	Схема недостаточно гибкая, ресурсы частот ограничены	Сложные приемники, необходим контроль мощности передатчиков
Замечания	Эта схема полезна лишь в комбинации со схемами TDMA, FDMA или CDMA	Стандартная схема для стационарных сетей, в комбинации со схемами FDMA/SDMA применяется во многих мобильных сетях	Объединяется со схемами TDMA (схема перескоков частот) и SDMA (повторное использование частоты)	Сталкивается с некоторыми проблемами: высокая сложность, снижение ожидаемых показателей; будет интегрироваться со схемами TDMA/FDMA

Гибкая емкость систем CDMA означает, что в ячейку можно добавлять все больше и больше пользователей, т.е. жесткого предела количества присутствующих в ней пользователей не существует. В системах TDMA/FDMA такой жесткий предел есть — если в системе больше не остается свободных временных или частотных интервалов, то подключение в ячейку новых пользователей не допускается. При добавлении нового пользователя в ячейку системы CDMA ячейка уменьшается, и в ней повышается уровень шума, но пользователь все равно может связываться с остальными станциями. Однако, сокращение размеров ячейки может вызвать проблемы, поскольку в результате некоторые

пользователи могут "выпасть" из ячейки. Поэтому планировать размеры ячеек в системах CDMA сложнее, чем в более жестких схемах TDMA/FDMA.

Мобильные телефонные системы, использующие комбинации схем SDMA/TDMA/FDMA или SDMA/CDMA, являются централизованными — базовая станция управляет многими мобильными станциями. Однако, в произвольных беспроводных коммуникационных системах необходимо использовать другие алгоритмы доступа к среде. В предыдущих разделах было показано, что в большинстве существующих распределенных систем применяется несколько версий базовой схемы Aloha. Для обеспечения фиксированной задержки доступа и требуемой пропускной способности используются схемы Aloha с разделением и механизмы резервирования. Каждая из существующих схем обладает своими преимуществами и недостатками. Схема CSMA очень эффективна при слабой загрузке, с помощью схемы MACA можно преодолеть трудности со скрытыми и открытыми терминалами, схема с опросом обеспечивает фиксированную пропускную способность. Отдельной схемы, которая имела бы все эти преимущества, не существует. Именно этот факт послужил причиной того, что стандарт IEEE 802.11 объединяет все три схемы (см. раздел 7.3). Опрос применяется для установления временной структуры через базовую станцию. Схема CSMA используется для доступа во время некоординируемых периодов, а схема MACA применяется для решения проблемы скрытых терминалов и при отсутствии базовых станций.

3.7. Упражнения

1. Что является основной физической причиной недееспособности многих схем MAC, разработанных для проводных сетей, при использовании их в беспроводной среде? Что делается в проводных сетях для избежания этого эффекта?
2. Вспомните проблему скрытых и открытых терминалов. Что происходит с такими терминалами при использовании схем Aloha, Aloha с разделением, Aloha с резервированием, MACA?
3. Как влияют эффекты приближенности/удаленности на системы TDMA? Что происходит в этом случае в системах CDMA? Какие контрмеры принимаются в системах TDMA и CDMA?
4. Как реализуется алгоритм MAC в схеме SDMA? Какие роли могут играть мобильные станции, базовые станции, планирование сетевого провайдера?
5. Что является основной предпосылкой для применения схемы FDMA? Почему этот фактор увеличивает сложность систем FDMA в сравнении с системами TDMA? Какие схемы управления доступом к среде применяются при рассмотрении всего частотного пространства (представленного в главе 1)?
6. Какие альтернативные варианты дуплексных каналов можно реализовать в беспроводных сетях? А в обычных проводных сетях?
7. Какие преимущества есть у схемы фиксированного TDM в сравнении со схемой случайного TDM, управляемого по запросам? Сравните эффективность этих схем в случае нескольких соединений с фиксированными и изменяющимися потоками данных. Объясните, почему в обычных мобильных телефонных системах используются фиксированные схемы, а в компьютерных сетях чаще применяются случайные схемы. В будущем большая часть данных будет создаваться с помощью компьютеров. Как этот факт повлияет на мобильные телефонные системы?
8. Объясните возникновение помех в пространственной, временной и частотной областях, а также в кодовом пространстве. Какие контрмеры принимаются в системах SDMA, TDMA, FDMA и CDMA?

9. Представьте, что каждая станция может принимать любую из остальных станций. Пусть некоторой станции необходимо передать данные. Она детектирует свободную несущую частоту. Почему после начала передачи все равно может возникнуть конфликт?
10. Какими преимуществами обладают схемы с резервированием? Как в этих схемах можно избежать конфликтов при передаче, почему вероятность возникновения конфликтов здесь ниже, чем в классической схеме Aloha? Назовите недостатки схем с резервированием.
11. Почему схема MACA может не работать при наличии скрытых/открытых терминалов? При рассмотрении учитывайте не только наличие мобильных станций, но и изменение характеристик передачи.
12. Какая из схем MAC предоставляет твердые гарантии относительно общей пропускной способности системы и фиксированной задержки доступа?
13. Каким образом создаются защитные интервалы между отдельными пользователями в системе CDMA?
14. Измените пример упрощенной схемы CDMA из раздела 3.5, добавив к переданному сигналу (–2, 0, 0, –2, +2, 0) случайный "шум". Добавьте, например, шумовую последовательность (1, –1, 0, 1, 0, –1). Что в этом случае будет детектировать приемник при приеме сигналов передатчиков A и B? При рассмотрении дополнительно учитывайте вопросы приближенности/удаленности станций. Как это усложняет ситуацию? Что можно предпринять в этом случае?

3.8. Рекомендуемая литература

Abramson, N. [1996] *Wideband random access for the last mile*, IEEE Personal Communications, 3 (6).

Abramson, N. [1997] *The throughout of packet broadcasting channels*, IEEE Transactions on Communication, COM-25, (1).

Halsall, F. [1996] *Data communications, computer networks and open systems*, Addison Wesley Longman.

Kleinrock, L.; Tobagi, F. [1975] *Packet switching in radio channels: part 1 — carrier sense multiple-access modes and their throughput-delay characteristics*, IEEE Transactions on Communications, COM-23, (12).

Viterbi, A. [1995] *CDMA: principles of spread spectrum communication*, Addison Wesley Longman.

Системы телекоммуникаций 4

Цифровые сотовые сети — наиболее быстро развивающаяся часть рынка мобильных и беспроводных устройств. Они являются расширениями традиционных сетей PSTN и ISDN и обеспечивают общегосударственный, а иногда и всемирный роуминг с одного и того же мобильного телефона. В настоящее время эти системы в основном применяются для передачи речевой информации. Однако, объемы обмениваемой информации непрерывно увеличиваются, поэтому в данной главе рассмотрено несколько технологий беспроводной передачи данных в системах сотовой связи.[1]

Представленные системы берут свое начало не от компьютерных сетей. Их архитектура — это архитектура традиционной телефонной связи. Поэтому их основные разновидности обычно реализуются как коммутационные службы с акцентом на передаче речи и скоростью передачи данных не более, например, 9,6 Кбит/с. Однако, службы поддерживают скорость до 250 км/ч (к примеру, услугами сети GSM можно пользоваться в движущемся автомобиле), когда большинство беспроводных систем перестает действовать.

Всемирный рынок сотовых сетей выглядит следующим образом. Система GSM является самой распространенной цифровой системой и занимает приблизительно 40% рынка. (Эта система представлена в разделе 4.1). Аналоговая система AMPS все еще удерживает порядка 30% рынка, японская система PDC — 15%. Остаток разделен между системами на основе CDMA, TDMA и другими технологиями. В **Европе** большинство людей (свыше 60 миллионов) используют цифровую систему GSM, хотя некоторые все еще пользуются аналоговыми системами (меньше 10 миллионов). Почти все продаваемые сейчас телефоны — цифровые. В **США** и **Канаде** ситуация несколько отличается. Здесь более 80% людей используют аналоговую систему AMPS, которая в 1998 году имела больше 60 миллионов абонентов. Рынок цифровых технологий разделен между системами на основе TDMA (5 миллионов абонентов), CDMA (3,3 миллиона) и GSM (2 миллиона). В то время как в Европе действует одна цифровая система, в США их существует несколько. Это приводит к серьезным проблемам, связанным с покрытием и доступностью услуг, и является одним из немногих случаев, когда рыночные механизмы не привели к улучшению услуг (в сравнении с общим европейским стандартом).

На рис. 4.1 приводится количество абонентов различных технологий мобильной телефонной связи во всем мире [GSMData, 1998]. Различные версии одной и той же технологии здесь объединены (например, системы GSM, работающие на частотах 900, 1800 и 1900 МГц). Прогнозы на 1999–2001 годы необходимо рассматривать с осторожностью.

В этой главе в качестве главного примера полностью цифровой системы мобильной телефонной связи используется система GSM. Причина — не только успех на рынке, но и ее архи-

---

[1] Все системы, представленные здесь, являются цифровыми. Читателю, интересующемуся старыми аналоговыми системами (например, US AMPS), можно порекомендовать хорошую книгу [Goodman, 1997].

тектура, которая позже легла в основу многих других систем. Управление мобильностью, мобильное автоматическое переключение и еще несколько основоположных идей других технологий заимствованы именно у системы GSM [Goodman, 1997]. Кроме того, в спецификациях американских систем внимание обычно сосредоточено на радиоинтерфейсах. В отличие от них в спецификации системы GSM определены различные открытые интерфейсы и сетевые объекты. Первый подход позволяет компаниям разрабатывать собственные решения, а второй предоставляет компаниям-поставщикам сетевых услуг возможность выбора между многими продуктами различных производителей.

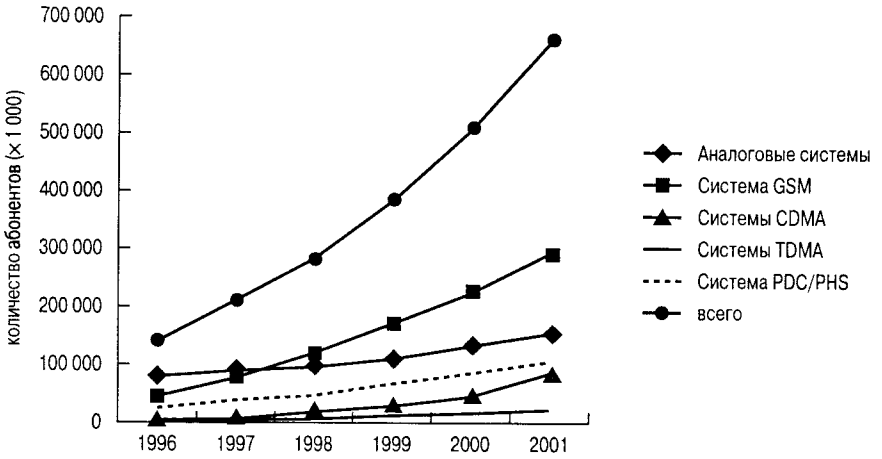

Рис. 4.1. Количество абонентов различных технологий мобильной телефонной связи во всем мире

После описания системы GSM обсуждаются DECT — цифровой стандарт для беспроводных телефонов, и TETRA — цифровые радиомагистрали. Основное внимание уделяется службам передачи данных. В связи с этим рассматривается эволюция системы GSM в направлении передачи данных с более высокой скоростью, пакетно-ориентированная передача и, наконец, система UMTS.

4.1. Система GSM

Система GSM является самой успешной современной цифровой системой мобильных телекоммуникаций — ее используют свыше 100 миллионов человек более чем в 130 странах мира. В начале 80-х годов в Европе возникла проблема сосуществования многих аналоговых систем мобильной телефонной связи. Часто они были основаны на похожих стандартах (например, на стандарте NMT 450), но их несущие частоты слегка отличались. Чтобы избежать подобной ситуации с полностью цифровой системой второго поколения, в 1982 году была основана Специальная группа мобильной связи (**Groupe Spéciale Mobile — GSM**). В настоящее время система, разработанная этой группой, называется **Глобальной системой мобильной связи (Global System for Mobile Communications — GSM)**. Спецификации этой системы разрабатываются институтом ETSI [ETSI, 1999], [GSM MoU, 1999].

Главной целью группы GSM было создание системы мобильной телефонной связи, службы передачи речи которой были бы совместимыми с сетью ISDN и другими системами PSTN. Кроме того, будущая система должна была обеспечивать роуминг по всей Европе. Спецификация исходной системы заняла более 5 000 страниц; описание новых служб, в особенности служб передачи данных, содержит еще больше подробностей. Читатели, знакомые с эталонной

моделью ISDN, найдут здесь много похожих аббревиатур, опорных точек и интерфейсов. При стандартизации системы GSM стремились достичь максимальной совместимости.

Итак, GSM — это типичная система 2-го поколения, заменяющая аналоговые системы 1-го поколения. Однако, она не обеспечивает высокоскоростной передачи данных по всему миру, ожидаемой от систем 3-го поколения, в частности, от системы UMTS. Вначале в системе GSM для исходящих линий связи использовались частоты 890–915 МГц, а для нисходящих линий — диапазон 935–960 МГц. Такую систему сейчас иногда называют **GSM 900**, чтобы отличать ее от более поздних версий. Эти версии включают системы GSM, работающие на частотах 1800 МГц (1710–1785 МГц для исходящих, 1805–1880 МГц — для нисходящих каналов) и 1900 МГц (1850–1910 МГц для исходящих каналов и 1930–1990 МГц — для нисходящих). Первую систему еще называют **DCS 1800 (цифровой сотовой системой)**, вторую — **PCS 1900 (персональной службой связи)**.

В следующем разделе описана структура, а также службы и протоколы системы GSM, которые являются общими для всех трех решений — **GSM 900**, **GSM 1800** и **GSM 1900**. Хотя система GSM создавалась и используется, главным образом, для передачи речи, можно предвидеть, что многие будущие приложения мобильной связи будут управляться данными. Поэтому отношение информационного обмена к речевому будет постоянно расти.

4.1.1. Мобильные службы

Система GSM объединяет различные службы передачи речи и данных, а также взаимодействует с другими существующими сетями. Интерес к сети обусловлен предоставляемыми ею услугами. В системе GSM определено три категории служб: службы-носители, телекоммуникационные и дополнительные службы. Они описаны в следующих разделах. На рис. 4.2 приведена эталонная модель служб системы GSM. **Мобильная станция MS** соединена с **мобильной наземной сетью GSM общего пользования (public land mobile network — PLMN)** через интерфейс U_m. (Сеть GSM-PLMN является инфраструктурой, необходимой для сети GSM). Эта сеть связана с транзитными сетями, например, с **цифровой сетью с интеграцией услуг (integrated services digital network — ISDN)** или с обычной **коммутируемой телефонной сетью общего пользования (public switched telephone network — PSTN)**. Перед соединением с **терминалом TE** может существовать еще одна сеть — сеть источник/адресат. **Службы-носители** обеспечивают сквозную передачу данных между интерфейсами сети, т.е. интерфейсом S мобильной станции и подобным интерфейсом другого терминала (например, интерфейсом S_0 для терминалов сети ISDN). Такие интерфейсы сети ISDN, как U, S и R, определены не для всех сетей. Поэтому эталонные интерфейсы для сквозной передачи данных в разных сетях могут быть различными. В классической модели системы GSM службы-носители действуют на основе соединений и используют коммутацию каналов или пакетов. Для этих служб необходимы лишь первые три уровня эталонной модели ISO/OSI.

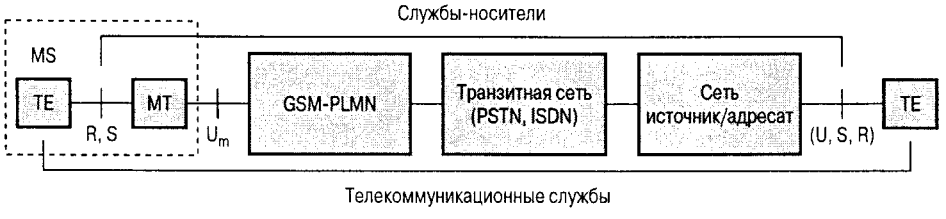

Рис. 4.2. Эталонная модель служб-носителей и телекоммуникационных служб

Внутри мобильной станции все специфические сетевые функции (доступ по схемам TDMA или FDMA, кодирование и т.п.) выполняются **мобильным оконечным устройством (MT)**. Оно

также предоставляет интерфейс для передачи данных (S) на терминал TE. Этот интерфейс может быть независимым от сети. В зависимости от характеристик терминала TE могут потребоваться и другие интерфейсы, например, интерфейс R (для эталонной модели сети ISDN) [Halsall, 1996]. **Телекоммуникационные службы** зависят от приложений, поэтому они могут использовать все семь уровней эталонной модели ISO/OSI. Эти службы являются сквозными, т.е. передача осуществляется с одного терминала TE на другой.

4.1.1.1. Службы-носители

В технологии GSM определены различные механизмы передачи данных. В исходной системе GSM услуги по передаче данных характеризовались скоростью до 9600 бит/с. Службы-носители допускают сквозную, несквозную, синхронную и асинхронную передачу. В **сквозных службах-носителях** используется только физический уровень (уровень 1). Для этого случая, если при передаче не происходит ошибок, характерна фиксированная задержка передачи и пропускная способность. Единственным механизмом улучшения качества передачи является **прямое исправление ошибок (forward error correction — FEC)**. При использовании этой схемы в поток данных код вносит избыточную информацию, которая при возникновении ошибок передачи используется для восстановления исходных данных. В зависимости от схемы FEC скорость передачи данных может составлять 2,4 Кбит/с, 4,8 Кбит/с или 9,6 Кбит/с. При использовании сквозных служб-носителей данные, потерянные в результате экранирования или переключения базовой станции, не восстанавливаются.

Несквозные службы-носители для коррекции ошибок и управления потоком используют протоколы второго и третьего уровней. Такие службы объединяют сквозные службы-носители с **протоколом радиоканала (radio link protocol — RLP)**. Он включает все механизмы **высокоуровневого протокола управления каналом (high-level data link control — HDLC)** [Halsall, 1996] и специальные механизмы выборочного отказа, необходимые для повторной передачи ошибочных данных. При этом можно добиться, чтобы частота появления ошибок не превышала 10^{-7}, однако, пропускная способность и задержки могут изменяться в зависимости от качества передачи.

В системе GSM с помощью сквозных и несквозных служб определено несколько служб-носителей для взаимодействия с сетями PSTN, ISDN и PSPDN (в частности, сетями X.25), которые доступны во всем мире. Скорость синхронной полнодуплексной передачи данных может составлять 1,2 Кбит/с, 2,4 Кбит/с, 4,8 Кбит/с или 9,6 Кбит/с. Для асинхронной полнодуплексной передачи она находится в пределах 300–9 600 бит/с [ETSI, 1991a]. Столь низкая скорость отражает тот факт, что услуги по передаче данных составляют лишь малую часть общего информационного обмена. В настоящее время это действительно так, но соотношение между службами передачи данных и речи в сетях GSM изменяется, причем все более важными становятся именно данные. Эта тенденция отражена в новых службах передачи данных (раздел 4.1.8).

4.1.1.2. Телекоммуникационные службы

Телекоммуникационные службы системы GSM в основном ориентированы на передачу речи. Они включают шифрованную голосовую передачу, службы сообщений и элементарный обмен данными с терминалами, известными по сетям PSTN и ISDN (например, факс). Главной службой, однако, является **телефония**. Основной задачей системы GSM было обеспечение высококачественной цифровой передачи речи с полосой пропускания, по крайней мере, 3,1 кГц (типичное значение для аналоговых телефонных систем). Для передачи речи используются специальные кодеки (кодер/декодер). При передаче аналоговых данных на компьютерные модемы (использующиеся, в частности, в факсах) применяются другие кодеки.

Аварийный номер — одна из услуг, предлагаемых системой GSM. Во всей Европе используется один и тот же номер. Эта услуга бесплатна и обязательна для всех поставщиков услуг. Та-

кое соединение автоматически устанавливается с ближайшим аварийным центром и имеет наибольший приоритет — возможно даже прерывание других соединений.

Служба коротких сообщений (short message service — SMS) обеспечивает передачу сообщений длиной меньше 160 символов. Для передачи сообщений SMS применяются не стандартные каналы данных системы GSM, а свободные ресурсы сигнальных каналов (раздел 4.1.3.1). Поэтому посылать и получать сообщения SMS можно даже во время передачи речи или данных. В настоящее время служба SMS обычно применяется для отображения состояния дороги, заголовков электронных писем, котировок ценных бумаг и т.п.

Еще одной неречевой телекоммуникационной службой, доступной во всем мире, является **факс группы 3**. В этой службе по аналоговой телефонной сети с помощью модемов как цифровые данные передаются факсимильные сообщения. Передача происходит согласно стандартам T.4 и T.30 института ITU-T. В большинстве случаев используется сквозная служба факсов, т.е. факсимильные данные и сигналы факса передаются с помощью сквозной службы-носителя. При низком качестве передачи служба-носитель автоматически адаптируется к низкой скорости и большому количеству избыточной информации, необходимой для схемы FEC.

4.1.1.3. Дополнительные службы

Кроме телекоммуникационных служб и служб-носителей, поставщики услуг GSM могут предлагать и **дополнительные службы**. Как и в сетях ISDN, они являются различными усовершенствованиями обычной службы телефонии. Такие службы у различных поставщиков могут быть разными. Обычно это **идентификация** пользователя, а также **переадресация** и **перенаправление** звонков. К тому же, могут быть доступными стандартные возможности сети ISDN, например, **закрытые группы пользователей** или **многосторонние** соединения. Закрытые группы пользователей представляют особый интерес, поскольку с их помощью можно создать подсеть GSM для отдельной фирмы, причем доступ к такой подсети будут иметь лишь члены группы.

4.1.2. Архитектура системы

Система GSM, как и все телекоммуникационные системы, имеет сложную иерархическую структуру. Она включает много объектов, интерфейсов и аббревиатур. На рис. 4.3 приводится упрощенная схема функционирования системы GSM, установленная стандартом [ETSI, 1991б]. Система GSM состоит из трех подсистем: **подсистемы радиосвязи (radio subsystem — RSS)**, **подсистемы сетей и коммутации (network and switching subsystem — NSS)**, а также **операционной подсистемы (operation subsystem — OSS)**. Каждая из подсистем подробно рассматривается в следующих разделах. Как правило, абоненты системы GSM видят лишь малую часть общей сети — мобильные станции (mobile station — MS) и несколько антенных вышек базовых трансиверов (base transceiver — BTS).

4.1.2.1. Подсистема радиосвязи

Как видно из самого названия, **подсистема радиосвязи** включает все радиообъекты, т.е. **мобильные станции (MS)** и **подсистему базовых станций (base station subsystem — BSS)**. На рис. 4.3 показана связь между подсистемами RSS и NSS через **интерфейс A** (сплошные линии) и связь с подсистемой OSS через **интерфейс O** (пунктирные линии). В основу интерфейса A положены системы PCM-30 с коммутацией каналов (2,048 Мбит/с), поддерживающие до 30 соединений с пропускной способностью 64 Кбит/с. В интерфейсе O используется Сигнальная система №7 (Signaling System No 7 — SS7). Она основана на передаче управляющей информации через сеть X.25 к подсистеме RSS и обратно.

- **Подсистема базовых станций (BSS)**. Сеть GSM содержит много подсистем BSS. Каждая из них управляется контроллером базовых станций (base station controller — BSC). Подсистемы

BSS отвечают за поддержку радиосоединений с мобильными станциями, кодирование/декодирование речи, а также за регулирование скорости передачи через беспроводную сеть. Кроме контроллера BSC, подсистема BSS включает несколько станций BTS.

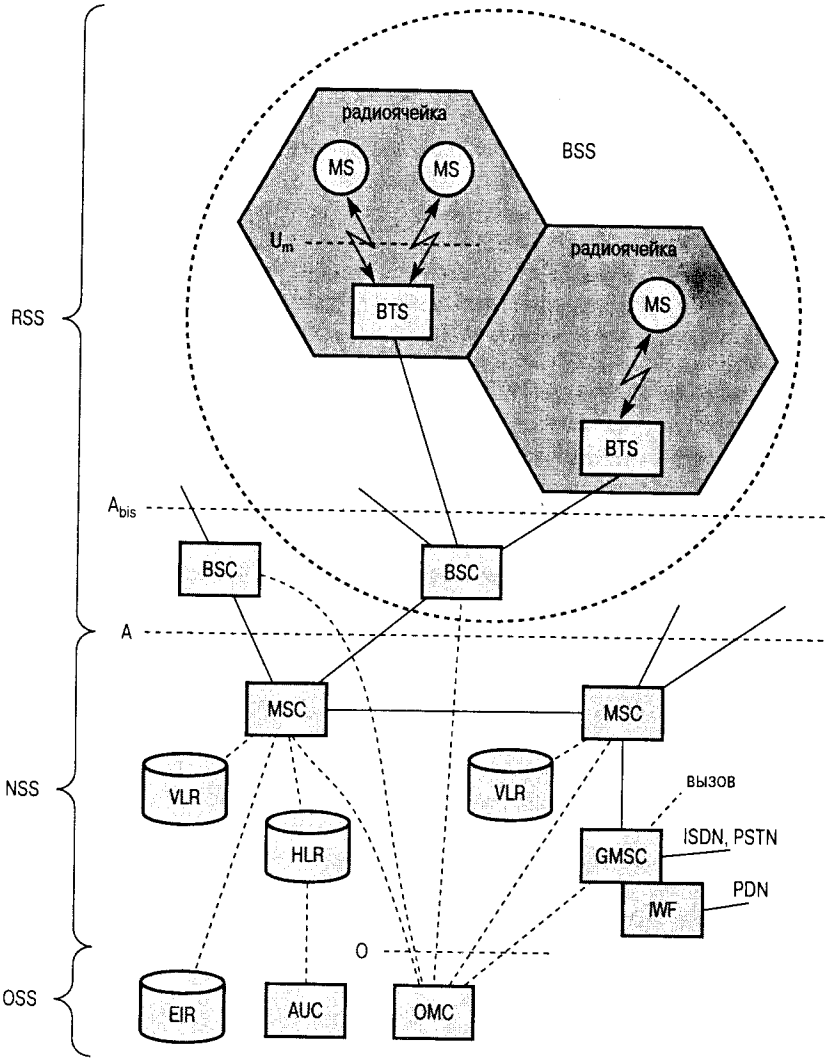

Рис. 4.3. Функциональная архитектура системы GSM

- **Базовая трансиверная станция (BTS)**. В состав станции BTS входит все оборудование для радиосвязи, т.е. антенны, устройства обработки сигналов и усилители, необходимые для радиопередачи. Станция BTS может создавать одну или несколько (при использовании секторных антенн, см. раздел 2.8) ячеек. Она соединена с мобильной станцией через интерфейс U_m (интерфейс U сети ISDN, приспособленный к мобильному окружению) и с контроллером BSC — через интерфейс A_{bis}. Интерфейс U_m включает все механизмы, необходимые для беспроводной передачи (схемы TDMA, FDMA и т.п.) и обсуждается ниже. Интерфейс A_{bis} состоит из соединений с пропускной способностью 16 или 64 Кбит/с. Размер

ячейки GSM может изменяться от нескольких сотен метров до 35 километров в зависимости от окружения (зданий, открытого пространства, гор и т.д.) и ожидаемой интенсивности информационного обмена.

- **Контроллер базовых станций (BSC)**. Контроллер BSC управляет станциями BTS. С его помощью происходит резервирование частот и переключение с одной станции BTS на другую, осуществляется поиск мобильных станций. Кроме того, контроллер BSC уплотняет радиоканалы для соединений со стационарными сетями через интерфейс A.

В табл. 4.1 приводится обзор задач контроллера BSC и станции BTS, а также функций, в которых они поддерживают другие объекты сети.

Таблица 4.1. Функции станции BTS и контроллера BSC внутри подсистемы BSS

Функция	BTS	BSC
Управление радиоканалами		x
Перескок частоты	x	x
Управление наземными каналами		x
Соединение наземных каналов с радиоканалами		x
Кодирование и декодирование каналов	x	
Адаптация скорости передачи	x	
Шифрование и дешифрование	x	x
Поиск мобильных станций	x	x
Измерение исходящего сигнала	x	
Измерение информационного обмена		x
Аутентификация		x
Регистрация и обновление данных о местонахождении		x
Управление переключением базовых станций		x

- **Мобильная станция (MS)**. В состав мобильной станции входят все пользовательское оборудование и программное обеспечение, необходимое для соединения с сетью GSM. Сюда же можно отнести независимое от пользователя аппаратное и программное обеспечение, а также **модуль идентификации абонента (subscriber identity module — SIM)**, в котором хранятся все индивидуальные данные. Мобильную станцию можно идентифицировать с помощью системы **международной идентификации мобильного оборудования (international mobile equipment identity — IMEI)**. В то же время пользователь может персонализировать любую мобильную станцию, используя свой модуль SIM. Иначе говоря, индивидуальные пользовательские механизмы, например, оплата и идентификация, зависят от модуля SIM, а не от самого устройства. В механизмах, зависящих от устройства (защита от краж), применяется схема опознавания оборудования (EI). Без модуля SIM возможны звонки лишь по аварийным номерам. Типичные мобильные станции системы GSM 900 характеризуются мощностью передачи 2 Вт. Для системы GSM 1800, благодаря меньшим размерам ячеек, достаточно мощности 1 Вт. Кроме телефонного интерфейса, мобильная станция может предоставлять и другие типы интерфейсов пользователям с дисплеями, громкоговорителями, микрофонами или программируемыми клавишами. Можно также упомянуть интерфейсы, включающие компьютерные модемы и технологии IrDA или Bluetooth. В SIM-карте содержится много идентификаторов и таблиц, например, тип карты, регистрационный номер, список абонентских услуг, **личный идентификационный номер (PIN)**, **ключ снятия блокировки PIN (PUK)**, **ключ аутентификации** K_i и **международная идентификация мобильного абонента (IMSI)** [ETSI, 1991в]. Номер PIN используется для разблокировки мобильной станции. После трехразового неправильного набора номера PIN модуль SIM блокируется. В таких случаях для его разблокирования нужен ключ PUK. Кроме того, при входе в систему GSM мобильная станция сохраняет динамическую

информацию, в частности, **ключ шифра K_c** и данные о местонахождении, состоящие из **временной идентификации мобильного абонента (TMSI)** и **идентификации положения (LAI)**. Типичные мобильные станции (например, мобильные телефоны) могут выполнять различные дополнительные функции, зависящие от производителя оборудования. Такими функциями могут быть использование отпечатков пальцев вместо номера PIN, наличие календаря или адресной книги, а также простые игры.

4.1.2.2. Подсистема сетей и коммутации

Подсистема сетей и коммутации (NSS) — ядро системы GSM. Она соединяет беспроводную сеть со стандартными сетями общего пользования и осуществляет переключение подсистем BSS. Подсистема NSS определяет местонахождение пользователей в любой точке земного шара и обеспечивает оплату, ведение счетов и роуминг пользователей между различными поставщиками услуг связи в разных странах. В ее состав входят следующие коммутаторы и базы данных.

- **Центр коммутации мобильных служб (mobile services switching center — MSC)**. Центры MSC являются высокопроизводительными цифровыми коммутаторами ISDN. Они устанавливают соединения с другими центрами MSC и контроллерами BSC через интерфейс A. Следовательно, центры MSC образуют стационарную магистральную сеть системы GSM. В большинстве случаев центр MSC управляет несколькими станциями BSC, находящимися в какой-то географической области. **Шлюз MSC (GMSC)** дополнительно соединен с другими стационарными сетями, в частности, с сетями **PSTN** и **ISDN**. При использовании дополнительных **межсетевых функций (IWF)** центр MSC может также подключаться к **сетям передачи данных общего пользования (PDN)**, например, к сети X.25. Центр MSC управляет всеми сигналами, нужными для установки соединения, разъединения и переключения соединений на другие центры MSC. В этих целях используется стандартная **Сигнальная система №7 (SS7)**. Кроме того, центр MSC выполняет все функции, необходимые для дополнительных служб, — перенаправления звонков, многосторонних соединений, обратной оплаты и т.п.
- **Регистр исходного положения (HLR)**. Регистр HLR — самая важная база данных системы GSM. В ней хранится вся информация о пользователях. Статическая информация включает **номер ISDN мобильного абонента (MSISDN)**, оплаченные услуги и ключ аутентификации K_i. Кроме того, нужна и динамическая информация, например, данные о текущей **зоне местонахождения (LA)** мобильной станции. Как только мобильная станция покидает текущую зону LA, данные в регистре HLR обновляются. Такая информация необходима для локализации пользователя во всемирных сетях GSM. Все эти элементы индивидуальной информации находятся в отдельных регистрах HLR для каждого пользователя. Они также нужны для взимания оплаты и ведения счетов.
- **Регистр местонахождения посетителей (VLR)**. Регистр VLR, связанный с каждым центром MSC, является динамической базой данных. В ней хранится вся важная информация, необходимая пользователям мобильных станций, находящихся внутри зоны LA центра MSC. Если в зоне LA, за которую отвечает регистр VLR, появляется новая мобильная станция, в регистр VLR из регистра HLR копируется вся информация об этом пользователе. Такая иерархия регистров VLR и HLR позволяет избежать частых обновлений HLR и передачи данных о пользователе на большие расстояния. Применение регистров HLR и VLR для локализации пользователей описано в разделе 4.1.5.

4.1.2.3. Операционная подсистема

Третья часть системы GSM — **операционная подсистема (OSS)** — выполняет все функции, необходимые для эксплуатации и обслуживания сети. Подсистема OSS обладает и собственными

сетевыми объектами, а доступ к другим получает через сигналы SS7 (см. рис. 4.3). В подсистеме определены следующие объекты.

- **Центр эксплуатации и обслуживания (OMC)**. Центр OMC отвечает за управление и контроль всех объектов сети через интерфейс O (для сетей X.25 — через систему SS7). К типичным функциям управления центра OMC можно отнести наблюдение за информационным обменом, управление безопасностью и абонентами, а также ведение счетов. Центры OMC используют понятие **сети управления телекоммуникациями (TMN)**, стандартизованной институтом ITU-T.
- **Центр аутентификации (AuC)**. Поскольку радиоинтерфейс и мобильные станции весьма уязвимы, для обеспечения конфиденциальности пользователей и защиты передачи данных был определен отдельный центр AuC. Он содержит алгоритмы аутентификации и ключи шифров. Кроме того, он генерирует значения, необходимые для аутентификации пользователя в регистре HLR. В действительности центр AuC может находиться в специальной защищенной части регистра HLR.
- **Регистр опознавания оборудования (EIR)**. Регистр EIR является базой данных для опознавания IMEI. Иначе говоря, в нем хранится идентификация всех устройств, зарегистрированных для данной сети. Поскольку станции портативны, их легко украсть. Любой человек, у которого есть действующий модуль SIM, может использовать украденную мобильную станцию. Поэтому в регистре EIR содержится "черный список" украденных (или блокированных) устройств. Таким образом, как только владелец заявляет о краже, мобильная станция становится бесполезной. В регистре EIR также содержатся списки действующего и неисправного оборудования.

4.1.3. Радиоинтерфейс

Самым интересным интерфейсом системы GSM является радиоинтерфейс U_m. В него входят многие механизмы уплотнения и доступа к среде, описанные в главах 2 и 3. В системе GSM схема SDMA реализуется с помощью ячеек и станций BTS. Каждой мобильной станции назначается станция BTS. Кроме того, для разделения исходящей и нисходящей линий связи используется дуплекс FDD. Этот механизм показан на рис. 4.4 (и на рис. 3.3). Для доступа к среде применяется объединение схем TDMA и FDMA. В системе GSM 900 схема FDMA использует 124 канала шириной по 200 кГц, в системе GSM 1800 — 374 канала. В следующем примере в качестве основы использована система GSM 900, однако, система GSM действует подобным образом и на частотах 1800 и 1900 МГц.

На рис. 3.3 главы 3 уже было проиллюстрировано применение частотного уплотнения в системе GSM. На рис. 4.4 дополнительно приводится схема использования временного уплотнения. Каждый из 248 каналов дополнительно разделяется по времени на **кадры TDMA**, т.е. каждая несущая шириной 200 кГц делится на кадры, непрерывно следующие один за другим. Продолжительность кадра составляет 4,615 мс. Кадр подразделяется на 8 **временных интервалов GSM**, причем каждый интервал представляет физический канал TDM и длится 577 мкс. Таким образом, каждый канал TDM каждые 4,615 мс занимает несущую шириной 200 кГц на 577 мкс.

Данные передаются небольшими порциями — **пакетами**. На рис. 4.4 показан так называемый **нормальный пакет**, который применяется для передачи данных внутри временного интервала (пользовательские и сигнальные данные). На диаграмме приводится пакет продолжительностью 546,5 мкс и емкостью 148 бит. Остающиеся 30,5 мкс используются как **защитный интервал**. Это делается, чтобы избежать перекрытия с другими пакетами из-за различных задержек на путях распространения и чтобы дать передатчику время на включение и выключение. Заполнение всего интервала позволяет передавать 156,25 бит за 577 мкс. Поэтому макси-

мальная физическая скорость передачи данных в каждом канале TDM составляет около 33,8 Кбит/с, а на каждой несущей радиочастоте через интерфейс U_m передается приблизительно 270 Кбит/с.

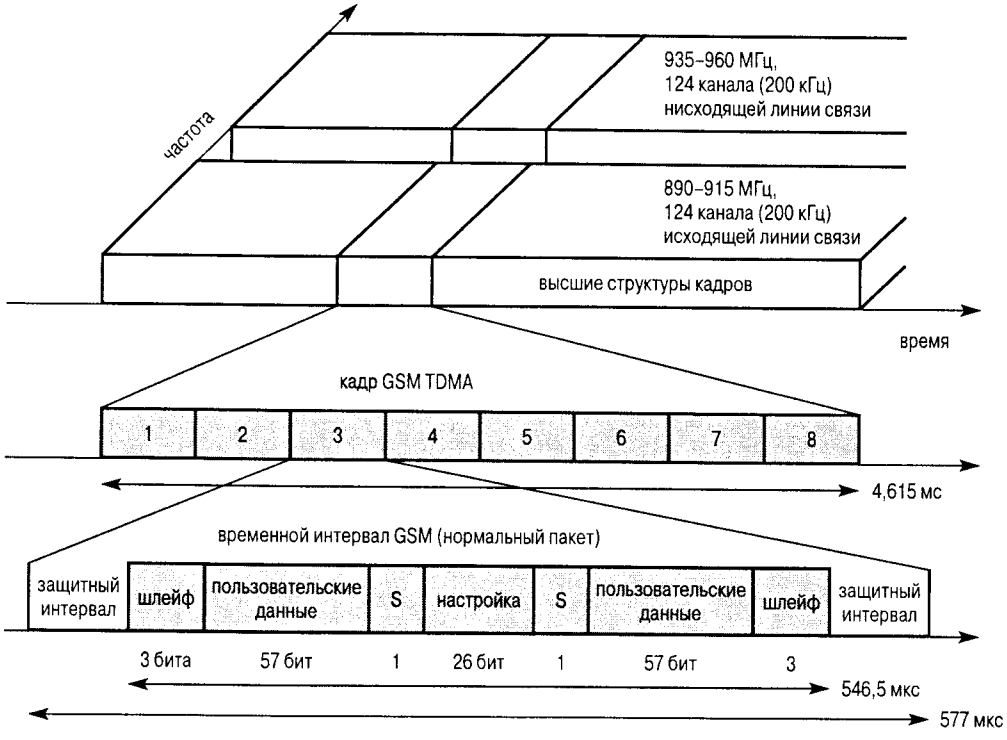

Рис. 4.4. Кадр TDMA системы GSM, интервалы и пакеты

Первый бит и последние три бита нормального пакета (**шлейф**) полагаются равными нулю. Их можно использовать для увеличения производительности приемника. **Настроечная** последовательность внутри интервала нужна для того, чтобы приспособить параметры приемника к текущим характеристикам пути распространения. При многолучевом распространении с ее помощью выбирается самый сильный сигнал. Метка S указывает на тип информации, содержащейся в поле **данных**. Эта информация может быть либо пользовательскими данными, либо данными управления сетью. Кроме нормального пакета, в стандарте [ETSI, 1993a] определено еще четыре типа пакетов для передачи данных. Пакет **коррекции частоты** позволяет мобильной станции настроить местный гетеродин так, чтобы избежать помех соседних каналов. **Пакет синхронизации** содержит расширенную настроечную последовательность и применяется для синхронизации во времени мобильной станции и станции BTS. **Пакет доступа** необходим для первоначальной установки соединения между мобильной и базовой трансиверной станцией. Наконец, **холостой пакет** используется в тех случаях, когда недоступны данные для передачи.

Возможность использования для передачи простого оборудования обусловлена двумя факторами. С одной стороны, интервалы для исходящего и нисходящего физического канала TDM разделены по частоте (в системе GSM 900 разность частот составляет 45 МГц, в системе GSM 1800 с дуплексом FDD — 95 МГц). С другой стороны, кадры TDMA сдвинуты по времени на три интервала. Если станция BTS посылает данные в первый интервал нисходящего канала, в момент времени t_0, мобильная станция получает доступ к первому интервалу исходяще-

го канала в момент времени $t_0 + 3 \cdot 577$ мкс. Поэтому мобильной станции не требуется полнодуплексный передатчик — достаточно и простого полудуплексного передатчика с переключением прием/передача.

Для избежания частотно-избирательного затухания в системе GSM вводится дополнительный механизм **медленного перескока частоты**. После каждого кадра мобильная станция и станция BTS могут изменять несущую частоту в соответствии с общей последовательностью перескоков. Мобильная станция изменяет свою частоту между интервалами исходящего и нисходящего каналов.

4.1.3.1. Логические каналы и иерархия кадров

В предыдущем разделе было показано, как происходит физическое разделение среды на $8 \cdot 124$ дуплексных канала. В этом разделе представлены логические каналы и иерархия кадров. В их основу положены физические каналы, состоящие из интервалов, повторяющихся каждые 4,615 мс. Представим себе логический канал C_1, использующий лишь каждый четвертый интервал, и другой логический канал C_2, за которым закреплены все остальные интервалы. Оба логических канала могут использовать один и тот же физический канал по схеме $C_1C_2xC_2C_1C_2xC_2C_1$ и т.д. (Здесь "x" означает, что в физическом канале остается неиспользованной некоторая пропускная способность.)

В системе GSM определены две основные группы логических каналов — каналы информационного обмена и каналы управления.

- **Каналы информационного обмена (TCH)**. Каналы TCH используются системой GSM для передачи пользовательских данных (например, речи и факсимильных сообщений). Они разделены на два основных класса — **полноскоростные (TCH/F)** и **полускоростные (TCH/H)** каналы. Каналы TCH/F характеризуются пропускной способностью 22,8 Кбит/с, а для каналов TCH/H эта величина составляет лишь 11,4 Кбит/с. Для речевых кодеков, применявшихся на начальном этапе стандартизации системы GSM, требовалась скорость передачи 13 Кбит/с. Оставшаяся емкость каналов TCH/F использовалась для коррекции ошибок (**TCH/FS**). Более новые коды обеспечивают лучшее кодирование речи. Появляется возможность применения каналов TCH/H. При этом производительность системы GSM в отношении передачи речи удваивается. Передача данных может характеризоваться различной пропускной способностью. Например, в каналах **TCH/F4.8** и **TCH/F9.6** скорость передачи составляет 4,8 Кбит/с и 9,6 Кбит/с соответственно, а в новой спецификации **TCH/F14.4** она равна 14,4 Кбит/с. Эти каналы отличаются схемами кодирования и возможностями коррекции ошибок.
- **Каналы управления (CCH)**. В системе GSM для контроля доступа к среде, распределения каналов информационного обмена и управления мобильностью используется много различных каналов CCH. Определено три группы каналов управления, причем каждая из них, в свою очередь, тоже разделяется на подгруппы (можно понять, почему для спецификации потребовалось больше 5 000 страниц).
  - **Широковещательный канал управления (BCCH)**. Станция BTS использует этот канал для передачи данных всем мобильным станциям, находящимся в ячейке. Такими данными могут быть, например, идентификатор ячейки, ее опции (перескок частоты) и частоты, доступные в данной и соседних ячейках. Информация о коррекции частоты и синхронизации во времени передается через **канал коррекции частоты (FCCH)** и **канал синхронизации (SCH)**, которые являются подканалами BCCH.
  - **Общий канал управления (CCCH)**. Через канал CCCH осуществляется обмен информацией при установке соединения между мобильной станцией и станцией BTS. Для поиска мо-

бильной станции при звонках на нее станция BTS использует **поисковый канал (PCH)**. Если же соединение устанавливает мобильная станция, она посылает данные на станцию BTS через **канал произвольного доступа (RACH)**. Для множественного доступа в канале RACH (все мобильные станции, находящиеся внутри ячейки, имеют доступ к этому каналу) применяется схема Aloha с выделением временных интервалов. На этом этапе могут возникать конфликты с другими мобильными станциями системы GSM. Через **канал предоставления доступа (AGCH)** станция BTS сообщает мобильной станции, что для дальнейшей установки соединения можно использовать канал TCH или SDCCH.

- **Выделенный канал управления (DCCH)**. В то время, как все рассмотренные выше каналы были однонаправленными, следующие каналы являются двунаправленными. Если мобильной станции не удалось установить соединение TCH со станцией BTS, она использует для передачи сигналов **автономный выделенный канал управления (SDCCH)** с низкой пропускной способностью (782 бит/с). Сигналы могут включать аутентификацию, регистрацию или другие данные, необходимые для установки соединения TCH. С каждым каналом TCH и SDCCH связан **медленный ассоциированный выделенный канал управления (SACCH)**. Через него осуществляется обмен информацией о системе, в частности, передача данных о качестве канала и уровне мощности сигнала. Наконец, когда нужно передать большее количество сигнальной информации, а канал TCH уже существует, устанавливается **быстрый ассоциированный выделенный канал управления (FACCH)**. Канал FACCH использует временные интервалы, которые в ином случае относились бы к каналу TCH. Он необходим, например, при переключениях, когда станция BTS должна обмениваться с мобильной станцией большим количеством данных за меньший промежуток времени.

Стоит отметить, что перечисленные каналы не могут использовать временные интервалы произвольным образом. В системе GSM применяется довольно сложная схема уплотнения, объединяющая несколько иерархий кадров. Пусть передача пользовательских данных проводится через простой канал TCH/F. С каждым таким каналом связан канал SACCH медленной передачи сигналов. Если же требуется быстрая передача сигналов, канал FACCH использует временные интервалы канала TCH/F. Типичная схема использования физического канала для передачи данных выглядит примерно так (здесь T обозначает пользовательскую передачу в канале TCH/F, а S — передачу сигналов в канале SACCH):

```
TTTTTTTTTTTTSTTTTTTTTTTTTx
TTTTTTTTTTTTSTTTTTTTTTTTTx
```

За двенадцатью интервалами пользовательских данных следует интервал передачи сигналов. Далее опять идут 12 интервалов с пользовательскими данными, затем следует свободный интервал. Эта схема из 26 интервалов повторяется раз за разом. В описанном случае канал TCH/F занимает лишь 24 из 26 интервалов. Напомним, что нормальный пакет переносит 114 бит пользовательских данных и повторяется каждые 4,615 мс. Это дает пропускную способность 24,7 Кбит/с. Поскольку канал TCH/F использует 24 из 26 интервалов, окончательная скорость передачи составляет 22,8 Кбит/с, что является нормой для канала TCH/F. Отсюда следует, что канал SACCH характеризуется пропускной способностью 950 бит/с.

Эта периодическая схема из 26 интервалов встречается во всех кадрах TDMA каналов TCH. Комбинация таких кадров называется **мультикадром информационного обмена**. На рис. 4.5 показано логическое объединение 26 кадров (кадров TDMA с продолжительностью 4,615 мс) в мультикадр продолжительностью 120 мс. Этот тип мультикадров используется в каналах TCH, SACCH и FACCH. Поскольку все такие логические каналы связаны с пользовательским информационным обменом, мультикадр называется мультикадром информационного обмена.

Кадры TDMA, содержащие (сигнальные) данные других логических каналов, объединяются в **мультикадр управления**. Мультикадры управления состоят из 51 кадра TDMA и имеют продолжительность 235,4 мс.

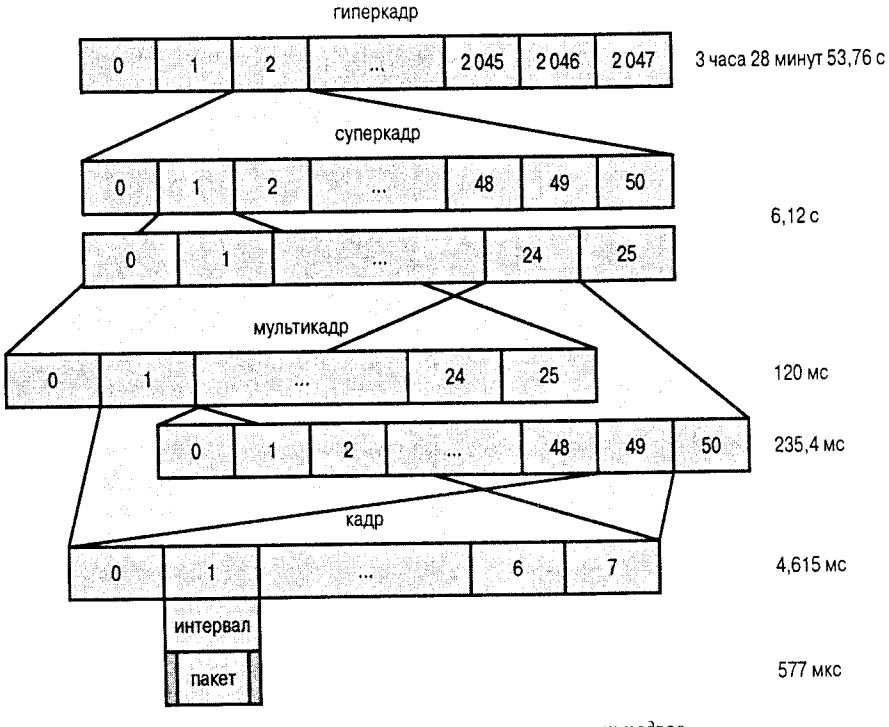

Рис. 4.5. Структурирование времени с помощью иерархии кадров

Эту иерархию логических кадров можно продолжить. Объединение 26 мультикадров с 51 кадром или 51 мультикадра с 26 кадрами образует **суперкадр**. **Гиперкадр** имеет продолжительность почти 3,5 часа и содержит 2 048 суперкадров. Всего в состав гиперкадра входит 2 715 648 кадров TDMA. Такая широкая логическая структура нужна для шифрования — в системе GSM принимается во внимание каждый кадр, причем их номера положены в основу алгоритма шифрования. Более того, в системе GSM номер кадра и номер интервала единственным образом определяют любой временной интервал.

4.1.4. Протоколы

Архитектура протоколов системы GSM показана на рис. 4.6. Она включает сигнальные протоколы, интерфейсы и объекты, которые уже приводились на рис. 4.3. Опять же, основное внимание обращается на интерфейс U_m, так как другие интерфейсы встречаются и в стационарных сетях. **Уровень 1** (физический уровень) выполняет все функции **радиосвязи**. Сюда можно отнести формирование пакетов пяти различных форматов, **уплотнение** пакетов в кадр TDMA, **синхронизацию** со станцией BTS, детектирование свободных каналов и контроль **качества нисходящего канала**. На физическом уровне интерфейса U_m осуществляется цифровая **модуляция** по схеме GMSK и **шифрование/дешифрование** данных. Заметим, что шифрование не является сквозным, а проводится только через радиоинтерфейс между станцией BTS и мобильной станцией.

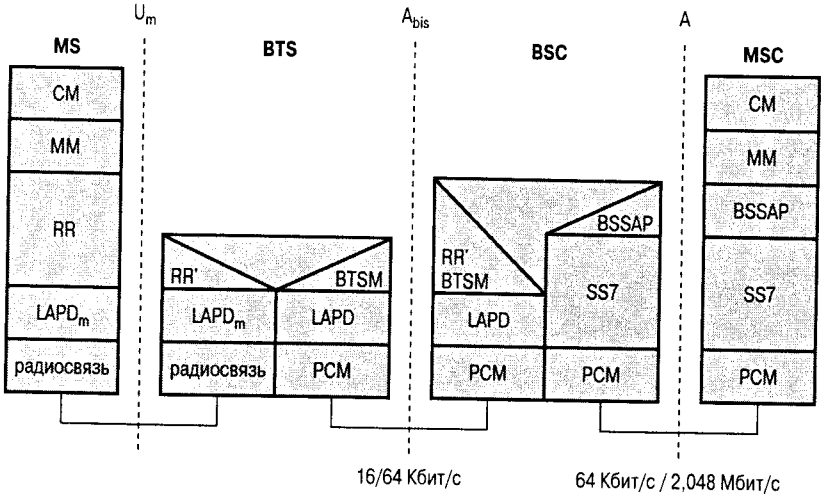

Рис. 4.6. Архитектура протоколов передачи сигналов

Синхронизация включает и коррекцию задержек на отдельных путях между станциями MS и BTS. Все мобильные станции внутри ячейки используют одну и ту же станцию BTS и поэтому должны синхронизироваться с ней. Станция BTS генерирует временную структуру кадров, интервалов и т.п. Сложным вопросом в этом отношении являются различные периоды кругового обращения сигнала (RTT). Мобильная станция, находящаяся близко к станции BTS, имеет очень короткий период RTT. В то же время период RTT для станции MS, удаленной на расстояние 35 км, составляет уже около 0,23 мс. Чтобы удаленная мобильная станция могла использовать структуру интервалов без коррекции, требуются широкие защитные интервалы. Такое условие легко понять: 0,23 мс — это больше 40 процентов всего интервала (0,577 мс). Поэтому в действительности поступают по-другому. Станция BTS посылает на станцию MS значение текущего периода RTT. Мобильная станция затем устанавливает такое время доступа, что все пакеты достигают станции BTS, не выходя за положенные пределы. Этот механизм позволяет уменьшить продолжительность защитного интервала до 30,5 мкс, что составляет лишь 5% всего интервала (см. рис. 4.4). Регулирование доступа проводится с помощью переменного **временного опережения**. Пакет может сдвигаться на величину, кратную 63 битам, причем длительность каждого бита составляет 3,69 мкс (что приводит к требуемой продолжительности 0,23 мс).

Канальное кодирование и **обнаружение/коррекция ошибок** являются главными задачами физического уровня. Они непосредственно связаны с механизмами кодирования. При канальном кодировании находят широкое применение схемы **прямой коррекции ошибок (FEC)**. В схемах FEC к пользовательским данным добавляется избыточная информация, которая позволяет обнаруживать и исправлять некоторые ошибки. Производительность схемы FEC зависит от количества избыточной информации, алгоритма кодирования и дальнейшего чередования данных, необходимого для уменьшения числа пакетных ошибок. Использование схемы FEC также объясняет то, что обнаружение и коррекция ошибок происходят на первом, а не на втором уровне эталонной модели ISO/OSI. В системе GSM попытки исправления ошибок производятся на физическом уровне, а на высшие уровни ошибочные данные не передаются.

Логические каналы системы GSM используют разные схемы кодирования с различными возможностями коррекции ошибок. В речевых каналах для достижения скорости передачи 22,8 Кбит/с (которая включает пропускную способность 13 Кбит/с речевого кодека, избыточ-

ную информацию, биты CRC и чередование [Goodman, 1997]) требуется дополнительное кодирование голосовых данных. Оно производится после превращения аналоговых сигналов в цифровые. Поскольку передача речи является главной услугой, предоставляемой системой GSM, физический уровень тоже выполняет специальные функции, в частности, **детектирование активности речи (VAD)**. Этот механизм позволяет передавать голосовые данные лишь в том случае, когда есть голосовой сигнал. Во время пауз (например, когда пользователю нужно подумать, прежде чем ответить) физический уровень генерирует **комфортный шум**, имитирующий соединение (абсолютная тишина может смутить пользователя). На самом деле в эти периоды передача данных не происходит.

Для минимизации помех, вызванных пакетными ошибками, применяется чередование данных. Этот механизм, а также схема повторения логического канала, приводят к **задержке** передачи. В каналах TCH/FS такая задержка составляет около 60 мс, а в каналах TCH/F9.6 — приблизительно 100 мс. Эти интервалы необходимо добавлять к задержке передачи в тех случаях, когда для связи вместо обычных стационарных станций (телефонов, компьютеров и т.д.) используются мобильные станции. Такие интервалы могут повлиять на производительность протоколов высших уровней (используемых, например, для передачи данных между компьютерами).

Для передачи сигналов между объектами сети GSM необходимы высшие уровни (см. рис. 4.6). Протокол **$LAPD_m$** для **второго уровня** в интерфейсе U_m определен именно с этой целью. Он происходит от процедуры доступа к D-каналу (**LAPD**) в системах ISDN — версии протокола HDLC [Goodman, 1997], [Halsall, 1996]. Протокол $LAPD_m$ является "облегченным" в сравнении с протоколом LAPD, так как для обнаружения ошибок ему не нужны ни метки синхронизации, ни расчет контрольных сумм. (За выполнение этих задач отвечает физический уровень системы GSM.) Протокол $LAPD_m$ обеспечивает надежную передачу информации, повторное упорядочение кадров данных и управление потоком [ETSI, 1993б], [ETSI, 1993в]. Поскольку между первым и вторым уровнями нет буферизации, протокол $LAPD_m$ использует структуру кадров, повторяющиеся последовательности битов и т.п., определенные для интерфейса U_m. Дополнительные службы, предоставляемые протоколом $LAPD_m$, включают сегментацию и дефрагментацию данных, а также передачу с подтверждением/без подтверждения.

В состав **третьего** сетевого уровня системы GSM входит несколько подуровней, также показанных на рис. 4.6. Самый нижний подуровень отвечает за **управление радиоресурсами (RR)**. Часть **RR'** этого подуровня размещена в станции BTS, оставшаяся часть реализована в контроллере BSC. Функции RR' поддерживаются контроллером BSC через **управление станцией BTS (BTSM)**. Главными задачами механизма RR являются установка, обслуживание и освобождение радиоканалов. Протоколы управления радиоресурсами также имеют непосредственный доступ к радиоинформации на физическом уровне и обеспечивают надежное соединение с вышестоящим уровнем.

Протокол управления мобильностью (MM) отвечает за регистрацию, аутентификацию, идентификацию, обновление информации о местонахождении и обеспечение **временной идентификации мобильного абонента (TMSI)**. Идентификация TMSI заменяет **международную идентификацию мобильного абонента (IMSI)** и необходима для того, чтобы с помощью радиоинтерфейса скрыть истинную личность пользователя мобильной станции. Механизм IMSI идентифицирует пользователя, а механизм TMSI действует лишь в текущей зоне местонахождения регистра VLR. Протокол управления мобильностью предоставляет надежное соединение со следующим высшим уровнем.

Наконец, уровень **управления соединениями (CM)** содержит три объекта — управление вызовами (**CC**), **службу коротких сообщений (SMS)** и **дополнительные службы (SS)**. Служба SMS позволяет передавать сообщения с помощью каналов SDCCH и SACCH (если не нужно посы-

лать сигнальные данные). Службы SS были описаны в разделе 4.1.1.3. Механизм CC используется на высших уровнях для установки и отключения соединения, а также для изменения его параметров. Этот уровень также выполняет функции передачи внутриполосных тонов, что иначе называется **двухтональным многочастотным набором (DTMF)**, через сеть GSM. Такие тона применяются, например, в банковском деле для удаленного управления автоответчиками и ввода номеров PIN, а также для набора номеров в обычной аналоговой телефонной системе. Эти тона нельзя послать непосредственно через голосовой кодек мобильной станции системы GSM, так как кодек их исказит. Поэтому они передаются как сигналы, а затем в стационарной части сети GSM опять преобразуются в тона.

В интерфейсах A_{bis} и A используются дополнительные протоколы (внутренние интерфейсы системы GSM здесь не представлены). Передача данных на физическом уровне обычно осуществляется с помощью систем **импульсно-кодовой модуляции (PCM)**. Системы PCM в большинстве случаев предоставляют сквозные каналы с пропускной способностью 64 Кбит/с. В системе GSM возможно также уплотнение четырех каналов со скоростью передачи 16 Кбит/с в отдельный канал со скоростью передачи 64 Кбит/с (пропускной способности 16 Кбит/с достаточно для пользовательских данных мобильной станции). Физический уровень интерфейса A чаще всего содержит выделенные линии емкостью 2,048 Мбит/с. На втором уровне интерфейса A_{bis} применяется протокол LAPD, а для управления станцией BTS используется протокол BTSM.

Передача сигналов между центром MSC и контроллером BSC осуществляется через **систему SS7**. С помощью этого же протокола передается вся управляющая информация между центрами MSC, AuC и OMC, регистрами HLR, VLR и EIR. Кроме того, центр MSC может управлять подсистемой BSS через **прикладную часть BSS (BSSAP)**.

4.1.5. Локализация и вызовы

Одной из ключевых особенностей системы GSM является автоматическая всемирная локализация пользователей. Системе всегда известно, где находится пользователь, поэтому телефонные номера действительны в любой точке земного шара. Для обеспечения такой услуги в системе GSM информация о местонахождении пользователя периодически обновляется. Это происходит даже в том случае, когда абонент не пользуется мобильной станцией (станция MS остается зарегистрированной в сети GSM и полностью не отключается). В регистре HLR постоянно содержатся данные о текущем местонахождении, а с помощью регистра VLR, отвечающего за мобильную станцию, учитывается изменение ее положения. Когда станция MS попадает в зону действия другого регистра VLR, ему сразу передаются необходимые данные регистра HLR. Изменение регистров VLR, при котором все службы остаются доступными, называется **роумингом** (roaming). Роуминг может осуществляться в сети одного поставщика услуг, между двумя поставщиками услуг одной страны (общегосударственный роуминг обычно не поддерживается) и между различными поставщиками разных стран (международный роуминг).

Чтобы определить местонахождение мобильной станции и обратиться к ней, необходимо несколько номеров.

- **Международный номер ISDN мобильной станции (Mobile station international ISDN number — MSISDN)**[2]. Для абонента GSM единственным важным номером является телефонный номер. Напомним, что он связан не с каким-то определенным устройством, а с модулем SIM, принадлежащим пользователю. Структура номера MSISDN определяется стандартом E.164 института ITU-T. Этот же стандарт используется в стационарных сетях ISDN. Номер

[2] Помимо приведенной расшифровки аббревиатуры встречаются и другие (даже в первоначальных стандартах ETSI): "Mobile Subscriber ISDN Number" (номер ISDN мобильного абонента) или "Mobile Station ISDN Number" (номер ISDN мобильной станции)

MSISDN состоит из **кода страны (CC)** (например, в Германии применяется код 49; в некоторых странах используются специальные префиксы, в частности, 009 или 00), **национального кода адресата (NDC)** (т.е. адреса сетевого поставщика услуг и регистра HLR) и **номера абонента (SN)**.

- **Международная идентификация мобильного абонента (IMSI)**. В технологии GSM механизм IMSI применяется для однозначной идентификации абонента. Эта идентификация состоит из **мобильного кода страны (MCC)** (например, в Швеции используется код 240, во Франции — код 208), **мобильного кода сети (MNC)** (т.е. кода регистра HLR) и **идентификационный номер мобильного абонента (MSIN)**.
- **Временная идентификация мобильного абонента (TMSI)**. Идентификация IMSI обнаруживает личность пользователя, передающего сигналы через радиоинтерфейс. Чтобы скрыть эту информацию, в технологии GSM для локального опознавания абонента применяется идентификация TMSI (4 байта). Она выбирается текущим регистром VLR и действует в его зоне местонахождения лишь временно. К тому же, регистр VLR периодически изменяет идентификацию TMSI.
- **Номер роуминга мобильной станции (MSRN)**. Номер MSRN является еще одним временным адресом, скрывающим личность и местонахождение абонента. Этот адрес генерируется регистром VLR после запроса из центра MSC и сохраняется в регистре HLR. Номер MSRN включает **гостевой код страны (VCC)**, **гостевой государственный код адресата (VNDC)**, идентификацию текущего центра MSC и номер абонента. Таким образом, при входящих звонках номер MSRN помогает найти нужного абонента.

Все перечисленные номера нужны для определения местонахождения абонента и поддержки соединения с мобильной станцией. Рассмотрим **вызов мобильного оконечного устройства (MTC)**, т.е. ситуацию, когда какая-то станция вызывает мобильную станцию (вызывающая станция может находиться вне сети GSM или быть другой мобильной станцией). Основные этапы соединения вызывающей станции с мобильным пользователем показаны на рис. 4.7. На этапе 1 пользователь набирает телефонный номер абонента GSM. Стационарная сеть (PSTN) устанавливает (с помощью кода адресата), что номер относится к пользователю сети GSM и переадресовывает установку соединения на шлюз GMSC (2). Шлюз GMSC определяет регистр HLR абонента (он закодирован в телефонном номере) и посылает в этот регистр сигнал установки соединения (3). Регистр HLR проверяет существование нужного номера и подписку пользователя на требуемые услуги. Затем из текущего регистра VLR запрашивается номер MSRN (4). После его получения (5) регистр HLR определяет, какой из центров MSC отвечает за мобильную станцию. Эта информация направляется на шлюз GMSC (6), который, в свою очередь, переадресовывает запрос установки соединения на нужный центр MSC (7).

Начиная с этого места, за все дальнейшие шаги отвечает центр MSC. Прежде всего он запрашивает из регистра VLR информацию о текущем состоянии мобильной станции (8). Если мобильная станция доступна, центр MSC инициирует ее поиск во всех своих ячейках (10) (поиск одной нужной ячейки забрал бы слишком много времени). Станции BTS всех подсистем BSS посылают поисковый сигнал на мобильную станцию (11). После ответа (12 и 13) регистр VLR должен произвести проверку безопасности (установить шифрование и т.п.). Затем регистр VLR посылает сигнал об установке соединения с мобильной станцией центру MSC (15–17).

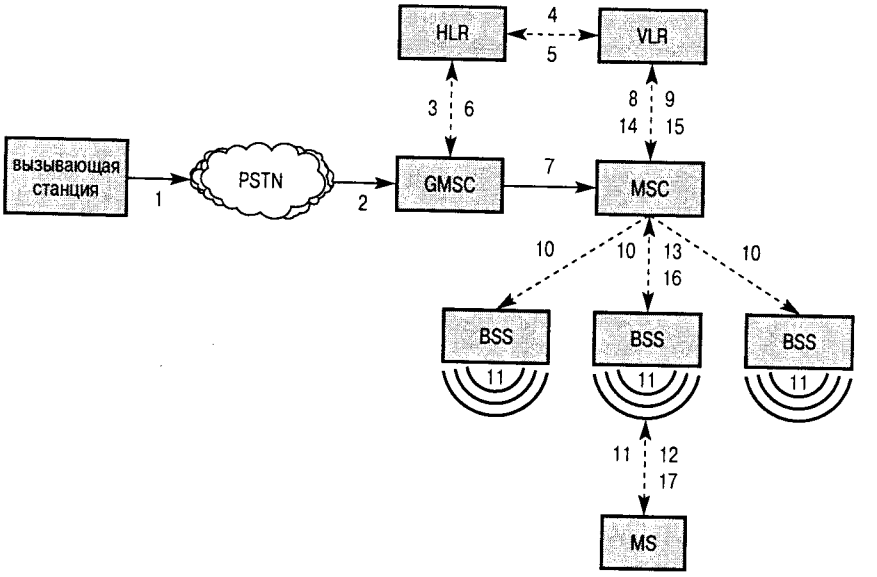

Рис. 4.7. Вызов мобильного оконечного устройства (MTC)

Вызов с мобильного источника (MOC) совершается намного проще (рис. 4.8). Мобильная станция посылает запрос на установку нового соединения (1). Подсистема BSS переадресовывает этот запрос центру MSC (2). Центр MSC проверяет, может ли данный абонент пользоваться требуемой службой (3 и 4). Далее производится проверка доступности ресурсов сетей GSM и PSTN. Если необходимые ресурсы доступны, центр MSC устанавливает соединение между мобильной станцией и стационарной сетью.

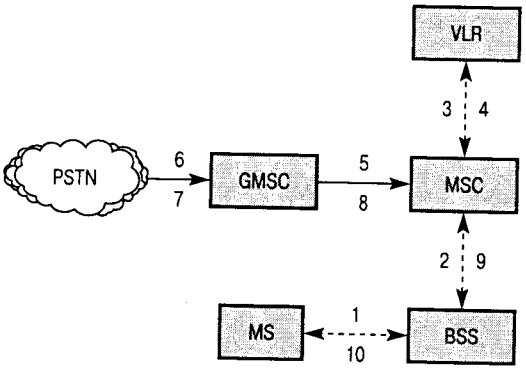

Рис. 4.8. Вызов с мобильного источника (MOC)

При установке соединения, кроме описанных выше шагов, происходит обмен некоторыми сообщениями между мобильной станцией и станцией BTS (в обоих направлениях). На рис. 4.9 приведены сообщения для вызовов MTC и MOC. Поиск необходим только при вызове MTC. Дальнейшие сообщения для обоих типов вызовов похожи. Первым шагом в этом отношении является получение доступа к каналу с последующим выделением канала. С этой целью используется канал случайного доступа (RACH). Присваиваться может канал информационного обмена (TCH) или более медленный сигнальный канал SDCCH.

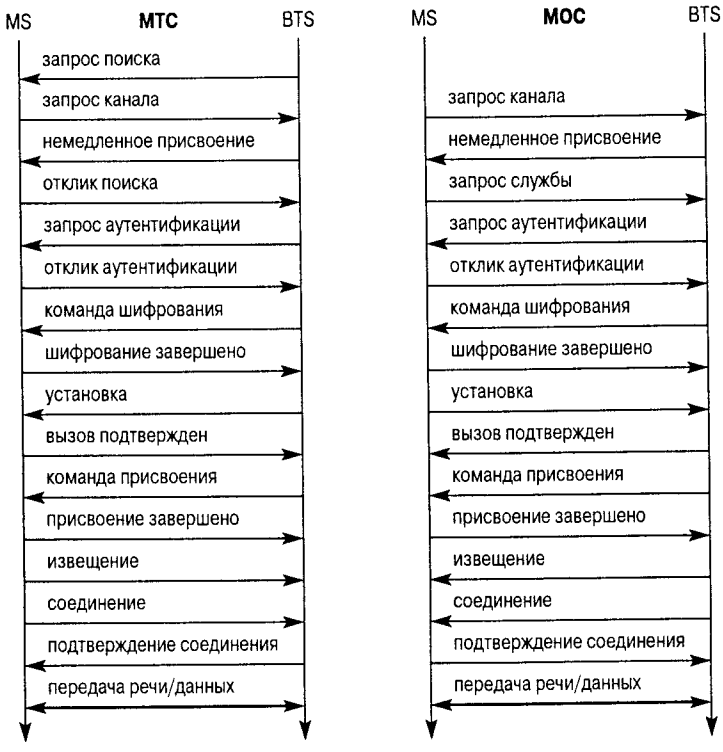

Рис. 4.9. Поток сообщений для вызовов МТС и МОС

Следующие этапы необходимы для обеспечения безопасности соединения. Они включают аутентификацию мобильной станции и переключение на шифрованную связь. Затем система присваивает канал TCH (если это не было сделано ранее). Присвоение канала TCH на этом этапе дает следующее преимущество: в этом случае на первых шагах установки соединения используется лишь канал SDCCH. Если установка пройдет неудачно, канал TCH не будет блокирован. С другой стороны, применение канала TCH с самого начала дает выигрыш с точки зрения скорости.

Дальнейшие шаги зависят от типа вызова (МОС или МТС). При вызове мобильной станции она отвечает "извещением" (звуковой сигнал мобильной станции) и "соединением" (когда пользователь нажимает кнопку "connect"). То же самое происходит, если вызов был инициирован мобильной станцией. После подтверждения соединения обе стороны могут обмениваться данными.

Окончание соединения включает сообщение о разъединении, посылаемое одним из пользователей, за которым следует освобождение радиоканала.

4.1.6. Переключение

Отдельные ячейки не покрывают всей области действия службы. Их зона действия ограничена, например, 35-километровым расстоянием до антенны [Tripathi, 1998]. Поэтому в сотовых системах необходимо **переключение**. Чем меньше размер ячейки и быстрее движение мобильной станции, тем больше требуется переключений текущих звонков. Однако, переключение не должно вызывать обрыва соединения, также называемого **выпадением звонков**.

Есть две главных причины переключения (в стандарте их определено около сорока).

- Мобильная станция **выходит из зоны действия** станции BTS или некоторой антенны станции BTS. Тогда принимаемый **уровень сигнала** непрерывно уменьшается до тех пор, пока не упадет ниже минимальных требований. В другом случае может увеличиться **количество ошибок**, обусловленных помехами, — расстояние до станции BTS может быть слишком велико (макс. 35 км). Все эти эффекты ухудшают **качество радиоканала**, и радиопередача вскоре становится невозможной.
- Проводная инфраструктура (центр MSC, контроллер BSC) может решить, что **информационный обмен в одной из ячеек является слишком интенсивным**, и переместить мобильную станцию в другую ячейку с более слабой загрузкой (если это возможно). Таким образом, переключение может быть обусловлено **балансированием загрузки**.

На рис. 4.10 показано четыре возможных типа переключения в системе GSM:

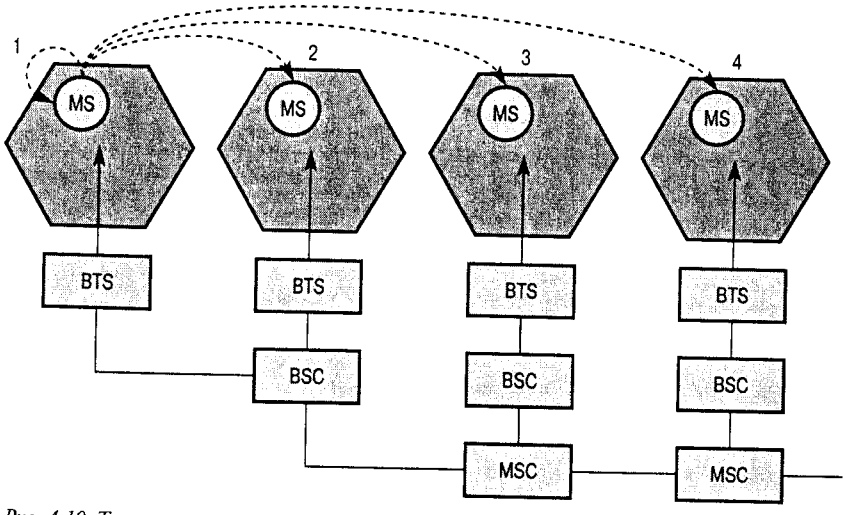

Рис. 4.10. Типы переключения в системе GSM

- **Переключение внутри ячейки**. Внутри ячейки передача на некоторой частоте может стать невозможной из-за узкополосных помех. В этом случае контроллер BSC может изменить несущую частоту (тип 1).
- **Переключение ячеек одного контроллера BSC**. Это наиболее распространенное переключение. Мобильная станция переходит из одной ячейки в другую, но находится в зоне управления одного и того же контроллера BSC. Он осуществляет переключение, выделяет новый радиоканал в новой ячейке и освобождает старый канал (тип 2).
- **Переключение контроллеров BSC одного центра MSC**. Поскольку контроллер BSC управляет ограниченным числом ячеек, в системе GSM иногда необходимо осуществлять переключения между ячейками, управляемыми различными контроллерами BSC. Такое переключение производится центром MSC (тип 3). Эту ситуацию иллюстрирует и рис. 4.12.
- **Переключение центров MSC**. Наконец, может потребоваться переключение двух ячеек, относящихся к разным центрам MSC. В этом случае переключение осуществляется совместно — обоими центрами MSC (тип 4).

Станция BTS и мобильная станция периодически измеряют качество исходящего и нисходящего каналов соответственно. (В качество канала входит уровень сигнала и количество оши-

бок.) Это делается для получения всей информации, необходимой при переключении из-за слабого соединения. Результаты измерений передаются мобильной станцией каждые полсекунды. Они содержат информацию не только о качестве текущего канала передачи, но и о качестве некоторых каналов в соседних ячейках (каналы BCCH).

Рис. 4.11 иллюстрирует типичное поведение уровня принимаемого сигнала в ситуации, когда мобильная станция удаляется от одной станции BTS (BTS_{old}) и приближается к другой (BTS_{new}). В этом случае регулирование переключения зависит не от действительного уровня принимаемого сигнала, а от его среднего значения. Поэтому контроллер BSC получает со станций MS и BTS все значения (количества ошибок за единицу времени и уровней сигналов в исходящем и нисходящем канале) и усредняет их. Затем результаты сравниваются с пороговыми значениями, т.е. с границей переключения (HO_MARGIN). Эта граница для избежания эффекта пинг-понга включает определенное запаздывание [Wong, 1997]. (Без запаздывания переключение могут активировать даже краткосрочные помехи, в частности, экранирование сигнала зданиями.) Однако, даже при использовании границы HO_MARGIN технология GSM не защищена от эффекта пинг-понга — слишком большое значение может вызвать обрыв соединения, а слишком низкое — привести к большому числу переключений.

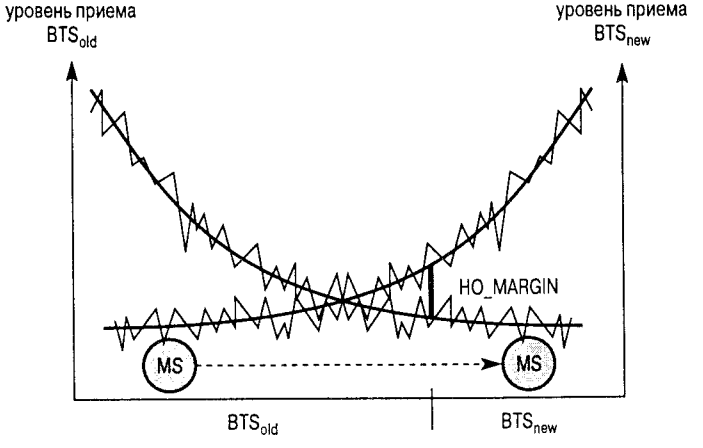

Рис. 4.11. Регулирование переключения в зависимости от уровня приема

На рис. 4.12 приводится поток сигналов при переключении контроллеров BSC одного центра MSC. Мобильная станция периодически передает свои измерения. Станция BTS_{old} объединяет их с собственными измерениями и перенаправляет эти данные на контроллер BSC_{old}. Основываясь на этих значениях и, например, на условиях информационного обмена, контроллер BSC_{old} может решить произвести переключение. Тогда он передает центру MSC сообщение "требуется переключение". Затем центр MSC посылает запрос ресурсов, необходимых для переключения, на другой контроллер BSC_{new}. Этот контроллер BSC проверяет доступность ресурсов (частот или временных интервалов) и активирует физический канал станции BTS_{new}, подготавливая прием мобильной станции.

Станция BTS_{new} подтверждает успешную активацию канала. Контроллер BSC_{new}, в свою очередь, подтверждает запрос переключения. Затем центр MSC отдает команду переключения и переадресовывает ее мобильной станции. Станция MS разрывает старое соединение и получает доступ к новой станции BTS. Далее происходит установка соединения (сюда входит установка соединения на втором уровне и передача мобильной станцией сообщения "переключение завершено"). На этом этапе можно считать, что переключение мобильной станции завершилось. Тем не менее, нужно еще осво-

бодить ресурсы старого контроллера BSC_{old} и базовой станции BTS_{old}, а также подтвердить успешное переключение сообщениями о завершении переключения и очистки.

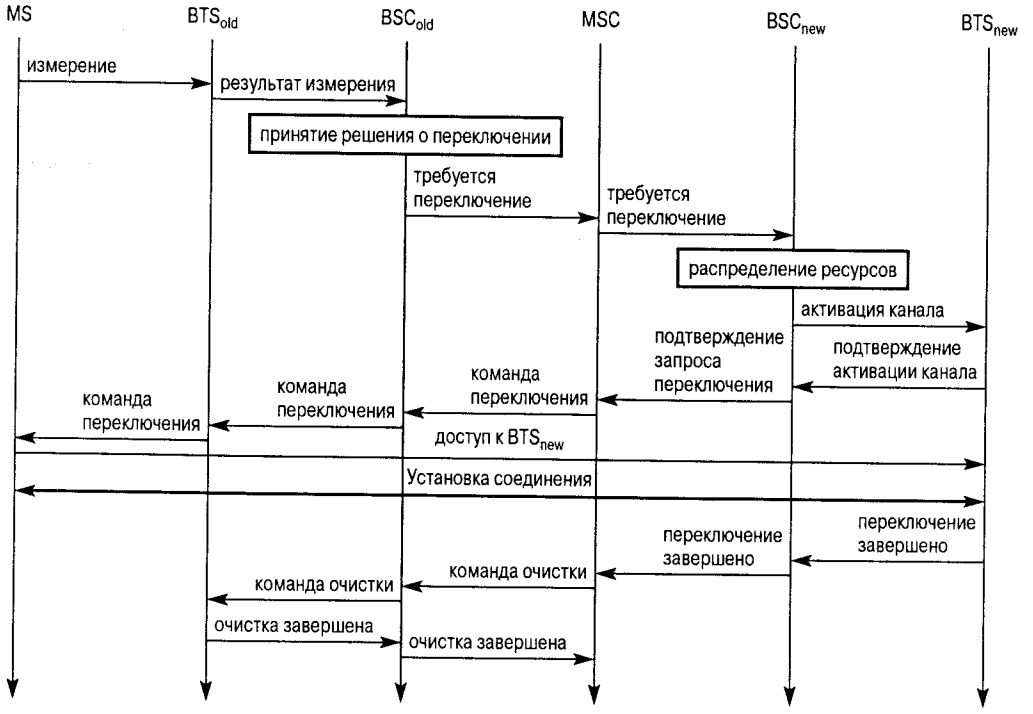

Рис. 4.12. Переключение контроллеров BSC одного центра MSC

В будущем станет возможным прямое переключение между различными сетями, например, между системами GSM, DECT и спутниковыми службами. Для этого нужны многорежимные мобильные станции и более сложная инфраструктура роуминга.

4.1.7. Безопасность

Технология GSM включает несколько служб безопасности. Они используют конфиденциальную информацию, содержащуюся в центре AuC и индивидуальном модуле SIM (который можно подключать к любой мобильной станции). Ранее уже упоминалось, что в модуле SIM хранятся личные, секретные данные. (К ним относится, в частности, секретный ключ K_i, применяющийся для аутентификации и шифрования). Номер PIN защищает модуль SIM от несанкционированного использования. Службы безопасности, предоставляемые системой GSM, можно описать следующим образом.

- **Контроль доступа и аутентификация**. Первый этап включает аутентификацию владельца модуля SIM. Для получения доступа к модулю SIM пользователю необходим секретный номер PIN. Следующий шаг — аутентификация абонента (см. рис. 4.9). В основу этого механизма положена схема "вызов-отклик", представленная в разделе 4.1.7.1.
- **Конфиденциальность**. Все пользовательские данные шифруются. После аутентификации станции MS и BTS осуществляют шифрование речи, данных и сигналов. Более подробно эти вопросы обсуждаются в разделе 4.1.7.2. Конфиденциальность обеспечивается не во всей стационарной сети GSM (телефонной сети), а лишь между мобильной станцией и станцией BTS.

- **Анонимность**. Чтобы обеспечить анонимность пользователя, производится шифрование всех данных. Идентификаторы пользователя, с помощью которых можно обнаружить его личность, при радиопередаче не используются. В технологии GSM вместо них передается временный идентификатор (TMSI). Он присваивается регистром VLR и изменяется всякий раз при обновлении местонахождения. К тому же регистр VLR и сам в любое время может изменить идентификатор TMSI.

Службы безопасности системы GSM используют три алгоритма. Для **аутентификации** применяется **алгоритм A3**, для **шифрования** — **алгоритм A5**, для **генерации ключей шифров** — **алгоритм A8**. В стандарте GSM общедоступным являлся лишь алгоритм A5. Алгоритмы A3 и A8 держались в секрете, но стандартизовались с открытыми интерфейсами. Сейчас оба алгоритма уже не секретны — в 1998 году они были опубликованы в сети Internet. Этот пример показывает, что с помощью утаивания можно добиться лишь иллюзии безопасности. Описанные алгоритмы оказались не очень надежными. Однако, сетевые поставщики услуг могут использовать для аутентификации более серьезные алгоритмы. Кроме того, может применяться более "прочное" сквозное шифрование.

4.1.7.1. Аутентификация

Перед использованием любой службы сети GSM абоненты должны пройти аутентификацию. В ее основу положены данные, содержащиеся в модуле SIM — **индивидуальный ключ аутентификации K_i** и **идентификация пользователя IMSI**. Кроме того, для аутентификации применяется алгоритм A3 и схема "вызов-отклик". Последний механизм действует следующим образом. Блок управления доступом (AC) в качестве оклика генерирует случайное число **RAND**. Модуль SIM мобильной станции отвечает отмеченным откликом (**SRES**) (рис. 4.13). Центр AuC осуществляет генерацию случайных значений RAND, отмеченных откликом SRES и ключей шифров K_c для каждой идентификации IMSI. Затем эта информация направляется в регистр HLR. Здесь она доступна для запросов текущего регистра VLR.

При аутентификации регистр VLR посылает случайное значение RAND на модуль SIM. Обе стороны — сеть и модуль абонента — выполняют с номером RAND и ключом K_i одинаковые операции (это и есть алгоритм A3). Мобильная станция посылает отклик SRES, генерируемый модулем SIM, обратно в регистр VLR. Здесь оба значения сравниваются. При их совпадении регистр VLR принимает абонента, в ином случае абонент получает отказ.

4.1.7.2. Шифрование

В системе GSM все сообщения, содержащие пользовательскую информацию, кодируются. После аутентификации станции MS и BTS начинают шифрование с использованием ключа шифра K_c. Ключ K_c генерируется из индивидуального ключа K_i и случайного номера с помощью алгоритма A8. Заметим, что сеть и модуль SIM мобильной станции получают одинаковые ключи K_c, основываясь на случайном значении RAND. Сам ключ K_c через радиоинтерфейс не передается.

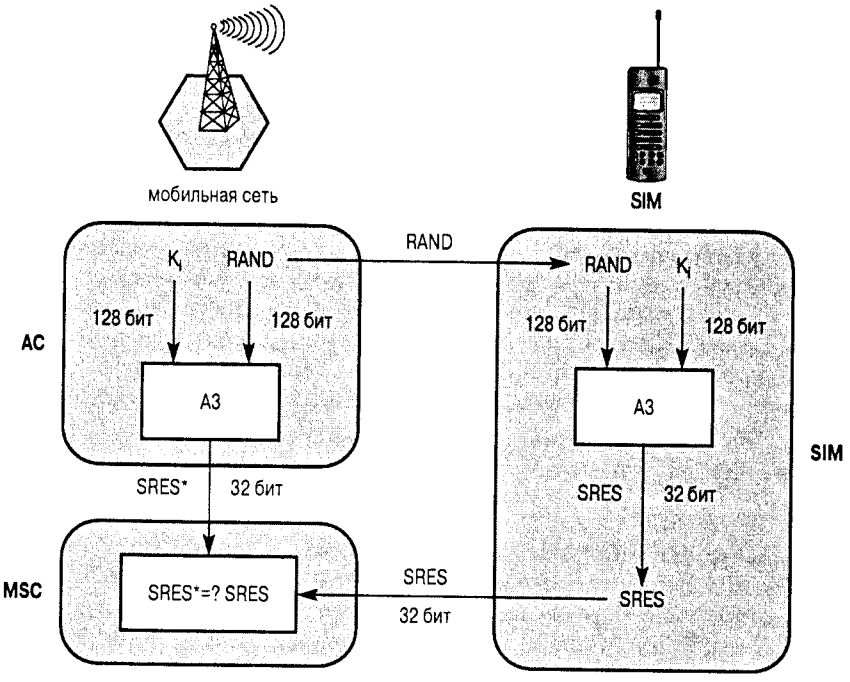

Рис. 4.13. Аутентификация абонента

Итак, станции MS и BTS могут выполнять шифрование и дешифрование с помощью алгоритма A5. Ключ K_c имеет размер 64 бит (рис. 4.14). Это не очень надежный ключ, но его достаточно, по крайней мере, для защиты от простого прослушивания. Публикация алгоритмов A3 и A8 в сети Internet показала, что в некоторых случаях 10 из 64 бит полагаются равными нулю и, соответственно, шифрование становится менее надежным.

4.1.8. Новые службы передачи данных

Стандартная пропускная способность 9,6 Кбит/с (14,4 Кбит/с у некоторых поставщиков услуг), доступная при передаче данных, не удовлетворяет требованиям современных компьютеров. Когда система GSM только разрабатывалась, мало кто мог предвидеть столь быстрое развитие обмена данными (по отношению к развитию голосовой связи). В то время скорость 9,6 Кбит/с являлась большой или, по крайней мере, достаточной для обычных факсимильных аппаратов группы 3. Однако, в последние годы потребности пользователей возросли. Оказалось, что для просмотра документов WWW, загрузки файлов и даже для интенсивного обмена электронными письмами с присоединенными файлами, пропускной способности 9,6 Кбит/с недостаточно.

Улучшить производительность системы GSM в отношении передачи данных можно двумя способами. В основу технологии GSM положены каналы информационного обмена, ориентированные на установление соединения. Пропускная способность каждого канала составляет, например, 9,6 Кбит/с. Можно объединить несколько каналов, тем самым увеличив пропускную способность. Система, действующая таким образом, представлена в следующем разделе (технология HSCSD). Более прогрессивным шагом является включение в технологию GSM пакетно-ориентированного информационного обмена. В этом случае акцент переносится с соединений/телефонной связи на пакетную передачу/Internet. Соответствующая система (GPRS) обсуждается в разделе 4.1.8.2.

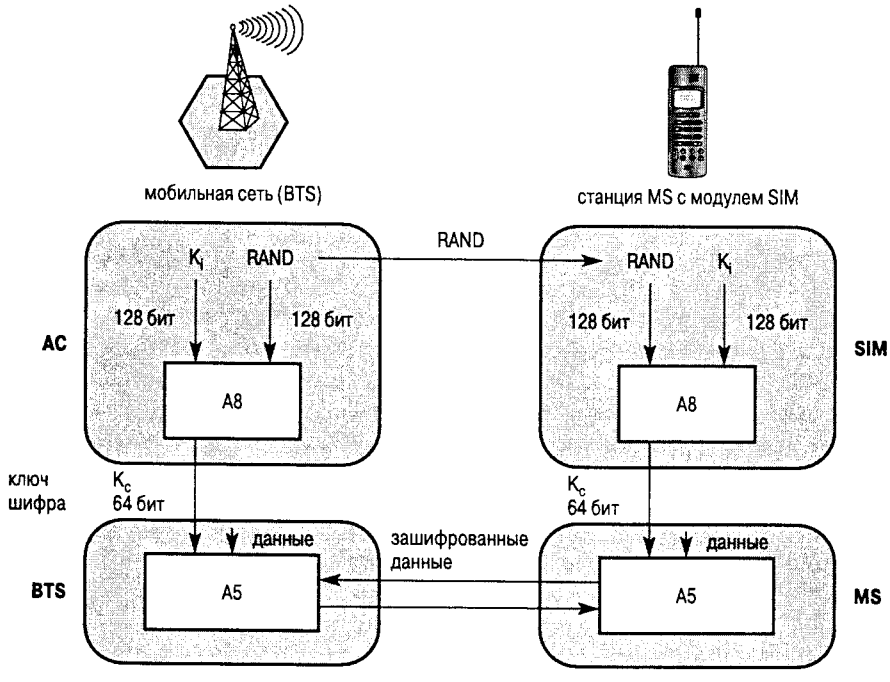

Рис. 4.14. Шифрование данных

4.1.8.1. HSCSD

Технология **высокоскоростной передачи данных с коммутацией каналов (HSCSD)** позволяет улучшить характеристики передачи данных в системе GSM. Эта технология уже сейчас используется некоторыми поставщиками услуг. Более высокая скорость передачи достигается за счет объединения нескольких каналов TCH. Мобильная станция может запрашивать для передачи более одного канала TCH сети GSM, т.е. одной станции выделяется несколько интервалов кадра TDMA. Распределение каналов может быть асимметричным. В этом случае для нисходящей линии связи выделяется больше временных интервалов, чем для исходящей. Такой механизм отражает типичный режим работы пользователей — данные в основном принимаются, а не отправляются. Для реализации системы HSCSD необходимо усовершенствовать лишь программное обеспечение станции MS и центра MSC (чтобы они могли разбивать информационный поток на несколько потоков, использующих отдельные каналы TCH, а затем объединять их снова).

Теоретически мобильная станция может использовать все восемь интервалов кадра TDMA. При этом достигается **пользовательская скорость передачи через радиоинтерфейс (AIUR)**, равная, например, 115,2 Кбит/с для 8 каналов TCH/F14.4 [ETSI, 1998д]. Однако, возникают и проблемы — мобильная станция должна посылать и принимать данные одновременно. Обычная система GSM не позволяет этого. В ней исходящий канал сдвинут относительно нисходящего на три интервала. В стандарте [ETSI, 1997a] скорость AIUR принимает значение 57,6 Кбит/с (дуплекс). Для этого используются четыре интервала в исходящем и нисходящем канале (для несквозных служб разрешенные комбинации каналов информационного обмена и соответствующие интервалы приведены в табл. 4.2).

Таблица 4.2. Возможные значения скорости передачи в системе HSCSD

AIUR	TCH/F4.8	TCH/F9.6	TCH/F14.4
4,8 Кбит/с	1	–	–
9,6 Кбит/с	2	1	–
14,4 Кбит/с	3	–	1
19,2 Кбит/с	4	2	–
28,8 Кбит/с	–	3	2
38,4 Кбит/с	–	4	–
43,2 Кбит/с	–	–	3
57,6 Кбит/с	–	–	4

Несмотря на всю свою кажущуюся привлекательность, технология HSCSD обладает существенными недостатками. В ней все еще применяются механизмы системы GSM, ориентированные на установление соединения и неэффективные при (пакетном) информационном обмене компьютерных данных. Для загрузки большого файла могут потребоваться все зарезервированные каналы. В то же время, при обычной работе с системой WWW каналы большую часть времени остаются свободными. Распределение каналов непосредственно влияет на стоимость услуг, так как зарезервированные каналы, даже если они свободны, не могут использоваться другими абонентами.

Кроме того, при использовании технологии HSCSD с *n* каналами переключение, установка соединения и освобождение каналов занимают в *n* раз больше времени. Каждый канал рассматривается отдельно. Вероятность блокирования или ухудшения услуг при переключении тоже возрастает, поскольку в этом случае контроллер BSC должен найти ресурсы уже для *n* каналов. В общем, для постоянного информационного обмена (в частности, загрузки файла) система HSCSD представляет привлекательное промежуточное решение с более высокой пропускной способностью. Однако, для пакетного информационного обмена Internet она не подходит (пользователь должен оплачивать каждый канал, выделенный для связи).

4.1.8.2. GPRS

Проблем технологии HSCSD можно избежать в системе, полностью сориентированной на пакетную передачу. Поэтому следующим шагом в направлении более гибкой передачи данных стало создание **общей пакетной радиослужбы (GPRS)**. Она предоставляет пакетный режим приложениям, использующим частую передачу небольших объемов информации (запросы пользователей) или менее частую передачу в больших объемах (ответы Web-сервера). Требования к системе GPRS приведены в спецификации [ETSI, 1998a]. По сравнению с существующими службами передачи данных, в системе GPRS пакетные приложения используют сетевые ресурсы более эффективно. Пользователям службы обеспечивается возможность выбора параметров качества обслуживания. Кроме того, система GPRS предоставляет услуги широковещания, а также многоадресной и одноадресной передачи. Общая цель этого подхода — обеспечение более эффективной, а следовательно, и более дешевой службы с пакетным режимом передачи. Такая служба особенно необходима Internet-приложениям, которые обычно используют именно пакетную передачу. Сетевые поставщики услуг могут поддерживать эту модель, взимая оплату не за время соединения (как это делается сейчас в службах передачи данных системы GSM и технологии HSCSD), а за объем переданной информации. Очевидно, появление системы GPRS вызвано потрясающим успехом пакетно-ориентированной сети Internet, а также новыми приложениями и моделями информационного обмена. Однако, далее будет показано, что для технологии GPRS необходимы дополнительные сетевые элементы, т.е. программное и аппаратное обеспечение. Поэтому, в отличие от системы HSCSD, для объединения каналов в системе GPRS недостаточно одного лишь обновления программ.

Система GPRS основана на следующих идеях [ETSI, 1998б]: для новых радиоканалов GPRS система GSM может выделить от одного до восьми интервалов кадра TDMA. Они распределены не фиксированным образом, а управляются запросами. Все временны*е* интервалы делятся между активными пользователями, исходящие и нисходящие каналы рассматриваются отдельно. Распределение интервалов зависит от текущей загрузки и предпочтения оператора связи. В зависимости от кодирования скорость передачи может достигать 150 Кбит/с. В частности, если использовать кодирование каналов с пропускной способностью 14,4 Кбит/с, то при выделении всех временны*х* интервалов получится канал со скоростью передачи 115,2 Кбит/с. Однако, сама идея, положенная в основу системы GPRS, не зависит от характеристик и типа канала (будь это обычный канал информационного обмена или канал управления) и не ограничивает максимальную скорость передачи (ее ограничивает лишь транспортная система GSM). Все службы GPRS можно использовать параллельно с традиционными службами.

Система GPRS первого поколения предоставляет услуги **двухточечной (PTP)** пакетной передачи [ETSI, 1998в]. Одна из предлагаемых версий PTP — **сетевая служба PTP, ориентированная на установление соединения (PTP-CONS)**. Она предоставляет виртуальное устройство, необходимое для изменения ячеек в сети GSM. Этот тип службы соответствует всемирно известному протоколу X.25 — пакетно-ориентированному протоколу передачи с коммутацией каналов. Сетевая служба PTP, действующая **без установки соединения (PTP-CLNS)**, является еще одной версией технологии PTP. Она поддерживает приложения, основанные на протоколе IP. Многоадресная передача, также называемая **многоточечной (PTM)** службой, обеспечивается системой GPRS второго поколения.

Пользователи системы GPRS могут определять **профиль качества обслуживания**. Он содержит **приоритет службы** (высокий, средний, низкий), **класс надежности**, **класс задержки** передачи и **пропускную способность канала пользовательских данных**. Радиоресурсы распределяются так, чтобы удовлетворить этим пользовательским требованиям. Модули данных службы (SDU) могут дублироваться или следовать в неправильном порядке. Также может произойти потеря SDU или передача искаженных SDU на высшие уровни. Вероятности соответствующих процессов для разных классов надежности приводятся в табл. 4.3. Первый класс надежности необходим приложениям, весьма чувствительным к ошибкам и неспособным исправлять ошибки собственными силами. Класс 2 подходит для приложений, более устойчивых к ошибкам. Наконец, приложения, способные к коррекции ошибок или вообще нечувствительные к ним, могут использовать класс 3.

Таблица 4.3. Классы надежности технологии GPRS, установленные стандартом [ETSI, 1998в]

Класс надежности	Вероятность потери SDU	Вероятность дублирования SDU	Вероятность неправильного порядка модулей SDU	Вероятность искажения SDU
1	10^{-9}	10^{-9}	10^{-9}	10^{-9}
2	10^{-4}	10^{-5}	10^{-5}	10^{-6}
3	10^{-2}	10^{-5}	10^{-5}	10^{-2}

Кроме задержки доступа к каналу, в стационарной и беспроводной частях сети GPRS существуют задержки передачи. (Задержки во внешних стационарных сетях здесь не рассматриваются.) Буферизация пакетов в системе GPRS, в отличие от сетей с промежуточным хранением, не приводит к дополнительным задержкам. При использовании технологии GPRS пакеты переадресовываются настолько быстро, насколько это вообще возможно. В табл. 4.4 приводятся максимальные средние и 95-процентные значения задержек для пакетов с размерами 128 и 1024 байт. Легко заметить, что все задержки, вне зависимости от класса, на несколько порядков превышают соответствующие значения для стацио-

нарных сетей. Этот факт очень важно учесть при реализации в сетях GPRS протоколов высших уровней (например, протокола TCP), обсуждающихся в главе 10.

Таблица 4.4. Классы задержек технологии GPRS, установленные стандартом [ETSI, 1998в]

Класс задержки	Среднее значение (для SDU размером 128 байт)	95-процентное значение (для SDU размером 128 байт)	Среднее значение (для SDU размером 1024 байт)	95-процентное значение (для SDU размером 1024 байт)
1	< 0,5 с	< 1,5 с	< 2 с	< 7 с
2	< 5 с	< 25 с	< 15 с	< 75 с
3	< 50 с	< 250 с	< 75 с	< 375 с
4		не определено		

Система GPRS включает несколько **служб безопасности**. В ней проводится аутентификация, обеспечивается конфиденциальность пользователя и передаваемой им информации. Возможны даже полностью **анонимные услуги**, когда оплата взимается независимо от личности пользователя мобильной станции (как дорожный сбор).

Архитектура системы GPRS содержит новые элементы сети — так называемые **узлы поддержки GPRS (GSN)**. Все узлы GSN интегрированы в стандартную структуру сети GSM. Определено много новых интерфейсов (рис. 4.15). Взаимодействие сети GPRS с внешними **сетями с пакетной передачей (PDN)** осуществляется через **шлюзовый узел поддержки GPRS (GGSN)**. Этот узел содержит данные о маршрутизации, выполняет преобразования адресов и, с помощью инкапсуляции, туннелирует данные пользователям. Узлы GGSN связаны с внешними сетями (например, сетями IP или X.25) через интерфейс G_i. Передача пакетов узлам SGSN производится через магистральную сеть GPRS (интерфейс G_n), основанную на протоколе IP.

Еще одним новым элементом является **обслуживающий узел поддержки GPRS (SGSN)**, взаимодействующий с мобильной станцией через интерфейс G_b. Узлы SGSN, в частности, запрашивают адреса пользователей из **регистра GPRS (GR)** и отслеживают местонахождение отдельных мобильных станций. Они также отвечают за сбор информации об оплате и выполняют несколько функций, связанных с безопасностью (например, контроль доступа). Узлы SGSN взаимодействуют с контроллерами BSC посредством ретрансляции кадров. Они находятся примерно на том же иерархическом уровне, что и центры MSC. В регистре GR, обычно являющемся частью регистра HLR, хранятся все данные, относящиеся к системе GPRS.

На рис. 4.15 показано, как пакетные данные передаются из сети PDN через узлы GGSN и SGSN в подсистему BSS и, в конце концов, достигают мобильной станции. В обычной сети GSM с коммутацией каналов центр MSC отвечает за передачу данных. Однако, в системе GPRS он используется только для передачи сигналов. Дополнительные интерфейсы, необходимые для взаимодействия с другими сетевыми элементами и сетями PLMN, интересующийся читатель может найти в стандарте [ETSI, 1998б].

Любые данные до передачи через сеть GPRS должны быть закреплены за какой-то мобильной станцией. Это делается с помощью механизмов **управления мобильностью**. Процедура закрепления включает присвоение **временной идентификации логического канала (TLLI)** и **порядкового номера ключа шифрования (CKSN)**, используемого для кодирования данных. Для каждой станции MS устанавливается **окружение GPRS**. Оно сохраняется в мобильной станции и соответствующем узле SGSN. В его состав входит состояние мобильной станции (которая может быть свободной, готовой к работе или находиться в режиме ожидания [ETSI, 1998б]), номер CKSN и метка индикации сжатия. Окружение GPRS также содержит данные

маршрутизации (идентификацию TLLI, область маршрутизации RA, идентификаторы ячейки и канала пакетных данных PDCH). Кроме закрепления и отделения, управление мобильностью отвечает за аутентификацию, контроль местонахождения и шифрование (границы шифрования находятся между станцией MS и узлом SGSN, что расширяет возможности обычной системы GSM).

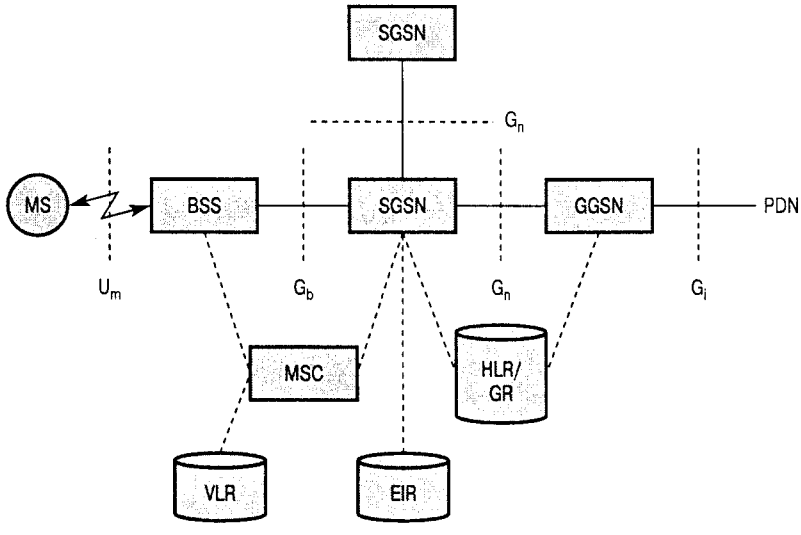

Рис. 4.15. Эталонная модель архитектуры GPRS

На рис. 4.16 показана архитектура протоколов матрицы передачи для системы GPRS. Структура сигнальных матриц описана в стандарте [ETSI, 1998б]. Все данные в магистральной сети GPRS, т.е. между узлами GSN, передаются с помощью **протокола туннелирования GPRS (GTP)**. Он может использовать два транспортных протокола — надежный протокол **TCP** (необходим для качественной передачи пакетов X.25) и менее надежный протокол **UDP** (применяется при передаче пакетов IP). В качестве сетевого протокола в магистральных сетях GPRS применяется протокол **IP** (использующий любой из нижних уровней). Между мобильной станцией и центром SGSN используется **зависящий от подсети протокол сближения (SNDCP)**. Его задача — согласование характеристик различных сетей. Над уровнями протоколов SNDCP и GTP производится туннелирование пользовательских пакетных данных от мобильной станции к узлу SGSN и обратно. Для достижения высокой надежности пакетной передачи между станцией MS и узлом SGSN применяется специальный протокол LLC. Он включает механизмы ARQ и FEC, необходимые для служб PTP (и позднее PTM).

Протокол подсистемы базовых станций GPRS (BSSGP) используется для передачи информации о маршрутизации и качестве обслуживания между подсистемой BSS и узлом SGSN. Этот протокол не выполняет коррекцию ошибок и действует над уровнем сети с **ретрансляцией кадров**. Наконец, для передачи данных через интерфейс U_m нужны протоколы, зависящие от радиоканала. **Протокол радиоканала (RLC)** обеспечивает надежное соединение, а протокол **MAC** управляет доступом с помощью сигнальных процедур и отображения кадров LLC в физические каналы системы GSM. Радиоинтерфейс U_m системы GPRS не требует фундаментальных изменений соответствующего интерфейса обычной системы GSM [Brasche, 1997], [ETSI, 1998г]. Однако, были определены некоторые новые логические каналы и соответствующие им физические ресурсы. К примеру, одной мобильной станции может выделяться до восьми **каналов обмена пакетными данными (PDTCH)**. Пропускную способность можно распределять со-

гласно запросам. Можно также разделить ее между каналами с коммутацией каналов и системой GPRS. Такое распределение может производиться динамически (с контролем загрузки) или быть фиксированным.

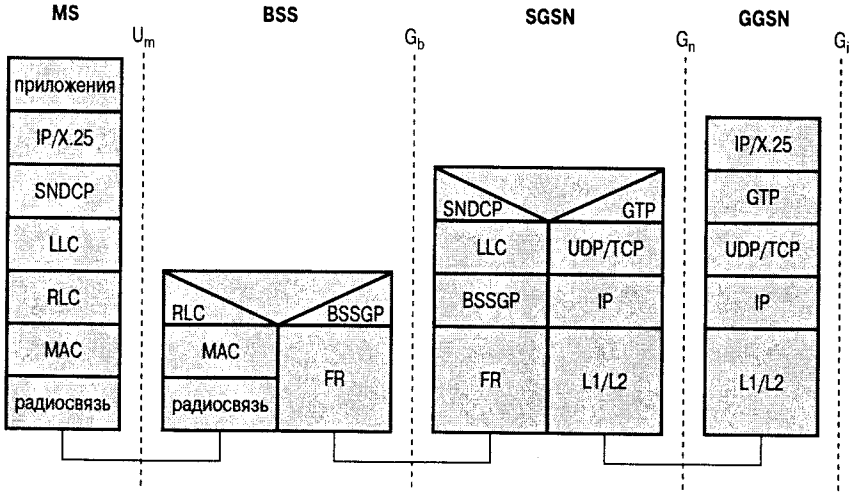

Рис. 4.16. Эталонная модель протоколов матрицы передачи в системе GPRS

Очень важным свойством сквозных приложений является то, что они не "замечают" никаких подробностей инфраструктуры GSM/GPRS. Пусть, к примеру, приложение использует протокол TCP над протоколом IP. Пакеты IP туннелируются на узел GGSN, а тот, в свою очередь, направляет их в сеть PDN. Все сети PDN переадресовывают свои пакеты (для пользователей системы GPRS) узлу GGSN. Узел GGSN запрашивает с узла SGSN параметры туннелирования и через него направляет пакеты мобильной станции.

4.2. Система DECT

Цифровые расширенные беспроводные коммуникации DECT (Digital Enhanced Cordless Telecommunications, ранее — Digital European Cordless Telephone и Digital European Cordless Telecommunications) — являются еще одной полностью цифровой сотовой сетью. Ее подробное описание приведено в стандартах [ETSI, 1999; 1998к, л], [DECT Forum, 1999]. Технология DECT заменяет старые аналоговые телефонные системы, в частности, системы CT1 и CT1+. Такие аналоговые системы не гарантируют полной безопасности, поскольку в них при передаче не производится шифрование данных. Они также характеризуются относительно низкой пропускной способностью. Возможности цифровой системы CT2, используемой главным образом в Великобритании, тоже не столь широкие (в сравнении со стандартом DECT, действующим в остальных европейских странах). В основном связь по технологии DECT используется в офисах и учебных заведениях, на аукционах и дома. Кроме того, точки доступа к сети PSTN могут устанавливаться на железнодорожных станциях, в больших государственных учреждениях и больницах. При этом услуги телефонной связи обходятся значительно дешевле, чем в системе GSM. Технологию DECT можно также применять на последних нескольких сотнях метров между новым оператором сети и ее абонентами. С помощью таких местных линий связи, действующих в небольшой области, новые компании могут предлагать свои услуги, даже не имея собственных установленных линий. В системах DECT существует много устройств сопряжения с сетями GSM, ISDN и сетями передачи данных. В настоящее время используется

более 20 миллионов устройств DECT. По этой технологии работает более 30 процентов всего оборудования беспроводных местных линий связи (WLL).

Системы GSM и DECT существенно отличаются диаметром и емкостью ячеек. Сеть GSM, в которой диаметр ячеек может достигать 70 км, предназначена для наружного использования. Зона действия технологии DECT ограничена 300-метровым расстоянием до базовой станции (в зданиях это расстояние уменьшается до 50 м из-за наличия стен). Вследствие такой ограниченной зоны действия и благодаря дополнительным технологиям уплотнения, система DECT может предоставлять свои услуги нескольким десяткам тысяч пользователей на квадратном километре. Эта ситуация типична для большого города, где в близко расположенных небоскребах находятся тысячи офисов. В сети DECT тоже существуют базовые станции, однако, их стоимость (вместе с мобильной станцией) составляет лишь несколько сотен евро. Для сравнения можно вспомнить, что базовая станция сети GSM стоит несколько десятков тысяч евро. В системе DECT тоже происходит переключение, но она не работает при высоких скоростях (например, при скорости 250 км/ч, когда система GSM еще функционирует).

Технология DECT действует в диапазоне частот 1880–1990 МГц и предоставляет 120 полнодуплексных каналов. В этой технологии применяются дуплексные линии с временным разделением (TDD), использующие кадры продолжительностью 10 мс. Диапазон частот с помощью схемы FDMA разделяется на 10 несущих частот. Каждый кадр разбивается на 24 временных интервала согласно схеме TDMA. В механизме TDD 12 интервалов используются для исходящих каналов и 12 — для нисходящих. Для осуществления цифровой модуляции применяется схема GMSK. Максимальная мощность передачи каждой станции равна 250 мВт, а средняя мощность составляет лишь 10 мВт.

4.2.1. Архитектура системы

Существует много различных вариантов физической реализации системы DECT, зависящих от ее конкретного применения. Различные объекты сети DECT могут быть интегрированы в одном физическом блоке. Они могут также дублироваться или быть рассредоточенными. Однако, в основе всех реализаций находится одна и та же логическая эталонная модель архитектуры системы, изображенная на рис. 4.17. **Глобальная сеть** соединяет местные структуры связи с внешним миром и предоставляет свои услуги через интерфейс D_1. В качестве глобальных сетей могут использоваться цифровые сети с интеграцией услуг (ISDN), коммутируемые телефонные сети общего пользования (PSTN), мобильные сети общего пользования (PLMN) (к ним относится и система GSM), а также сети общего пользования с коммутацией пакетов (PSPDN). Услуги, предлагаемые этими сетями, включают передачу данных, преобразование адресов и маршрутизацию данных между локальными сетями.

Локальные сети (в рассматриваемом контексте) отвечают за локальные телекоммуникационные услуги. Диапазон таких услуг достаточно широкий — от простой коммутации до программируемого перенаправления звонков и преобразования адресов. В качестве примеров локальных сетей можно взять аналоговые и цифровые телефонные сети частного пользования (PBX) или сети ЛВС, в частности, семейство ЛВС стандарта IEEE 802.x. Поскольку система DECT сама по себе является очень простой, все обычные сетевые функции интегрируются в локальные или глобальные сети. Там же размещается **база данных об исходном положении (HDB)** и **база данных посетителей (VDB)**. Обе базы данных поддерживают мобильность и выполняют примерно те же функции, что и регистры HLR и VLR в системе GSM. Входящие звонки автоматически направляются в ту из подсистем, которая отвечает за нужного абонента DECT. Текущая база данных VDB информирует базу HDB об изменениях местонахождения пользователя.

Внутренняя сеть системы DECT состоит из **стационарных оконечных радиоустройств (FT)** и **портативных оконечных радиоустройств (PT)**. Ее основная задача — обеспечение уплотнения. Устройства FT и PT охватывают уровни с первого по третий в стационарной и мобильной частях сети.

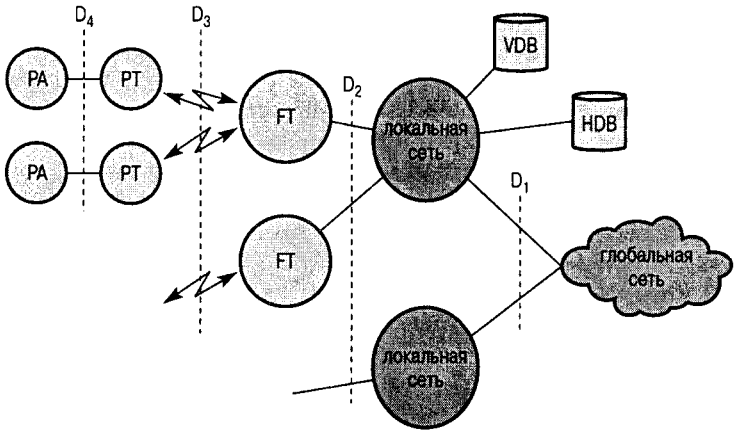

Рис. 4.17. Эталонная модель архитектуры системы DECT

4.2.2. Архитектура протоколов

Архитектура протоколов системы DECT подчиняется эталонной модели OSI. Охватываемые стандартом уровни приведены на рис. 4.18. Они включают физический уровень, а также уровни управления каналом и доступа к среде. Уровни управления каналом для **матрицы управления (C-матрицы)** и для **пользовательской матрицы (U-матрицы)** отличаются между собой. Для C-матрицы определен дополнительный сетевой уровень. С его помощью пользовательские данные с уровня 2 направляются сразу на U-матрицу. Управление охватывает все нижние уровни системы DECT.

4.2.2.1. Физический уровень

В системе DECT, как и во всех беспроводных сетях, **физический уровень** отвечает за проведение модуляции/демодуляции, детектирование поступающего сигнала и синхронизацию приемника с передатчиком. На физическом уровне также осуществляется сбор информации о состоянии для матрицы управления. Кроме того, этот уровень генерирует структуру физического канала с определенной гарантированной пропускной способностью. Канал для передачи данных выделяется физическим уровнем после запроса с уровня MAC.

На рис. 4.19 показаны типичные пакеты данных и стандартная структура кадров TDMA, используемая в системе DECT. Продолжительность каждого кадра составляет 10 мс. В одном кадре содержится 12 интервалов для исходящих каналов и 12 — для нисходящих (в режиме **базового соединения**). Если мобильный узел получает данные в интервале s, то обратная передача происходит в интервале $s + 12$. В режиме **расширенного соединения** доступны и другие схемы распределения. Каждый интервал имеет продолжительность 0,4167 мс и может содержать несколько разных физических пакетов. Обычно данные занимают 420 бит, а оставшиеся 52 мкс используются как **защитные интервалы**. 420 бит данных, в свою очередь, включают 32-битную **синхронизирующую последовательность** и поле **данных** D.

Поля передачи данных используют оставшиеся 388 бит для **управления сетью** (поле A), а также для пересылки **пользовательских данных** (поле B) и **качества передачи** (поле X). Данные управления сетью передаются со скоростью 6,4 Кбит/с (64 бит за 10 мс). В то же время скорость передачи пользовательских данных зависит от дополнительных механизмов коррекции ошибок. **Симплексные носители в незащищенном режиме** обеспечивают скорость передачи 32 Кбит/с. Использование в **защищенном режиме** 16-битовой **контрольной суммы** CRC (C) для блока данных размером 64 бит приводит к уменьшению скорости до 25,6 Кбит/с. Объединив два симплексных носителя, можно получить **дуплексный носитель**. В технологии DECT определены и другие типы носителей. В них более высокая

скорость передачи достигается за счет объединения интервалов. К примеру, **двойной дуплексный носитель** предоставляет полнодуплексный канал с пропускной способностью 80 Кбит/с.

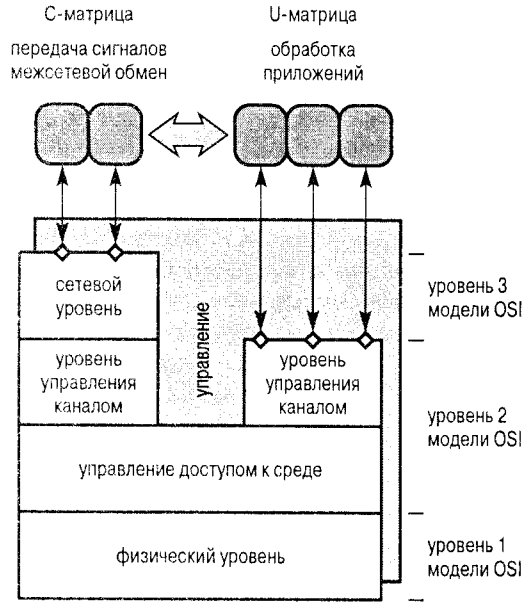

Рис. 4.18. Уровни протоколов системы DECT

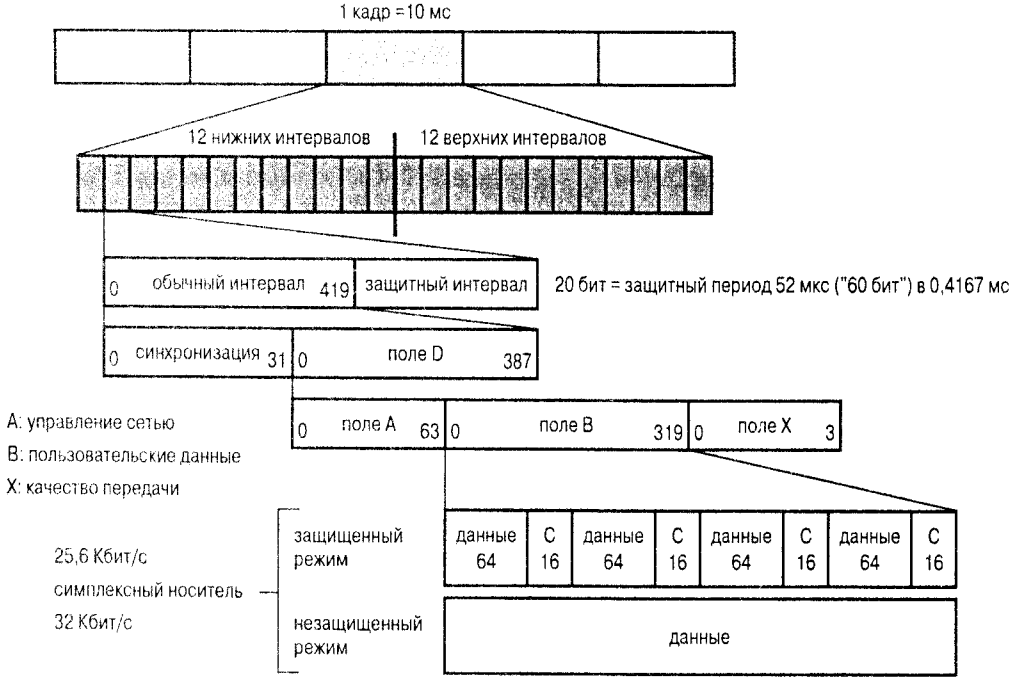

Рис. 4.19. Структура кадров и уплотнение в системе DECT

4.2.2.2. Уровень управления доступом к среде

Уровень **управления доступом к среде (medium access control — MAC)** устанавливает, поддерживает и освобождает каналы высших уровней с помощью активации и деактивации физических каналов. Кроме того, уровень MAC отвечает за уплотнение нескольких логических каналов в один физический канал. Логические каналы необходимы для управления сетью передачи сигналов, поиска мобильных станций, а также для передачи пользовательских данных и широковещательных сообщений. Дополнительные услуги включают сегментацию/дефрагментацию пакетов и защиту от ошибок/коррекцию ошибок.

4.2.2.3. Уровень управления каналом

Уровень **управления каналом (data link control — DLC)** должен создавать и поддерживать надежные соединения мобильных терминалов с базовой станцией. Для **C-матрицы** определено две службы — **широковещательная служба** поиска мобильных станций без организации соединений (называемая **Lb**) и **двухточечный** протокол **LAPC+Lc**. Этот протокол выполняет почти те же функции, что и протокол LAPD в сетях ISDN, но приспособлен к схеме MAC системы DECT.

Для **U-матрицы** тоже существует несколько служб. Среди них — сквозная незащищенная передача (по существу, это нулевая служба), прямое исправление ошибок, службы дальнейших усовершенствований и регулирование скорости передачи. Если данные передаются через сеть ISDN, например, со скоростью 64 Кбит/с, то сначала система DECT устанавливает такую же скорость 64 Кбит/с. Однако, при большом количестве ошибок система DECT увеличивает скорость до 72 Кбит/с, включает схему FEC и помещает в буфер до восьми блоков (для использования запроса ARQ). Это приводит к дополнительной задержке, которая может достигать 80 мс.

4.2.2.4. Сетевой уровень

Сетевой уровень системы DECT напоминает соответствующие уровни в сетях ISDN и GSM и существует только для C-матрицы. Этот уровень отвечает за запросы, проверку, резервирование, контроль и освобождение ресурсов стационарной станции (соединение со стационарной сетью, беспроводное соединение) и мобильного терминала (беспроводное соединение). Протокол **управления мобильностью (MM)** производит аутентификацию, а также управляет идентификацией и базами данных о местонахождении. Протокол **управления вызовами (CC)** необходим для переговоров, установки и отключения соединений. Для обмена данными с межсетевым устройством, соединяющим систему DECT с внешним миром, используются две **службы сообщений**, действующие **на основе соединений (COMS)** и **без организации соединений (CLMS)**.

4.3. TETRA

Еще одним средством беспроводной передачи данных являются магистральные радиосистемы. Они используют много различных несущих частот. Определенная частота присваивается пользователю лишь на короткий промежуток времени. Распределение частот управляется запросами. В обычных системах службам спасения, транспортным компаниям с управляемым передвижением, такси и т.д. выделяются отдельные несущие частоты. В магистральных радиосистемах упомянутые службы могут работать совместно в одном и том же диапазоне. С помощью такого механизма обеспечивается более эффективное многократное использование частот в технологиях FDM и TDM. Хотя радиомагистрали и не являются сетями общего пользования, в большинстве случаев в них существуют интерфейсы для взаимодействия со стационарной телефонной сетью, т.е. службы передачи речи и данных. Установка и эксплуатация этих систем обходится относительно дешево, поскольку им не нужно покрывать целые страны. Они должны действовать лишь в тех областях, где находятся локальные пользователи (например, городские службы такси). К тому же радиомагистрали значительно проще других сетей.

В 1991 году институтом ETSI был разработан стандарт общеевропейской системы **TETRA** (terrestrial trunked radio — наземная радиомагистраль)[3] [ETSI, 1999], [TETRAMoU, 1999]. Она предназначалась для замены систем MODACOM, MOBITEX и COGNITO, которые действовали в отдельных европейских государствах и в большинстве случаев соединялись с пакетными сетями X.25. (В США подобные функции выполняла система ARDIS.) В системе TETRA существует два типа служб — служба **Речь+Данные (V+D)** [ETSI, 1998м] и служба **с оптимизацией пакетов (PDO)** [ETSI, 1998н]. В то время как служба V+D осуществляет передачу (с коммутацией каналов) речи и данных, в службе PDO предусмотрена лишь пакетная передача, действующая либо на основе соединений с сетями X.25, либо без организации соединений (для сетевой службы ISO CLNS). Последняя служба может быть двухточечной или многоточечной. Типичные значения задержки для коротких сообщений (128 бит) не превышают 100 мс. Соединения V+D включают одноадресную и многоадресную передачу, многосторонние соединения отдельной защищенной группы пользователей, а также прямую эпизодическую связь без участия базовых станций. Однако, в зависимости от приоритета, задержки коротких сообщений здесь могут достигать 500 мс и больше.

В системе TETRA существуют и службы-носители. Их пропускная способность при незащищенной передаче данных равна 28,8 Кбит/с, а при защищенной — 9,6 Кбит/с. Примерами сквозных служб являются перенаправление, исключение и захват звонков. Сюда же можно отнести идентификацию, объединение групп, службы экстренных вызовов и приоритетные звонки. Системы TETRA и GSM имеют похожую структуру. Интерфейс U_m соединяет **мобильную станцию (MS)** с **инфраструктурой управления и коммутации (SwMI)**, содержащей пользовательские базы данных (HDB, VDB), базовую станцию и интерфейсы к сетям PSTN, ISDN и PDN. Однако, систему TETRA реализовать намного проще, поскольку в ней нет необходимости в переключении. Автомобили такси большую часть времени находятся внутри некоторой области, которую можно охватить одной ячейкой TETRA.

Для систем TETRA с дуплексом FDD выделено несколько диапазонов частот (в частности, 380–390 МГц/410–420 МГц для исходящих каналов и 390–400 МГц/420–430 МГц — для нисходящих). Каждый канал характеризуется пропускной способностью 36 Кбит/с и полосой частот шириной 25 кГц. Для модуляции применяется схема DQPSK. В службе V+D на одной несущей передается до четырех каналов TDMA (речь или данные). Служба PDO использует статистическое уплотнение. Доступ к каналам производится по схеме Aloha с выделением временных интервалов.

На рис. 4.20 показана **структура кадров TDMA** в системе TETRA. Каждый **кадр** состоит из четырех интервалов (именно поэтому в службе V+D на одной несущей передается четыре канала). Продолжительность кадра равна 56,67 мс. В каждом **интервале** переносится 510 бит за 14,17 мс, т.е. скорость передачи составляет 36 Кбит/с. Объединение 16 кадров с **кадром управления (CF)** образует **мультикадр**. Наконец, **гиперкадр** состоит из 60 мультикадров. Чтобы избежать одновременной передачи и приема, в системе TETRA исходящий канал сдвинут относительно нисходящего на два интервала.

Система TETRA, как и GSM, предусматривает **каналы информационного обмена (TCH)** и **каналы управления (CCH)**. Обычно для передачи речи используются каналы TCH/S, а для передачи данных — каналы TCH/7.2, TCH/4.8 или TCH/2.4 (в зависимости от применяемых схем FEC).

---

[3] Ранее аббревиатура TETRA расшифровывалась как "trans European trunked radio" (трансевропейская радиомагистраль), но в связи с выходом на международный рынок была принята нынешняя трактовка этого названия (ср. с эволюцией расшифровки аббревиатуры DECT).

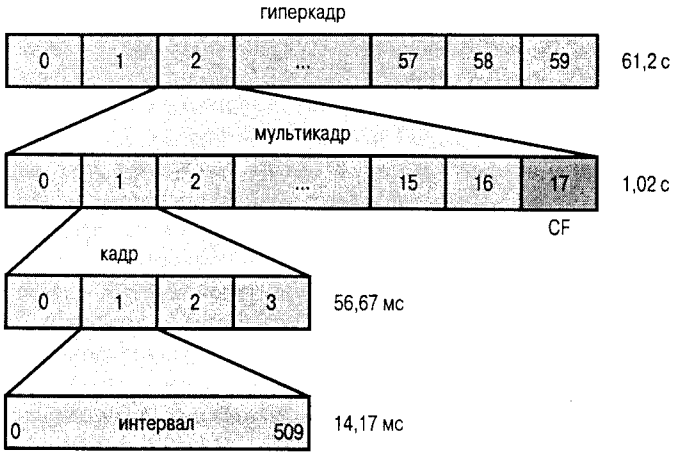

Рис. 4.20. Структура кадров в системе TETRA

Система TETRA предоставляет некоторые услуги, недоступные в системе GSM. Среди них — обычное и подтверждаемое групповое соединение, многоадресная передача и выборочное прослушивание. Аварийным службам необходимо устанавливать групповые соединения в жестких условиях, а иногда — при отсутствии всей инфраструктуры. Этого не позволяет ни система GSM, ни другие обычные сети мобильной телефонной связи.

4.4. UMTS и IMT-2000

Международный телекоммуникационный союз (International Telecommunication Union — ITU) объявил конкурс технологий радиопередачи (RTT) для программы "Международные мобильные телекоммуникации (IMT) 2000" [ITU, 1999], [Callendar, 1997], [Shafi, 1998]. Система IMT-2000 раньше называлась системой будущих мобильных телекоммуникаций общего пользования (FPLMTS). В ней пытались установить такую всемирную систему связи, которая бы поддерживала идею универсальных личных соединений (UPT) и обеспечивала мобильность терминалов и пользователей. Союзом ITU были разработаны некоторые рекомендации для систем FPLMTS, касающиеся, в частности, сетевой архитектуры (M.817), требований к радиоинтерфейсу (M.1034) и структуры служб FPLTMS (M.816). Число 2000 в названии IMT-2000 обозначает приблизительный год начала эксплуатации системы (около 2000) и используемый спектр (полоса частот около 2000 МГц). В программе IMT-2000 определены различные среды: использование внутри зданий и автомобилей, применение спутников и т.д. На конференции WRC в 1992 году для новых всемирных систем IMT-2000 были выделены диапазоны частот 1885–2025 МГц и 2110–2200 МГц (рекомендация института ITU-R M.1036).

В 1998 году было предложено несколько вариантов RTT для использования внутри и вне зданий. Эти предложения могут использовать пешеходы и автомобилисты. Кроме того, в них предусмотрена и возможность спутниковой связи. Предложения поступали от разных организаций. К примеру, консорциум универсальных беспроводных коммуникаций (США) разработал на базе стандарта IS-136 систему третьего поколения **UWC-136**. Стандарт IS-95 (США) был положен в основу системы **cdma2000**. Можно упомянуть и широкополосную сеть CDMA с коммутацией пакетов (**WP-CDMA**), совместимую с предложенной в Европе системой UTRA. Вообще предложения выдвигаются организациями трех больших регионов. В Европе эти функции выполняет институт ETSI, в США — институт ANSI (Национальный институт стандартизации). В Японии существует сразу две организации — ассоциация радиоиндустрии и

широковещания (ARIB) и совет телекоммуникационных технологий (TTC). Союз ITU еще не принял окончательного решения относительно возможных значений RTT, но продолжающиеся дискуссии носят скорее политический, чем технический характер.

Система IMT-2000, предложенная европейским институтом ETSI, называется **Универсальной мобильной телекоммуникационной системой (UMTS)** [Dasilva, 1997], [Ojanperä, 1998]. Для RTT радиоинтерфейса отдельно была разработана схема **наземного радиодоступа UMTS (UTRA)** [ETSI, 1998o], [UMTS Forum, 1999]. Система UMTS не является полностью новой. Она скорее отражает переход от второго поколения систем GSM к третьему поколению. Такой подход оправдан экономически — введение новой системы обойдется дороже, чем усовершенствование существующей. Выдвинуто несколько решений, обеспечивающих плавный переход между системами GSM и UMTS [GSMMoU, 1998], [Nilsson, 1998]. Одно из предложений включает **ускоренную передачу данных для развития GSM (EDGE)**. С помощью расширенных схем модуляции и других технологий здесь в одном канале достигается скорость передачи 384 Кбит/с (т.е. 48 Кбит/с на один временной интервал). При этом используются те же частоты, что и в системе GSM. Несущая, как и раньше, характеризуется шириной 200 кГц. Технологию EDGE можно вводить постепенно, предоставляя в некоторых каналах возможность коммутации между системами EDGE и GSM/GPRS. Кроме увеличения скорости передачи, можно также обеспечить поддержку интеллектуальных сетей. Эти функции выполняют **индивидуализированные приложения с мобильной расширенной логикой (CAMEL)**. С их помощью можно, например, создавать для всех абонентов-посетителей **виртуальную исходную среду (VHE)**. Меморандум [GSMMoU, 1999] включает много предложений относительно роуминга, ведения счетов, служб, сетей доступа, внутренних сетей и безопасности. В нем также рассматриваются некоторые вопросы радиосвязи, качества обслуживания, эксплуатации системы и ее расширения. Кроме того, обсуждаются требования к терминалам и области применения системы.

Система UMTS входит в состав более широкой структуры, разработанной институтом ETSI, которая называется **Глобальной мультимедийной мобильностью (GMM)**. Эта структура объединяет мобильные и стационарные **терминалы**, различные **сети доступа** (GSM BSS, DECT, ISDN, UMTS, LAN, WAN, CATV, MBS) и несколько **внутренних транспортных сетей** (GSM NSS+IN, ISDN+IN, B-ISDN+TINA, TCP/IP) [ETSI, 1999]. Институтом ETSI были выдвинуты **основные требования** к системе UMTS и радиоинтерфейсу UTRA [ETSI, 1998s]. Среди ключевых требований — минимальная пропускная способность 144 Кбит/с при внешнем доступе из сельской местности (хотя конечной целью является 384 Кбит/с), а также функционирование при скорости 500 км/ч. В пригородной зоне пропускная способность должна составлять не менее 384 Кбит/с (цель — 512 Кбит/с) при скорости 120 км/ч. Внутри зданий и в городе, когда расстояния до базовых станций небольшие, нужно обеспечить пропускную способность 2 Мбит/с при скорости 10 км/ч (ходьба).

Кроме того, система UMTS должна предоставлять несколько служб-носителей и служб, действующих в режиме реального времени. Необходимо обеспечить возможность передачи данных, использующей коммутацию каналов или пакетов, с различной скоростью. Переключение должно происходить не только между ячейками системы UMTS, а и между ячейками сети GSM и спутниковых систем. Система должна быть совместимой с сетями GSM, ATM, IP и ISDN. Чтобы учесть типичный асимметричный характер соединений, соотношение между скоростями передачи в исходящем и нисходящем каналах UMTS можно сделать переменным. И наконец, система UMTS должна соответствовать структуре IMT-2000 (для успеха системы этот фактор, возможно, является определяющим). Поскольку проект глобальной системы UMTS весьма амбициозен, на начальных этапах альтернативой ему могут стать городские ячейки UMTS, предоставляющие некоторые из перечисленных услуг.

Несколько компаний и заинтересованных групп подали свои предложения радиоинтерфейса UTRA [ETSI, 1998u]. В январе 1998 года два из них были выбраны институтом ETSI для

применения в системе UMTS. Для **спаренной полосы** (дуплекс FDD) будет использоваться широкополосная система CDMA (**W-CDMA**). Для **неспаренной полосы** (дуплекс TDD) предназначена система CDMA с временным разделением (**TD-CDMA**) [Adachi, 1998], [Dahlman, 1998], [ETSI, 1998o]. Спаренная полоса обычно используется поставщиками мобильных сетей общего пользования (к ним относится и сеть GSM), а неспаренная полоса предназначается для местной связи и связи внутри зданий (система DECT). В следующих разделах представлены ключевые свойства обоих компонентов системы UMTS. Существуют нерешенные проблемы с патентами на определенные технологии CDMA. Поэтому до сих пор неясно, будут ли две описанные технологии использоваться в окончательной системе.

4.4.1. Основы архитектуры системы UMTS

На рис 4.21 приводится упрощенная схема структуры системы UMTS. Оба решения радиоинтерфейса UTRA построены согласно этой схеме. **Сеть UTRA (UTRAN)** обеспечивает мобильность на уровне ячеек и включает несколько **сетевых радиоподсистем (RNS)**. Они осуществляют шифрование и дешифрование радиоканала, управляют переключением и распределяют радиоресурсы. Сеть UTRAN связана с **пользовательским оборудованием (UE)** через интерфейс U_u (его можно сравнить с интерфейсом U_m в системе GSM). Через интерфейс I_u (в системе GSM ему соответствует интерфейс A) сеть UTRAN соединяется с **внутренней сетью (CN)**. Внутренняя сеть осуществляет межсистемное переключение. В ее состав входят шлюзы к другим (стационарным или беспроводным) сетям. Хотя основные требования к интерфейсу I_u уже определены, сам интерфейс пока еще не стандартизован.

В телекоммуникационных стандартах рассматриваемого типа сигнальные и пользовательские данные обычно разделяются. С этой целью в системе UMTS вводятся C-матрица (управление) и U-матрица (пользовательские данные). Их архитектура включает три нижних уровня. Уровень радиоинтерфейса (уровень 1) отвечает за проведение уплотнения и разуплотнения, модуляцию, синхронизацию и расширение. На этом уровне также производится контроль мощности и осуществляется прямое исправление ошибок. Уровень 2 отвечает за соединения MAC, управление приоритетами и доступом к среде, разрешение конфликтов и диспетчеризацию пакетов данных. Наконец, все функции управления, необходимые для установки соединений, относятся к третьему уровню.

Рис. 4.21. Упрощенная структура системы UMTS

4.4.2. Режим UTRA FDD

В режиме FDD системы UTRA используется **широкополосная схема CDMA (W-CDMA)** с расширением методом прямой последовательности. Применение дуплекса FDD подразумевает, что исходящие и нисходящие линии связи действуют на различных частотах. Мобильные станции посылают данные через исходящие каналы на несущих частотах из диапазона 1920–1980 МГц. Для нисходящего канала базовая станция использует полосу частот между 2110 МГц и 2170 МГц. В этом режиме для пользовательского информационного обмена, т.е. передачи речи, выделяется около 250 каналов. Каждому физическому каналу соответствует некоторая несущая частота, код расширения (DS) и относительная фаза (только для исходящих каналов). Пропускная способность одного канала может достигать 2 Мбит/с.

Как и в системе GSM, здесь определено несколько физических и логических каналов. Сначала рассмотрим исходящие линии связи. Пользовательские данные (начиная со второго уровня и выше) передаются через **выделенный физический исходящий канал данных (исходящий канал DPDCH)**. **Выделенный физический исходящий канал управления (исходящий канал DPCCH)** предназначен для передачи данных управления (в частности, данных контроля мощности) на первом уровне. Оба канала передаются параллельно, причем в первом случае

(DPDCH) используется синфазная составляющая сигнала (см. главу 2), а во втором (DPCCH) — квадратурная. Для обоих каналов возможны различные коэффициенты расширения. Через **физический канал случайного доступа (PRACH)** передаются данные управления мобильных станций, необходимые для случайного доступа. В этом канале применяется схема Aloha с выделением временных интервалов, использующая интервалы продолжительностью 1,25 мс. О том, какие из интервалов случайного доступа в данное время свободны, мобильные станции внутри ячейки узнают через общий широковещательный канал.

Выделенный физический нисходящий канал (нисходящий канал DPCH) используется для передачи пользовательских данных и данных управления с первого уровня. При этом используется временное уплотнение. В схеме UTRA определены дополнительные общие физические каналы управления радиосвязью. На рис. 4.22 показана структура кадров системы UTRA с дуплексом FDD, а также некоторые из перечисленных каналов. **Суперкадр** состоит из 72 кадров и длится 720 мс. **Кадр** продолжительностью 10 мс включает 16 интервалов. Приведенный на рис. 4.22 пример иллюстрирует использование интервала тремя различными каналами. Исходящий канал DPCCH содержит **идентификатор транспортного формата (TFI)**, **контрольный сигнал**, необходимый для оценки качества канала, а также данные **управления мощностью** передачи (**TPC**). Исходящий канал DPDCH, используемый параллельно с каналом DPCCH, предназначен для пользовательских **данных**. В состав нисходящего канала DPCH входят те же самые поля (т.е. управляющий сигнал, идентификатор TFI, данные TPC и пользовательские данные), уплотненные по времени в одном интервале.

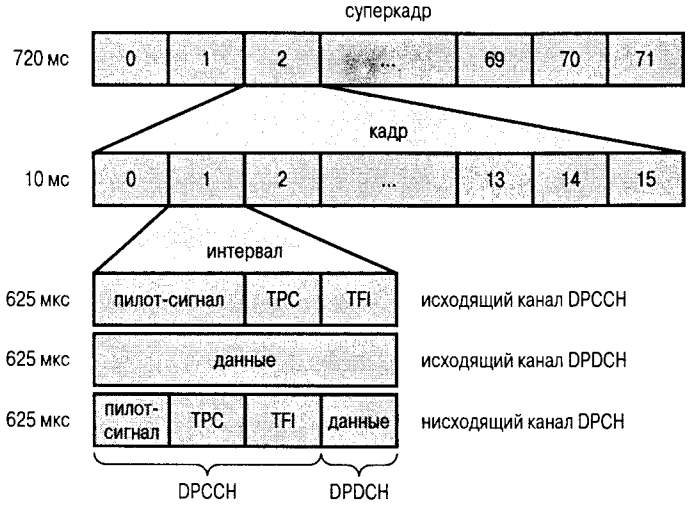

Рис. 4.22. Структура кадров в системе UMTS (режим FDD)

Скорость передачи элементарных сигналов в рассматриваемой системе составляет $4{,}096 \times 10^6$ элементарных сигналов/с. В будущих приложениях она может быть увеличена до $8{,}192 \times 10^6$ элементарных сигналов/с или $16{,}384 \times 10^6$ элементарных сигналов/с. Для модуляции битов данных применяется схема DQPSK. При использовании схемы CDMA возможно плавное переключение базовых станций [Wong, 1997] и даже "более плавное переключение" между различными секторами одной базовой станции. Кроме того, в этой схеме положение мобильных станций определяется с точностью до 18 м. Однако, для систем CDMA необходимы сложные схемы регулирования мощности, оперирующие со скоростью 1 600 тактов в секунду. Передатчик и приемник на протяжении соединения работают непрерывно.

4.4.3. Режим UTRA TDD

В режиме TDD системы UTRA для доступа к среде применяется **широкополосная схема TDMA/CDMA**. В этом режиме исходящий и нисходящий каналы работают на одинаковой частоте. Здесь предоставляется около 120 ортогональных каналов пользовательского информационного обмена со скоростью передачи 2 Мбит/с. Это составляет примерно половину пропускной способности режима FDD. Нужно помнить, однако, что в схеме TDD используется лишь половина соответствующей полосы частот. Физические каналы характеризуются интервалами и кодами расширения (метод прямой последовательности). Обычно в одном интервале содержится восемь пакетов различных каналов. На рис. 4.23 показана типичная структура пакетов внутри отдельного интервала (**пакеты информационного обмена**). Такие пакеты состоят из пользовательских **данных** и **промежуточных интервалов**, предназначенных для оценки качества канала. **Защитные периоды** (**GP**) с длительностью 23,4 мкс необходимы для того, чтобы различные интервалы не перекрывались. Каждый **интервал** содержит 2 560 элементарных сигналов и продолжается 625 мкс. 16 интервалов образуют **кадр** продолжительностью 10 мс.

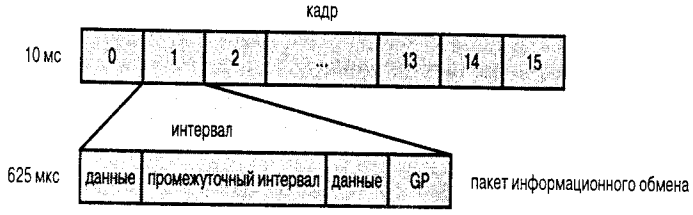

Рис. 4.23. Структура кадров в системе UMTS (режим TDD)

Кадры TDD могут быть **симметричными** или **асимметричными**. Иначе говоря, в состав кадра может входить одинаковое количество "исходящих" и "нисходящих" интервалов или произвольная их комбинация. Это отражает различные нужды пользователей в отношении скорости передачи. Кроме того, кадр может содержать одну или несколько **точек коммутации** исходящего и нисходящего каналов. Однако, для каждого из каналов необходимо выделить хотя бы по одному интервалу. Коэффициент расширения зависит от желаемого значения скорости передачи и может регулироваться системой.

В режиме TDD для модуляции используется схема QPSK. Регулирование мощности можно производить не столь часто (здесь достаточно 100–800 циклов регулирования в секунду). Передатчик и приемник работают лишь на протяжении выделенных им интервалов, т.е. не беспрерывно. Для успешного переключения необходима жесткая синхронизация в небольшой области (несколько ячеек). Это обусловлено применением временного уплотнения. В описанной схеме, как и в системе GSM, допускается только жесткое переключение.

Обе схемы — FDD и TDD — могут сосуществовать. Похожая структура кадров и их одинаковая продолжительность (10 мс) позволяют переключаться между этими схемами. К тому же мультикадровая структура режимов UTRA позволяет производить прямые переключения между системами UMTS и GSM.

4.5. Резюме

Большая часть этой главы была посвящена GSM — наиболее успешной цифровой сотовой системе. Исходно технология GSM предназначалась для передачи речи. Несмотря на это, в данной главе рассмотрена эволюция GSM в направлении передачи данных, осуществляемой через системы HSCSD и GPRS. Эта эволюция, кроме всего прочего, включает переход от сетей с комму-

тацией каналов к системам с коммутацией пакетов и, таким образом, сближает систему GSM с моделью Internet. В данной главе также представлены DECT — цифровой стандарт для беспроводных телефонов и TETRA — магистральные радиосистемы. Технологию DECT можно использовать для беспроводной передачи данных внутри зданий, в общежитиях, а также для местных беспроводных линий связи (WLL). Радиомагистрали TETRA применяются, в частности, в аварийных ситуациях. Они обеспечивают быструю установку соединения (даже группового) и могут работать как эпизодические сети, т.е. при отсутствии базовой станции.

В США состояние сферы мобильных телекоммуникаций несколько отличается от ситуации в Европе. Американской промышленностью была разработана система TDMA IS-54, основанная на аналоговой системе AMPS. Система IS-54 предоставляет, кроме аналоговых, и цифровые каналы информационного обмена. В ней используются двухрежимные мобильные телефоны. В этой технологии внедрена некоторая часть архитектуры системы GSM, например, ассоциированные каналы управления, процедуры аутентификации и шифрования, мобильное автоматическое переключение. Многие идеи, используемые в японской системе PDC, заимствованы из системы IS-54.

Следующая система, IS-136, является более эффективной. В ней используются цифровые каналы управления (в системе IS-54 использовались аналоговые каналы управления сети AMPS). Стало возможным применение полностью цифровых мобильных телефонов, появилось несколько новых служб. В качестве примеров можно привести речевую почту, ожидание звонка, идентификацию, групповой вызов, службу SMS. Сеть IS-136 иногда называют Североамериканской системой TDMA (NA-TDMA) или цифровой системой AMPS (D-AMPS). Она работает в диапазонах частот около 800 МГц и 1 900 МГц. Усовершенствования сети D-AMPS/IS-136 в направлении системы IMT-2000 включают расширенные технологии модуляции на несущей шириной 30 кГц и увеличение скорости передачи до 64 Кбит/с (первый этап, 136+). На втором этапе, 136HS (High Speed), должен появиться новый радиоинтерфейс, основанный на технологии EDGE.

В основу системы IS-95 положена схема CDMA — совершенно иная технология доступа к среде. До внедрения этой системы предполагали, что она превзойдет системы TDMA по многим параметрам. В частности, считалось, что количество пользователей в одной ячейке системы IS-95 может в 20 раз превышать емкость ячейки в системе AMPS. Сейчас поставщики систем CDMA делают более реалистические оценки (пятикратное увеличение емкости). В системе IS-95 обеспечивается гибкое переключение, позволяющее избежать эффекта пинг-понга, характерного для системы GSM [Wong, 1997]. С другой стороны, в системе IS-95 все базовые станции должны быть синхронизированы. Этого можно добиться с помощью спутников системы GPS. Однако, они являются военными спутниками и не контролируются сетевыми поставщиками услуг. Кроме того, в системе IS-95 нужно часто производить контроль мощности. Еще один ее недостаток — неполное покрытие — приводит к необходимости использования двухрежимных мобильных телефонов. Превосходство систем CDMA над системами TDMA до сих пор не доказано. Тем не менее, пример системы UMTS показывает, что основные идеи схемы CDMA будут интегрироваться во многие будущие системы.

Вопрос о том, какая из систем будет использоваться в мобильных телефонных сетях третьего поколения, все еще остается открытым. Система UMTS, представленная в этой главе, является одним из предложений. Это предложение на данный момент поддержано Европой и Японией. Какими будут системы будущего, обеспечивающие связь с любой точкой земного шара и плавное переключение между различными системами? К сожалению, это скорее политический, чем технический вопрос. В одной из возможных систем различные подходы объединены [Nilsson, 1998]. Здесь в качестве общей основы используется сеть Internet. Мобильные устройства являются двух- или трехрежимными и поддерживают, например, системы GSM и IS-136 или UMTS и CDMA2000. Каждая мобильная сеть содержит центры MSC и ядро коммутации пакетов различных сетей (например, систем

GSM и UMTS или сети D-AMPS и соответствующего решения IMT-2000). Для связи используется Internet и существующие телефонные сети. Регистры местонахождения должны взаимодействовать между собой через межсетевой регистр местонахождения (ILR).

4.6. Упражнения

1. Назовите несколько ключевых особенностей систем GSM, DECT и TETRA. Какие общие свойства этих систем? Почему определено три разных системы? В каких ситуациях они могут заменять друг друга? Какими специфическими преимуществами обладает каждая система?
2. Какие основные проблемы возникают при передаче данных в системах, предназначенных для голосовой передачи? Какие шаги можно предпринять для избежания этих проблем и увеличения эффективности? Как это может отразиться на оплате?
3. Какие типы служб предоставляет система GSM? Приведите несколько примеров и назовите причины разделения этих служб.
4. Радиоинтерфейс стандартной системы GSM в сравнении с предлагаемыми каналами TCH может обеспечить более высокую пропускную способность (33,8 Кбит/с). За счет чего уменьшается скорость передачи пользовательских данных?
5. Назовите основные элементы архитектуры системы GSM и опишите их функции. Какие преимущества предоставляет спецификация всех внутренних интерфейсов системы GSM (в дополнение к радиоинтерфейсу)?
6. Опишите функции станции MS и модуля SIM. Почему в системе GSM они разделены? Каким образом и где в системе GSM сохраняется информация о пользователе? Как пользовательские данные защищены от несанкционированного доступа, в частности, при передаче через радиоинтерфейс? Каким образом можно определить местонахождение мобильной станции (не только положение текущей станции BTS)? Подумайте об измерениях качества сигнала.
7. Как архитектура регистров HLR/VLR, используемых в системе GSM, ограничивает количество пользователей, в особенности движущихся пользователей?
8. Какая инфраструктура необходима для системы GPRS и не нужна в системе HSCSD? Какие ее составляющие являются новыми? Назовите их функции.
9. Какие ограничения на диаметр ячеек и скорость передачи (речи, данных) существуют в системах GSM, HSCSD, GPRS? Каким образом можно увеличить скорость передачи?
10. Какие схемы уплотнения используются в системе GSM? С какими целями? Кроме физического уровня, рассмотрите также и более высокие уровни.
11. Как в системе GSM достигается синхронизация? Почему она является столь важной? Какие протоколы отвечают за проведение синхронизации?
12. Назовите причины задержек обмена пакетными данными в системе GSM. Помните, что нужно различать пакетно-ориентированную передачу и передачу с коммутацией каналов.
13. Где и в каких случаях в системе GSM могут происходить конфликты? Сравните конфликты, возникающие при передаче данных в обычной системе GSM, системах HSCSD и GPRS.
14. Для чего и где необходимы разные каналы передачи сигналов? Какие отличия существуют между ними?
15. Каким образом в системе GSM происходит локализация, обновление местонахождения и роуминг? Как это отражается на базах данных? В каких ситуациях необходим роуминг?
16. Почему в системе GSM определено так много идентификаторов и адресов (MSISDN, TMSI, IMSI и т.д.)? Назовите причины, принимая во внимание различия между идентификаторами пользователей и системными идентификаторами.

17. Зачем в системе GSM необходимо переключение? Какие проблемы возникают при переключении? Перечислите различные этапы переключения и его типы. Какие ресурсы необходимы для переключения при передаче данных в системах HSCSD и GPRS? Можно ли при этом гарантировать качество обслуживания?
18. Какие функции выполняют аутентификация и шифрование в системе GSM? Как поддерживается безопасность в системе?
19. Как можно увеличить скорость передачи данных в стандартной системе GSM? Как в этих целях используются дополнительные схемы HSCSD, GPRS и EDGE? Какие различия существуют между перечисленными подходами? Одинакова ли их сложность? Какие проблемы остаются и после увеличения скорости передачи?
20. Чем ограничена скорость передачи в системах GPRS и HSCSD при использовании реальных устройств (в сравнении с теоретическими возможностями системы GSM)?
21. Рассмотрим самый высокий класс задержки. Пусть данные передаются со скоростью 115,2 Кбит/с. Сколько битов будет послано до того, как передатчик получит первое подтверждение приемника (допустим, что задержками в стационарной сети и системе приемника можно пренебречь)? Вспомним, что средний размер передачи для данных WWW составляет приблизительно 10 Кбайт. Что происходило бы над уровнем GPRS при использовании стандартного протокола TCP (эти вопросы обсуждаются в главах 10 и 11)? Теперь представим себе медленное начало сеанса связи и его связь с временем передачи и подтверждения приема. Какие изменения нужно сделать для реализации этой схемы?
22. Какая часть исходной сети GSM используется в системе GPRS? С помощью каких элементов осуществляется передача данных?
23. Каковы типичные значения скорости передачи в системе DECT? Как они достигаются при использовании кадров TDMA? Какие из схем уплотнения применяются в системе DECT и с какими целями? Сравните сложность систем DECT и GSM.
24. Приведите примеры типичных пользователей магистральных радиосистем. Что делает радиомагистрали особенно привлекательными для этих групп пользователей? Чем в этом отношении они отличаются от других существующих сетей? Почему магистральные радиосистемы являются более дешевыми, чем, например, системы GSM?
25. Перечислите основные свойства систем мобильной телефонной связи 3-го поколения. За счет чего в этих системах увеличивается емкость и скорость передачи? Каким образом в системе UMTS осуществляются асимметричные соединения и обеспечиваются различные значения пропускной способности?
26. Сравните текущее состояние мобильных телефонных сетей в Европе, Японии и Северной Америке. В чем отличия? Какие усилия предпринимаются для создания общей системы или, по крайней мере, совместимых систем?

4.7. Рекомендуемая литература

Adachi, F.; Sawahashi, M.; Suda, H. (1998) "Wideband DS-CDMA for next-generation mobile communication systems", *IEEE Communications Magazine*, 36, (9).

Brache, G.; Walke, B. (1997) "Concepts, services, and protocols of the new GSM phase 2+ General Packet Radio Service", *IEEE Communications Magazine*, 35, (8).

Callendar, M. (1997) "International Mobile Telecommunications-2000 standards efforts of the ITU", сборник статей в *IEEE Personal Communications*, 4, (4).

Dahlman, E.; Gudmundson, B.; Nilsson, M.; Skold, J. (1998) "UMTS/IMT-2000 based on wideband CDMA", *IEEE Communications Magazine*, 36, (9).

Dasilva, J; Ikonomou, D.; Erben, H. (1997) "European R&D programs on third-generation mobil communication systems", *IEEE Personal Communications*, 4, (1).
DECT Forum (1999), `http:/www.dect.ch/`.
ETSI (1991a) *Bearer services supported by a GSM PLMN*, European Telecommunications Standards Institute, GSM recommendations 02.02.
ETSI (1991б) *General description of a GSM PLMN*, European Telecommunications Standards Institute, GSM recommendations 01.02.
ETSI (1991в) *Subscriber Identity Modules, Functional Characteristics*, European Telecommunications Standards Institute, GSM recommendations 02.17.
ETSI (1993a) *Multiplexing and multiple access on the radio path*, European Telecommunications Standards Institute, GSM recommendations 05.02.
ETSI (1993б) *MS-BSS data link layer — general aspects*, European Telecommunications Standards Institute, GSM recommendations 04.05.
ETSI (1993в) *MS-BSS data link layer specification*, European Telecommunications Standards Institute, GSM recommendations 04.06.
ETSI (1997a) *High Speed Circuit Switched Data (HSCSD)*, Stage 1, European Telecommunications Standards Institute, GSM 02.34, V5.2.1.
ETSI (1998a) *General Packet Radio Service (GPRS); Requirements specification of GRPS*, European Telecommunications Standards Institute, TR 101 186, V6.0.0 (1998–04).
ETSI (1998б) *General Packet Radio Service (GPRS); Service Description; Stage 2*, European Telecommunications Standards Institute, EN 301 344, V6.1.1 (1998–08).
ETSI (1998в) *General Packet Radio Service (GPRS); Service Description; Stage 1*, European Telecommunications Standards Institute, EN 301 113, V6.1.1 (1998–11).
ETSI (1998г) *General Packet Radio Service (GPRS); Overall description of the GPRS radio interface; Stage 2*, European Telecommunications Standards Institute, TS 101 350, V6.0.1 (1998–08).
ETSI (1998д) *High Speed Circuit Switched Data (HSCSD); Stage 2*, European Telecommunications Standards Institute, TS 101 038, V5.1.0 (1998–07).
ETSI (1998e) *Universal Mobile Telecommunications System (UMTS); Concept groups for the definition of the UMTS Terrestrial Radio Access (UTRA)*, European Telecommunications Standards Institute, TR 101 397, V3.0.1 (1998–10).
ETSI (1998ж) *Universal Mobile Telecommunications System (UMTS); High level requirements relevant for the definition of the UMTS Terrestrial Radio Access (UTRA) concept*, European Telecommunications Standards Institute, TR 101 398, V3.0.1 (1998–10).
ETSI (1998з) *High level requirements relevant for the definition of the UMTS Terrestrial Radio Access (UTRA) concept*, European Telecommunications Standards Institute, TR 101 398, V3.0.1 (1998–10).
ETSI (1998и) *Concept groups for the definition of the UMTS Terrestrial Radio Access (UTRA)*, European Telecommunications Standards Institute, TR 101 397, V3.0.1 (1998–10).
ETSI (1998к) *Digital Enhanced Cordless Telecommunications (DECT), Generic Access Profile (GAP)*, European Telecommunications Standards Institute, EN 300 444, V1.3.2 (1998–03).
ETSI (1998л) *Digital Enhanced Cordless Telecommunications (DECT), Common Interface (CI)*, European Telecommunications Standards Institute, EN 300 175, V1.4.1 (1998–02).
ETSI (1998м) *Terrestrial Trunked Radio (TETRA), Voice plus Data (V+D)*, European Telecommunications Standards Institute, серия стандартов ETS 300 392.
ETSI (1998н) *Terrestrial Trunked Radio (TETRA), Packet Data Optimized (PDO)*, European Telecommunications Standards Institute, серия стандартов ETS 300 393.
ETSI (1998о) *The ETSI UMTS Terrestrial Radio Access (UTRA) ITU-R Radio Transmission Technologies (RTT) Candidate Submission*, European Telecommunications Standards Institute.
European Telecommunications Standards Institute (1999), `http://www.etsi.org/`.
Goodman, D. (1997) *Wireless Personal Communications Systems*. Addison Wesley Longman.

GSM Data, Intel Corporation (1998), http://www.gsmdata.com/.

GSM MoU (1998) *Vision for the evolution from GSM to UMTS*, GSM MoU Association, Permanent Reference Document, V 3.0.0.

GSM Memorandum of Understanding Association (1999) http://www.gsmworld.com/.

Halsall, F. (1996) *Data communications, computer networks and open systems*. Addison Wesley Longman.

International Telecommunication Union, International Mobile Telecommunications (1999), http://www.itu.int/imt/.

Nilsson, T. (1998) "Toward third-generation wireless communication", *Ericsson Review No. 2*, http://www.ericsson.com/.

Ojanpera, T.; Prasad, R. (1998) "An overview of third-generation wireless personal communications: A European perspective", *IEEE Personal Communications*, 5, (6).

Shafi, M.; Sasaki, A.; Jeong, D. (1998) "IMT-2000 developments in the Asia Pacific region", сборник статей, *IEEE Communications Magazine*, 36, (9).

TETRA Memorandum of Understanding (1999), http://www.tetramou.com/.

Tripathi, N.D.; Reed, J.H.; VanLandingham, H.F. (1998) "Handoffs in cellular systems", *IEEE Personal Communications*, 5, (6).

UMTS Forum (1999) http://www.umts-forum.org/.

Wong, D.; Lim, T. (1997) "Soft handoffs in CDMA mobile systems", *IEEE Personal Communications*, 4, (6).

Спутниковые системы 5

В этой главе представлена еще одна система, поддерживающая мобильную связь, — спутниковая связь. Спутники обеспечивают всемирное покрытие без затрат на прокладку кабеля между базовыми станциями и почти не зависят от изменяющейся плотности населения. В данной главе, после краткого изложения истории спутников и обсуждения различных областей их применения, рассматриваются основы спутниковых систем. Орбита, видимость, качество передачи и другие характеристики этих систем тесно связаны между собой. Далее объясняется, как различные ограничения и требования к использованию привели к возникновению трех основных классов спутников — GEO, MEO и LEO. Высокая скорость маловысотных спутников поднимает новые проблемы маршрутизации, определения местонахождения мобильных пользователей и переключения каналов связи. Некоторые аспекты этих вопросов представлены в отдельных разделах. Глава завершается рассмотрением четырех примеров планируемых и уже установленных систем глобальной спутниковой связи.

5.1. История

Спутниковую связь начали использовать после второй мировой войны. Ученые знали о возможности создания ракет, которые бы выносили радиопередатчики в космос. В 1945 году Артур К. Кларк (Arthur C. Clarke) опубликовал очерк "Внеземные ретрансляторы". Но только в 1957 году, после начала холодной войны, в Советском Союзе состоялся запуск первого спутника, который так и назывался — СПУТНИК. Это событие потрясло западный мир. СПУТНИК нельзя сравнить с современными аппаратами. Это был всего лишь небольшой передатчик, периодически посылавший сигнал "бип". Однако, этого было достаточно, чтобы США приложили максимум усилий для разработки своего первого спутника. Спустя три года, в 1960 году, состоялся запуск первого американского отражающего спутника связи. Он назывался ECHO и представлял собой зеркало. Связь осуществлялась с помощью отражения сигналов. Еще через три года был запущен первый геостационарный (или геосинхронный) спутник SYNCOM. Геостационарные спутники даже сейчас являются магистралями для трансляции новостей. Их главное преимущество — фиксированное расположение над землей (более подробно это объясняется в разделе 5.3.1). Такие спутники вращаются синхронно с Землей, поэтому кажутся зависшими над определенным местом.

Наконец, в 1965 году вступил в строй первый коммерческий геостационарный спутник связи INTELSAT 1, также известный как "Early Bird" ("ранняя пташка"). Он использовался 1,5 года, весил 68 кг и предоставлял 240 дуплексных телефонных каналов или, как альтернативу, один телевизионный канал. Спутники INTELSAT 2 и INTELSAT 3, запущенные в 1967 и 1969 годах соответственно, предоставляли уже 1 200 телефонных каналов. Для наземной связи всегда можно использовать провода. Однако, для морских судов это не выход. Поэтому в 1976 году вступили в строй три спутника MARISAT. Их задачей было обеспечение всемирной мор-

ской связи. Для передатчиков и приемников, установленных на судах, все еще были характерны большие антенны (размеры порядка 1,2 м, мощность передачи 40 Вт). Первая мобильная спутниковая система INMARSAT-A начала свою работу в 1982 году. Через шесть лет появилась система INMARSAT-C — первая спутниковая система, обеспечивающая мобильную телефонную связь и передачу данных. (Скорость передачи данных достигала 600 бит/с, существовали устройства сопряжения с сетями с коммутацией пакетов стандарта X.25.) В 1993 году спутниковые системы, наконец, стали полностью цифровыми. Однако, в действительности мобильность пользователей была относительной, поскольку устройства, необходимые для связи через стационарные спутники, были тяжелыми (несколько килограмм), а для достижения приемлемой скорости передачи была нужна высокая мощность. 1998 год ознаменовал начало новой эры спутникового обмена данными. В этом году начали действовать глобальные спутниковые системы с небольшими мобильными телефонами, в частности, системы Iridium и Globalstar (см. раздел 5.7). В настоящее время в коммерческом использовании находятся почти 200 геостационарных спутников, что иллюстрирует развитие спутниковой связи за последние 30 лет [Miller, 1998], [Maral, 1998], [Pascall, 1997].

5.2. Сферы применения

Традиционно спутники применяются в следующих областях.

- **Прогноз погоды**. Несколько спутников поставляют снимки Земли, сделанные в инфракрасном и видимом диапазоне. Без спутников предсказание ураганов стало бы невозможным.
- **Радиовещательные и телевизионные спутники**. При использовании спутника доступны сотни радио- и телепрограмм. Эта технология во многих местах вытесняет кабельное радиовещание и телевидение, так как ее установка дешевле. Кроме того, в большинстве случаев подобные услуги не требуют дополнительных расходов. В Европе современные спутниковые тарелки имеют диаметр 30–40 см, в северных странах диаметры немного больше.
- **Военные спутники**. Использование в разведывательных целях было одним из самых первых применений спутников. Кроме того, управление многими каналами связи происходит также через спутники, поскольку их сложнее атаковать.
- **Навигационные спутники**. Хотя вначале глобальная система навигации и определения положения GPS применялась только в военных целях, сейчас она широко известна и доступна каждому. Система позволяет точно определить местонахождение по всему миру, а при использовании некоторых дополнительных технологий точность достигает нескольких метров. Почти все корабли и самолеты полагаются на систему GPS (в дополнение к традиционным навигационным системам). Более того, появляется все больше грузовых и легковых автомобилей со встроенными приемниками системы GPS. Эта система также используется для управления автоперевозками и для определения местонахождения автомобиля при угоне.

В контексте мобильной связи особый интерес представляет применение спутников для передачи данных.

- **Глобальные телефонные магистрали**. Как уже упоминалось, одним из первых применений спутников для связи стало установление международных телефонных магистралей. Иногда запустить новый спутник можно быстрее, чем проложить кабель. Однако, спутники, действующие как магистрали телекоммуникационных сетей, хотя еще и находятся в использовании, но все чаще заменяются оптоволоконными кабелями, проложенными через океан. Главными причинами этого являются огромная пропускная способность оптоволоконных соединений (несколько десятков Гбит/с для уплотнения с разделением по длине волны) и

намного меньшие задержки, в сравнении со спутниками. Чтобы попасть от передатчика через геостационарный спутник к приемнику, сигнал должен пройти около 72 000 км. В оптоволоконных каналах, пересекающих Тихий или Атлантический океан, это расстояние обычно не превышает 10 000 км. К сожалению, скорость света конечна. При использовании геостационарных спутников это приводит к задержке распространения в 0,25 с. В телефонных разговорах подобные задержки требуют терпения, а иногда могут просто раздражать.

- **Соединения с удаленными местами или развивающимися областями**. Во многих местах невозможно использовать прямое проводное соединение с телефонной сетью или Internet из-за географического положения (например, при исследовании Антарктики). Другой причиной может быть неудовлетворительное состояние соответствующей инфраструктуры в какой-то стране. Спутники обеспечивают простое и быстрое подсоединение к глобальным сетям [Schwarz, 1996].
- **Глобальная мобильная связь**. Последней тенденцией в применении спутников является поддержка глобального мобильного обмена данными. Геостационарные спутники не подходят для этих целей из-за задержек, поэтому необходимы спутники, движущиеся по более низким орбитам (см. раздел 5.3). Основная задача спутников в мобильной связи — не замена существующих мобильных телефонных сетей, а их расширение на те области, куда покрытие не распространяется. Не охваченными обычно являются места с низкой плотностью населения, где установление базовой станции неоправданно дорого. При интеграции спутниковой связи мобильные телефоны могут коммутировать со спутниками, обеспечивая абоненту возможность соединения по всему миру [Jamalipour, 1998].

Если вначале спутники были простыми транспондерами, то сейчас они скорее похожи на летающие маршрутизаторы. Транспондеры принимают сигнал на некоторой частоте, усиливают его, а затем передают на другой частоте. Когда-то было возможным только аналоговое усиление. Однако, использование цифровых сигналов тоже позволяет восстанавливать данные. Спутник декодирует сигнал в поток битов, который затем опять кодируется в сигнал. Цифровое восстановление, в сравнении с чисто аналоговым усилением, обеспечивает более высокое качество сигнала, принимаемого на земле. В следующих разделах будет показано, что современные спутники связи осуществляют многие функции высших уровней связи, в частности, межспутниковую маршрутизацию, коррекцию ошибок и т.п.

Рис. 5.1 иллюстрирует ситуацию, типичную для спутниковых систем, поддерживающих глобальную мобильную связь [Lutz, 1998]. Каждый из спутников, в зависимости от типа, покрывает своим лучом некоторую площадь на земле — так называемую зону обслуживания (см. раздел 5.3). Внутри зоны обслуживания связь спутника и мобильных пользователей происходит через **каналы мобильных пользователей** (MUL). Базовая станция, действующая как шлюз между спутником и другими сетями, управляет спутником через **шлюзовый канал** (GWL). Кроме того, спутники иногда могут связываться непосредственно между собой через **межспутниковые каналы** (ISL). Это позволяет прямую связь между пользователями из различных зон обслуживания без использования базовых станций или других наземных сетей. Сохранение дополнительных спутниковых каналов позволяет уменьшить задержку речи и пакетов данных. В некоторых спутниках применяются специальные антенны для создания меньших ячеек с помощью сфокусированных лучей (например, в системе ICO на каждый спутник приходится 163 сфокусированных луча [ICO, 1998]).

Спутниковые системы являются и будут являться ценным дополнением ко многим уже существующим наземным сетям. Пользователи могут связываться между собой через сеть ISDN, другие сети PSTN или даже через сотовые сети, в частности, через систему GSM. Многие шлюзы обеспечивают прямое соединение между этими различными сетями. Настоящая проблема в описанном примере — плавное прямое переключение между сотовой сетью и спутниковой системой.

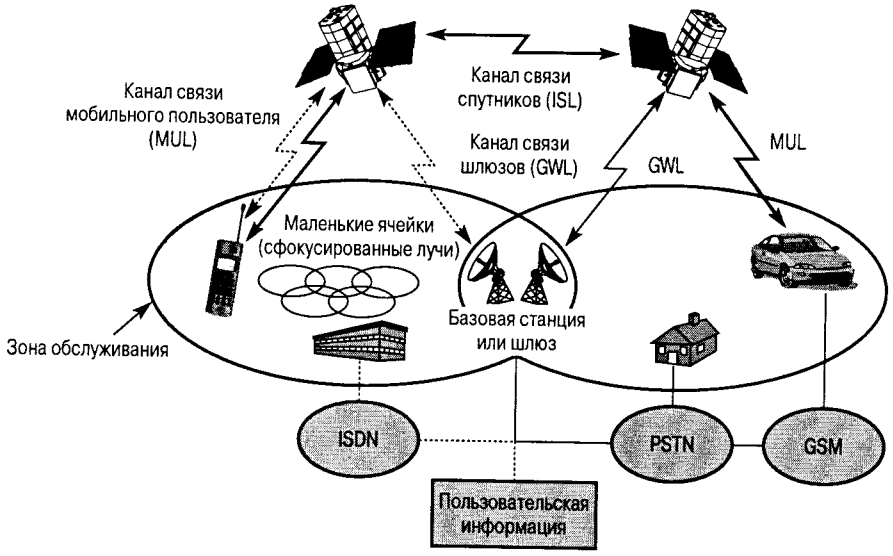

Рис. 5.1. Типичная спутниковая система глобальной мобильной связи

5.3. Основы

Спутники двигаются вокруг Земли по орбитам. В зависимости от приложения, эти орбиты могут быть круговыми или эллиптическими. Спутники на круговых орбитах всегда находятся на одинаковом расстоянии от поверхности земли и подчиняются простым законам.

- Сила гравитационного притяжения F_g Земли равна $mg(R/r)^2$.
- Центробежная сила F_c "выталкивает" спутник с орбиты и равна $mr\omega^2$.

Здесь используются следующие обозначения:

- m — масса спутника;
- R — радиус Земли, равный 6 370 км;
- r — расстояние от спутника к центру Земли;
- g — ускорение свободного падения, $g = 9{,}81$ м/с^2;
- ω — угловая скорость ($\omega = 2\pi f$), f — частота обращения.

Для устойчивой круговой орбиты должно выполняться следующее уравнение:

- $F_g = F_c$, т.е. силы должны быть одинаковы по величине. При рассмотрении этого соотношения можно сразу заметить, что масса спутника из него выпадает (сокращается).
- Решая уравнение относительно расстояния r от спутника к центру Земли, получим следующий результат:

$$\text{Расстояние } r = (gR^2/(2\pi f)^2)^{1/3}$$

Из последнего уравнения следует, что расстояние от спутника до поверхности Земли зависит от частоты его обращения. Эта зависимость, а также соответствующая относительная скорость спутника, приводятся на рис. 5.2. Одна из самых интересных точек диаграммы — точка, в которой период спутника равен 24 часам. В этом случае расстояние r равно 35 786 км. Если спутник характеризуется орбитальным временем 24 часа и находится над экватором, он называется геостационарным. (Спутники этого типа рассматриваются в следующем разделе.)

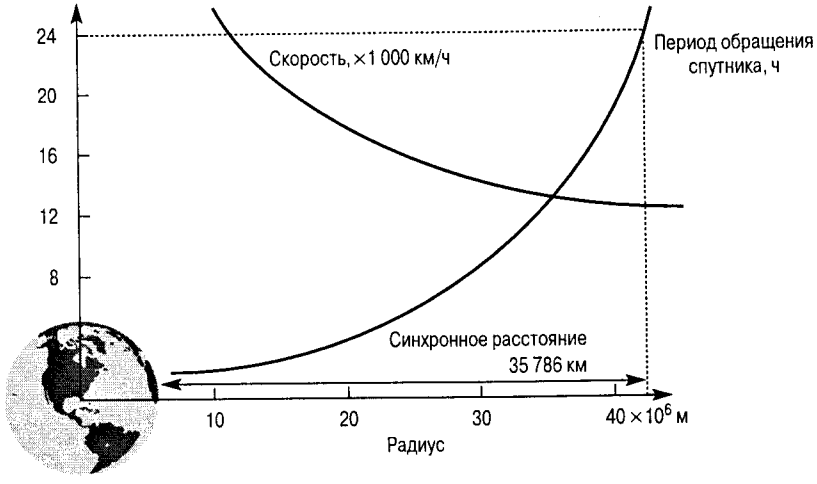

Рис. 5.2. Зависимость периода обращения спутника от расстояния до Земли

Углы склонения и возвышения являются важными параметрами в спутниковой связи. **Угол склонения** δ (см. рис. 5.3) определяется как угол между экваториальной плоскостью и плоскостью орбиты спутника. При угле склонения 0° спутник находится точно над экватором. Если орбита спутника не круговая, то ее ближайшая к Земле точка называется перигеем.

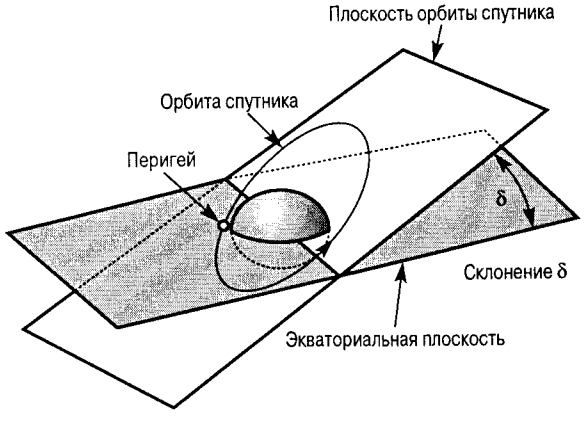

Рис. 5.3. Угол склонения спутника

Угол возвышения ε (см. рис. 5.4) определен как угол между центральным лучом спутника и касательной плоскостью к поверхности Земли. Далее, так называемая **зона обслуживания** — это область, в которой могут приниматься сигналы спутника.

Еще одной особенностью спутниковой связи являются потери на пути распространения. Ослабление мощности сигнала зависит от расстояния между наземным приемником и спутником, угла возвышения спутника и атмосферных условий. Потери L, зависящие от расстояния r между передатчиком и приемником, можно вычислить по формуле:

$$L = (4\pi rf/c)^2,$$

где f — несущая частота, а c — скорость света.

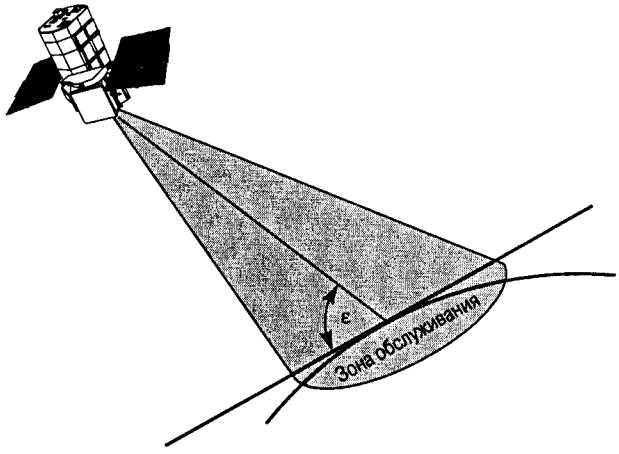

Рис. 5.4. Угол возвышения спутника

Это означает, что мощность принимаемого сигнала убывает пропорционально квадрату расстояния. При некоторых допущениях (относительно мощности передачи, диаметров антенн, рабочей частоты и т.п.) это также непосредственно влияет на максимально достижимую скорость передачи данных [Comparetto, 1997]. При использовании спутников, находящихся на расстоянии нескольких сотен километров, можно достичь скорости 10 Кбит/с (раздел 5.3.2) с антеннами обычных мобильных телефонов и несущей частотой 2 ГГц. В то же время скорость передачи для геостационарных спутников, находящихся на расстоянии 36 000 км, не превышает нескольких десятков бит/с.

Ослабление сигнала вследствие некоторых атмосферных условий происходит более сложным образом (см. рис. 5.5). В зависимости от угла возвышения сигналу нужно проникнуть через более или менее толстый слой атмосферы. Как правило, углы возвышения, меньшие 10 градусов, считаются бесполезными для связи. Поглощение радиоволн вследствие дождя является особенно сильным в зоне тропиков (количество ошибок передачи во время послеполуденных ливней резко возрастает).

Можно определить четыре различных типа орбит, показанных на рис. 5.6.

- **Геостационарные (или геосинхронные) околоземные орбиты (medium earth orbit — GEO).** Спутники GEO находятся на расстоянии около 36 000 км от поверхности Земли. Примерами таких спутников являются почти все теле- и радиовещательные спутники, многие метеорологические спутники, а также спутники, действующие как магистрали телефонных сетей (см. раздел 5.3.1).
- **Средние околоземные орбиты (medium earth orbit — MEO).** Спутники MEO движутся на расстояниях 5 000–12 000 км от поверхности Земли. До сегодняшнего времени этот класс содержал небольшое количество спутников, однако, некоторые системы будущего (в частности, система ICO) по различным причинам будут использовать именно его (см. раздел 5.3.3).
- **Низкие околоземные орбиты (low earth orbit — LEO).** В прошлом спутники LEO применялись главным образом в разведывательных целях. Сейчас этот класс, использующий высоты в диапазоне 500–1 500 км, положен в основу нескольких новых спутниковых систем (см. раздел 5.3.2).
- **Сильно эллиптические орбиты (highly elliptical orbit — HEO).** Этот класс включает все спутники с некруговыми орбитами. В настоящее время разрабатывается несколько коммерческих систем связи, использующих спутники с эллиптическими орбитами. Такие системы имеют перигей над большими городами, что позволяет улучшить качество связи.

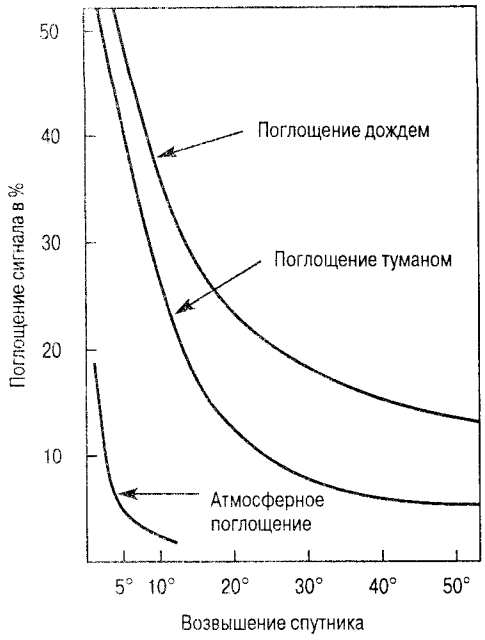

Рис. 5.5. Ослабление сигнала вследствие поглощения в атмосфере

Радиационные пояса Ван Аллена (Van Allen), состоящие из ионизированных частиц, находятся на высотах 2 000–6 000 км (внутренний пояс Ван Аллена) и 15 000–30 000 км (внешний пояс Ван Аллена). Соответственно, использовать эти орбиты для спутниковой связи крайне трудно.

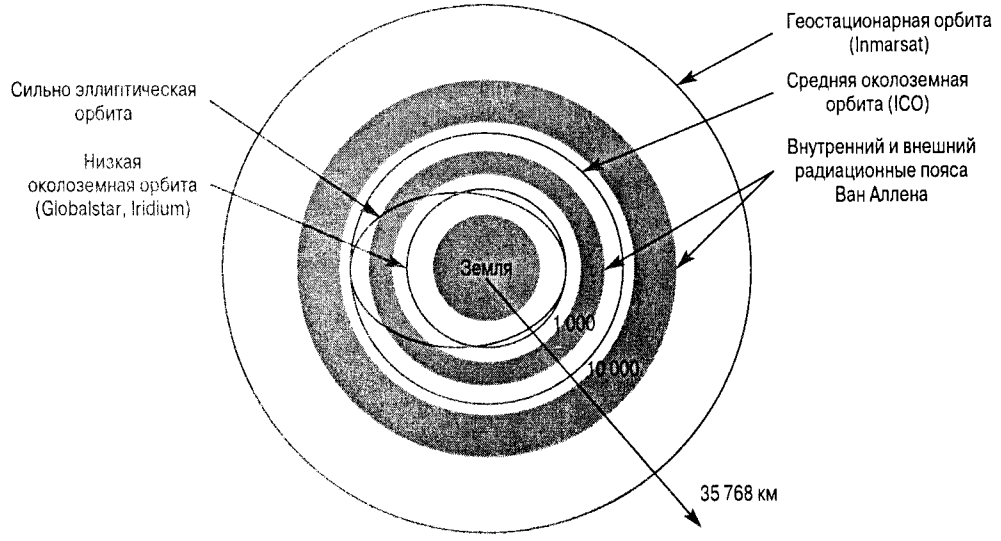

Рис. 5.6. Различные типы спутниковых орбит

5.3.1. Спутники GEO

Чтобы положение спутника в небе казалось фиксированным, он должен иметь период обращения 24 часа. Использовав уравнение для расстояния между Землей и спутником $r = (gR^2/(2\pi f)^2)^{1/3}$, можно показать, что соответствующее расстояние равно 35 786 км. Кроме того, склонение орбиты должно быть нулевым.

- **Преимущества**. Трех спутников GEO достаточно для покрытия любой точки земного шара. Положение антенн передатчиков и приемников можно фиксировать, регулирование не нужно. Поэтому спутники GEO идеально подходят для теле- и радиовещания. Срок их использования достаточно большой — около 15 лет. Для спутников GEO, из-за большой зоны обслуживания, переключение обычно не требуется. Кроме того, из-за отсутствия относительного движения они не подвергаются воздействию доплеровского сдвига.
- **Недостатки**. В северных и южных областях земного шара прием сигналов спутников GEO затруднен. Это обусловлено малыми углами возвышения для широт выше 60°. Таким образом, в этих регионах нужно использовать антенны большего радиуса. Для передачи необходима сравнительно высокая мощность (несколько десятков Вт), а при использовании устройств, работающих на батареях, этому требованию удовлетворить сложно. Таким образом, спутники GEO не могут использоваться небольшими мобильными телефонами. Самой крупной проблемой при передаче речи и данных является большая задержка (более 0,25 с). Поэтому многие схемы повторной передачи, известные из стационарных сетей, здесь не срабатывают. Кроме того, из-за большой зоны обслуживания нельзя повторно использовать частоты. В ином случае спутникам GEO необходимы специальные антенны, фокусированные на меньшие зоны обслуживания. Выведение спутника GEO на орбиту является очень дорогостоящей процедурой.

5.3.2. Спутники LEO

Спутники LEO двигаются по более низким орбитам, поэтому очевидно, что они характеризуются меньшими периодами (периоды спутников LEO обычно составляют от 95 до 120 минут). Кроме того, для более высокого качества каналов связи в системах LEO обеспечивается большой угол возвышения в любой точке земли. Каждый из спутников LEO виден с поверхности земли лишь в течении 10 минут. Можно произвести дальнейшее разделение спутников LEO на небольшие спутники со службами малой пропускной способности (несколько сотен бит/с), крупные спутники (несколько тысяч бит/с) и широковещательные спутники LEO (планируемая скорость достигает нескольких Мбит/с). Более подробное обсуждение этих вопросов интересующийся читатель найдет в статье [Comparetto, 1997].

- **Преимущества**. При использовании схем сжатия для речевой связи достаточно скорости передачи данных около 2 400 бит/с. Спутники LEO обеспечивают такую пропускную способность даже для мобильных терминалов с ненаправленными антеннами и низкой (1 Вт) мощностью передачи. Кроме того, задержка пакетов, доставляемых через системы LEO, относительно низкая (приблизительно 10 мс). Таким образом, эта задержка сравнима с задержкой для дальних проводных соединений (около 5–10 мс). Меньшие зоны обслуживания спутников LEO позволяют повторно использовать частоты аналогично тому, как это делается в сотовых сетях [Gavish, 1998]. Спутники LEO обеспечивают намного большую величину угла возвышения в полярных областях, а поэтому и лучшее глобальное покрытие.
- **Недостатки**. Самой крупной проблемой системы LEO является необходимость большого количества спутников для глобального покрытия. В нескольких схемах на орбите находятся 50–200, а то и больше спутников. Короткое время видимости с большим углом возвышения требует дополни-

тельных механизмов переключения соединения с одного спутника на другой. (Различные случаи переключения объясняются в разделе 5.4.) Из-за большого числа спутников, а также их быстрого движения, сложность всей спутниковой системы возрастает. Короткие сроки годности (5–8 лет), обуславливаемые сопротивлением атмосферы и излучением внутреннего пояса Ван Аллена, являются еще одной общей проблемой спутников LEO. Поэтому в системе из 48 спутников, рассчитанной на восемь лет (система Globalstar), новый спутник будет необходим каждые два месяца. Небольшие задержки одного спутника LEO — только часть проблемы. При установлении соединения с произвольной точкой земного шара может возникнуть необходимость маршрутизации пакетов данных от одного спутника к другому (или несколько раз от базовых станций к спутникам и обратно). Благодаря большой зоне обслуживания, в спутниках GEO маршрутизация этого типа обычно не требуется, так как приемники и передатчики с большой вероятностью находятся в одной зоне обслуживания.

5.3.3. Спутники MEO

Спутники MEO можно расположить где-то посередине между спутниками GEO и LEO. Это касается их орбит, а также их преимуществ и недостатков.

- **Преимущества**. При использовании орбит с радиусом 10 000 км системе необходимо около дюжины спутников. Это больше, чем в системе GEO, но значительно меньше, чем в системе LEO. Кроме того, эти спутники двигаются медленнее по отношению к вращению Земли, что облегчает проектирование системы (периоды спутников примерно равны шести часам). Спутник MEO, в зависимости от склонения, может покрыть довольно большую площадь, поэтому в системах MEO происходит меньше переключений.
- **Недостатки**. Опять же, из-за большого расстояния до поверхности Земли задержка возрастает до 70–80 мс. Спутникам необходимы более высокая мощность передачи и специальные антенны для меньших зон обслуживания.

5.4. Маршрутизация

Спутниковая система вместе со шлюзами и наземными сетями (см. рис. 5.1), как и любая другая сеть, должна маршрутизировать передачу данных от одного пользователя к другому. Маршрутизация в стационарной (наземной) части выполняется обычным способом. В то же время для спутниковой сети есть два различных решения. Если между спутниками существуют каналы ISL, маршрут информационного обмена можно проложить через них. Если же таких каналов нет, информация передается на землю, там маршрутизируется и передается обратно на спутник.

Предположим, что два пользователя спутниковой сети обмениваются данными. Рассмотрим случай, когда в спутниковой системе поддерживаются каналы ISL. Пусть один из пользователей посылает данные на спутник. Он, в свою очередь, направляет эти данные через остальные спутники на тот из них, который отвечает за нужный приемник. Этот последний спутник посылает данные на землю. Это означает, что для одного направления требуется лишь одна исходящая и одна нисходящая линии связи. Возможность маршрутизации внутри спутниковой системы уменьшает количество шлюзов, требуемых на земле.

Пусть теперь в спутниковой системе нет каналов ISL. Пользователь опять посылает данные на спутник, однако теперь они направляются на наземный шлюз. Затем происходит обычная маршрутизация в стационарных сетях, пока сигнал не достигает шлюза, отвечающего за нужный приемник. Данные опять посылаются на спутник, который перенаправляет их на приемник. Для реализации этого решения нужно две исходящих и две нисходящих линии связи. В зависимости от орбит и скорости маршрутизации в спутниковой сети, в сравнении с наземной сетью, задержка может быть

меньшей. Недостатком каналов ISL является высокая сложность системы, обусловленная дополнительными антеннами, а также аппаратными и программными средствами маршрутизации.

5.5. Определение местонахождения

Определение местонахождения пользователя в спутниковых сетях осуществляется почти так же, как и в наземных сотовых сетях. Одна дополнительная проблема обусловлена тем, что "базовые станции", т.е. спутники, тоже двигаются. Шлюзы спутниковых сетей поддерживают несколько регистров. В **регистре исходного положения** (HLR) сохраняется статическая информация о пользователе и его текущем местонахождении. Последнее известное местонахождение мобильного пользователя сохраняется в **регистре местонахождения посетителей** (VLR). Функции регистров VLR и HLR похожи на соответствующие функции регистров системы GSM (см. главу 4). В спутниковых сетях особенно важным является **регистр соответствия пользователей и спутников** (SUMR). В нем хранится информация о текущем местонахождении спутников и соответствии между каждым пользователем и спутником, через который осуществляется связь с этим пользователем.

Регистрация мобильной станции происходит следующим образом. Сперва мобильная станция посылает сигнал, принимаемый одним или несколькими спутниками. Спутники, получившие такой сигнал, сообщают об этом событии на шлюз. Шлюз определяет местонахождение пользователя через положение спутников. Затем из регистра HLR запрашиваются данные о пользователе, а регистры VLR и SUMR обновляются.

Вызов мобильной станции происходит подобно тому, как это делается в системе GSM. Звонок направляется на шлюз, который затем определяет местонахождение мобильной станции, используя регистры HLR и VLR. С помощью регистра SUMR определяется подходящий спутник связи и устанавливается соединение.

5.6. Переключение

Важным вопросом спутниковых систем, использующих спутники MEO или (особенно) LEO, является переключение. Представим себе сотовую мобильную телефонную сеть с быстро движущимися базовыми станциями. В упомянутых спутниковых системах имеем именно такую ситуацию — для мобильных телефонов каждый спутник представляет собой базовую станцию. Движение спутника может привести к необходимости введения дополнительных (в сравнении с наземными мобильными телефонными сетями) вариантов переключения.

- **Внутриспутниковое переключение**. С помощью специальных антенн спутник может создавать в своей зоне обслуживания несколько сфокусированных лучей. Внутриспутниковое переключение необходимо при перемещении пользователя из одного сфокусированного луча в другой луч того же спутника. Движение спутника может приводить к такому же эффекту.
- **Межспутниковое переключение**. Когда пользователь покидает зону обслуживания спутника или удаляется сам спутник, происходит переключение пользователя на следующий спутник. Переключение называется жестким, если коммутация осуществляется мгновенно. При гибком переключении (как в системах CDMA) в течение некоторого времени используются оба спутника (или даже больше). Межспутниковое переключение может также возникать, если система поддерживает каналы ISL. Частые переключения являются платой за высокое качество передачи. Чем выше необходимое качество, тем большими должны быть углы возвышения. Однако, большие углы возвышения подразумевают частые переключения, а это, в свою очередь, усложняет систему.
- **Переключение шлюзов**. В то время как пользователь может все еще иметь хорошее соединение со спутником, сам спутник может покинуть зону действия текущего шлюза. В этом случае спутнику нужно подсоединиться к другому шлюзу.

- **Межсистемное переключение**. Три упомянутых выше типа переключения происходят внутри спутниковой системы. Переключение, рассматриваемое в данном пункте, затрагивает различные системы. Обычно спутниковые системы используются в удаленных местах, когда все остальные сети недоступны. Как только становятся доступными традиционные сотовые сети, пользователи переключаются на этот тип связи, поскольку она обычно более дешевая и свободна от задержек. В современных системах возможно использование мобильных телефонов с двумя (и более) режимами. Однако, добиться прямого переключения между спутниковыми и наземными системами (и наоборот), к сожалению, пока не удалось.

5.7. Примеры

В табл. 5.1 приводится четыре примера современных спутниковых сетей (см. также [Miller, 1998] и [Lutz, 1998]). Система **Iridium** уже находится в эксплуатации. Изначально она предназначалась для 77 спутников (напомним, что атом иридия имеет 77 электронов). Сейчас она состоит из 66 основных и шести резервных спутников [Iridium, 1998]. Это первая коммерческая система спутников LEO со всемирным покрытием. Спутники двигаются на высоте 780 км, каждый из них весит около 700 кг. Последняя величина превосходит вес спутников конкурирующей системы Globalstar. Это обусловлено возможностью маршрутизации между спутниками системы Iridium через каналы ISL, т.е. спутнику необходима большая память, мощность обработки и т.п. Мобильные станции ("МС" в табл. 5.1) действуют согласно схеме FDMA/TDMA с дуплексом TDD на частоте 1,6138–1,6265 ГГц. Фидерные соединения со спутниками используют частоту 29,1–29,3 ГГц для исходящей линии связи и 19,4–19,6 ГГц — для нисходящей линии. Каналы ISL работают в диапазоне 23,18–23,38 ГГц.

Таблица 5.1. Примеры систем спутников MEO и LEO

	Iridium	Globalstar	ICO	Teledesic
Количество спутников	66 + 6	48 + 4	10 + 2	288
Высота, км	780	1 414	10 390	Прибл. 700
Покрытие	всемирное	±70° широты	всемирное	всемирное
Минимальный угол возвышения	8°	20°	20°	40°
Частоты, ГГц (приблизительно)	1,6 МС 29,2 ↑ 19,5 ↓ 23,3 ISL	1,6 МС ↑ 2,5 МС ↓ 5,1 ↑ 6,9 ↓	2 МС ↑ 2,2 МС ↓ 5,2 ↑ 7 ↓	19 ↓ 28,8 ↑ 62 ISL
Метод доступа	FDMA/TDMA	CDMA	FDMA/TDMA	FDMA/TDMA
Каналы ISL	Есть	Нет	Нет	Есть
Скорость передачи	2,4 Кбит/с	9,6 Кбит/с	4,8 Кбит/с	64 Мбит/с ↓ 2/64 Мбит/с↑
Количество каналов	4 000	2 700	4 500	2 500
Срок службы, лет	5–8	7,5	12	10
Стоимость	$4,4 млрд.	$2,9 млрд.	$4,5 млрд.	$9 млрд.

Система **Globalstar** является непосредственным конкурентом системы Iridium. Она включает меньшее число спутников с более низкими характеристиками. Это позволяет уменьшить вес спутников (450 кг) и сделать всю систему более дешевой. Система Globalstar не обеспечивает всемирного покрытия и не предоставляет каналов ISL, но зато характеризуется высокой скоростью передачи. В этой системе применяется схема CDMA и используется разнесение маршрутов рас-

пространения. Кроме того, за счет приема сигналов от нескольких спутников одновременно в системе Globalstar осуществляется гибкое переключение. Исходящие линии связи (от мобильных станций к спутникам) системы Globalstar работают в диапазоне 1,61–1,6265 ГГц, нисходящие линии — на частотах 2,4835–2,5 ГГц. Фидерные соединения шлюза со спутником используют диапазон 5,091–5,250 ГГц, спутника со шлюзом — частоты 6,875–7,055 ГГц.

Система промежуточных круговых орбит (ICO) [ICO, 1998], как явствует из названия, является системой спутников MEO. В то же время остальные три системы, представленные в табл. 5.1, состоят из спутников LEO. В системе ICO для достижения всемирного покрытия достаточно десяти основных и двух резервных спутников. Каждый спутник покрывает около 30 процентов поверхности Земли, средний угол возвышения в этой системе составляет 40°. Из-за высокой сложности спутников (т.е. больших антенн и опор солнечных элементов) их вес достигает 2 600 кг. Запуск спутников системы ICO обходится довольно дорого. Это обусловлено их крупным весом и большой высотой орбиты. В то же время у них срок службы (12 лет) больше, чем у спутников систем Globalstar и Iridium (восемь лет и меньше). Поэтому спутники системы ICO необходимо менять реже, что в конечном счете делает систему более дешевой.

В очень претенциозном проекте **Teledesic** (система спутников LEO) планируется предоставление высокоскоростных спутниковых соединений с высоким качеством обслуживания [Teledesic, 1998]. В отличие от других систем эта спутниковая сеть изначально разрабатывалась не для мобильной телефонной связи, а для возможности доступа к Internet через спутник во всем мире. Абонентами этой сети преимущественно будут предприятия, школы и т.д., расположенные в удаленных местах. Для нисходящих линий связи в системе Teledesic планируется обеспечить скорость 64 Мбит/с, для исходящих линий — скорость 2 Мбит/с. Со специальных терминалов для исходящих линий связи будет возможна даже скорость 64 Мбит/с. Приемники будут представлять собой, например, установленные на крышах терминалы размером с переносной компьютер, соединенные с локальными сетями внутри зданий. Работу системы планируется начать в 2003 году. Начальный план включал 840 основных и 84 резервных спутника. Позже это количество было снижено до 288 спутников, разделенных на 12 самолетов по 24 спутника. При десятилетнем сроке службы новый спутник необходимо будет запускать, по крайней мере, еженедельно. Для высокой скорости передачи нужны более высокие частоты, поэтому система Teledesic действует в сантиметровом диапазоне. Для исходящей линии связи используются частоты 28,6–29,1 ГГц, для нисходящей линии — диапазон 18,8–19,3 ГГц. При таких высоких частотах каналы связи легко блокируются дождем или другими помехами. По этой причине нужны углы возвышения, большие 40°. В системе Teledesic проводится маршрутизация с использованием каналов ISL, и осуществляется быстрая коммутация пакетов между спутниками.

Схема CDMA используется для доступа лишь в системе Globalstar. В основу остальных систем положены различные схемы TDMA/FDMA. Стоимость различных систем, приведенная в табл. 5.1, является лишь грубой оценкой, необходимой для сравнения. Эта стоимость непосредственно отражает сложность систем. Например, спутники системы ICO устроены более сложно, чем спутники системы Iridium, поэтому стоимость систем ICO и Iridium примерно одинакова. Простые и небольшие спутники системы Globalstar делают ее более дешевой, чем Iridium.

5.8. Резюме

Развитие спутниковых систем от геостационарных спутников конца 60-х до многих современных систем на различных орбитах произошло достаточно быстро. Сейчас наблюдается тенденция к отказу от крупных спутников GEO в пользу небольших спутников MEO и LEO. Главной причиной этого являются более короткие задержки, важные для речевой связи. Различные системы со всемирным покрытием предоставляют широкий выбор услуг — от простой передачи речи и данных с небольшой скоростью до высокоскоростных соединений с обеспечением каче-

ства обслуживания. Однако, спутниковые системы не предназначены для замены наземных систем связи, а являются дополнением к ним. Пока не ясно, насколько дорогостоящими будут эксплуатация и техническое обслуживание спутников и во сколько будет обходиться абонентам передача данных через них. В этом контексте проблемами спутников LEO являются высокая сложность и относительно короткий срок службы. Таким образом, стоимость исходной системы составляет лишь часть необходимых затрат.

Возможно появление новых систем, связывающих спутники, вращающиеся на низких орбитах, с наземными антеннами. Несколько компаний планируют использовать высотные самолеты или дирижабли для транспортировки антенн. Эти антенны будут размещаться над большими городами, обеспечивая высококачественную передачу данных с меньшими (в сравнении со спутниковыми системами) затратами.

5.9. Упражнения

1. Назовите основные применения спутниковой связи и главные тенденции ее развития.
2. Почему телекоммуникационные системы GEO сейчас заменяются оптоволоконными соединениями?
3. Как углы возвышения и склонения определяют использование спутника?
4. Какими характеристиками обладают различные орбиты? Какие их преимущества и недостатки?
5. Какие общие проблемы возникают при распространении сигнала от спутника к приемнику?
6. Какие функции может выполнять спутник, если рассматривать его как межсетевое устройство?
7. С какими проблемами столкнутся абоненты спутниковой системы мобильных телефонов в больших городах? Рассмотрите использование внутри зданий и небоскребов.
8. Почему в космосе не слишком много места для спутников GEO?

5.10. Рекомендуемая литература

Comparetto, G.; Ramirez, R. (1997) "Trends in mobile satellite technology", *IEEE Computer*, 30, (2), February.

Gavish, B.; Kalvenes, J. (1998) "The impact of satellite altitude on the performance of LEOS based communication systems", *Wireless Networks*, J. C. Baltzer, 4, (2).

Globalstar L. P. (1998) San Jose, CA, USA, `http://www.globalstar.com/`.

ICO Global Communications (1998), London, UK, `http://www.ico.com/`.

Iridium LLC, (1998) Washington, DC, USA, `http://www.iridium.com/`.

Jamalipour, A. (1998) *Low earth orbital satellites for personal communication networks*, Artech House.

Lutz, E.: "Issues in satellite personal communication systems", *Wireless Networks*, J. C. Baltzer, 4, (2).

Maral, G.; Bousquet, M. (1998) *Satellite communications systems: Systems, techniques and technology*, 3rd edition, John Wiley & Sons.

Miller, B. (1998) "Satellites free the mobile phone", *IEEE Spectrum*, March.

Pascall, S. C.; Withers, D. J. (1997) *Commercial satellite communication*, Focal Press, 1997.

Schwartz, R. (1996) *Wireless communications in developing counties: cellular and satellite systems*, Artech House.

Teledesic Corporation (1998) Kirkland, WA, USA, `http:/www.teledesic.com/`.

Широковещательные системы 6

В этой книге основное внимание уделено различным технологиям связи, обеспечивающим индивидуальные двусторонние соединения. В то же время очень важно понять, какую роль в сценариях будущих мобильных коммуникаций будут играть однонаправленные широковещательные системы. Типичные широковещательные системы, в частности, радио и телевидение, распределяют информацию независимо от нужд отдельных пользователей. Стоит отметить, что широковещание, как дополнение к двусторонним технологиям связи, может быть крайне рентабельным.

Телевидение и радиопередачи будущего будут полностью цифровыми. Уже сейчас несколько радиостанций создают и транслируют свои программы цифровым способом через Internet или цифровое радио (см. дальнейшие разделы этой главы). Проводится разработка цифрового телевидения. Цифровая передача обеспечивает распределение не только аудио- и видеоинформации, но и произвольных цифровых данных, т.е. радио- и телепрограммы могут сопровождаться мультимедийной информацией. Связанные с этим затраты значительно ниже, чем при индивидуальных беспроводных соединениях.

В следующих разделах дается общее введение в асимметричную связь и ее крайний случай — однонаправленное широковещание. Здесь одним из важных вопросов является циклическое повторение данных (рассматриваемое в разделах, посвященных широковещательным дискам). Подробно будет рассмотрена система цифрового аудиовещания (DAB), которая уже стандартизирована и используется. Интересной особенностью технологии DAB в отношении обмена данными является возможность передачи мультимедийной информации. Наконец, будет представлена система цифрового видеовещания (DVB), которая, возможно, заменит сегодняшнее телевидение. Будучи объединенной со спутниковой передачей и Internet, эта система предоставляет индивидуальным пользователям высокую скорость передачи при низких затратах [ETSI, 1999].

6.1. Обзор

Однонаправленные системы распределения или широковещательные системы являются крайними вариантами асимметричных систем связи. Очень часто из-за ограничений пропускной способности, различий в мощности передачи или из-за больших затрат система связи не может быть симметричной. В **симметричных системах связи** обеспечиваются одинаковые характеристики передачи для всех направлений связи. Иначе говоря, характеристики каналов (пропускная способность, задержки, затраты), проходящих от A до B и от B до A, одинаковы.

В качестве примера служб с симметричной связью можно взять обычную телефонную сеть или систему GSM со сквозными соединениями. В этих случаях не имеет значения, кто кому звонит — пропускная способность и задержки будут одинаковыми.

Для телефонной связи такая симметричность необходима, однако многие другие приложения не требуют одинаковых характеристик в обеих направлениях передачи информации. Рассмотрим типичную конфигурацию "клиент/сервер". В большинстве случаев клиенту нужно намного больше данных сервера, чем серверу — данных клиента. Самым известным примером этого механизма является система World Wide Web. Миллионы пользователей получают данные с Web-серверов через свои броузеры (клиенты). Пользователи посылают информацию на сервер лишь время от времени. Один сеанс запроса новых страниц — это передача порядка нескольких сотен байт, а ответ в среднем составляет несколько десятков килобайт.

Еще более яркими примерами являются телевизоры с приставками. Скорость потока видеоинформации при высоком разрешении может достигать нескольких Мбит/с. В то же время обычный пользователь лишь время от времени передает несколько байт, переключая каналы или посылая информацию, необходимую для покупки товаров в телемагазине.

Наконец, современные пейджеры и радиоприемники работают лишь в одном направлении. Эти устройства лишь принимают информацию. Чтобы послать какую-то информацию обратно, например, на радиостанцию, пользователю нужны дополнительные технологии связи (обычно в этих целях используется телефонная система).

Однонаправленные широковещательные системы являются частным случаем **асимметричных систем связи**. Здесь обычно имеют дело с высокоскоростными потоками данных от одного передатчика к многим приемникам. При этом возникает следующая проблема. Передатчик может оптимизировать посылаемые данные лишь для всей группы приемников, а не для каждого пользователя отдельно. На рис. 6.1 приводится простая схема широковещания. В передатчике производится попытка оптимизировать поток передаваемых пакетов для всех приемников без учета нужд каждого из них по отдельности. Затем пакеты с помощью широковещания передаются на приемники. Каждый приемник выбирает нужные пакеты и отсеивает остальные.

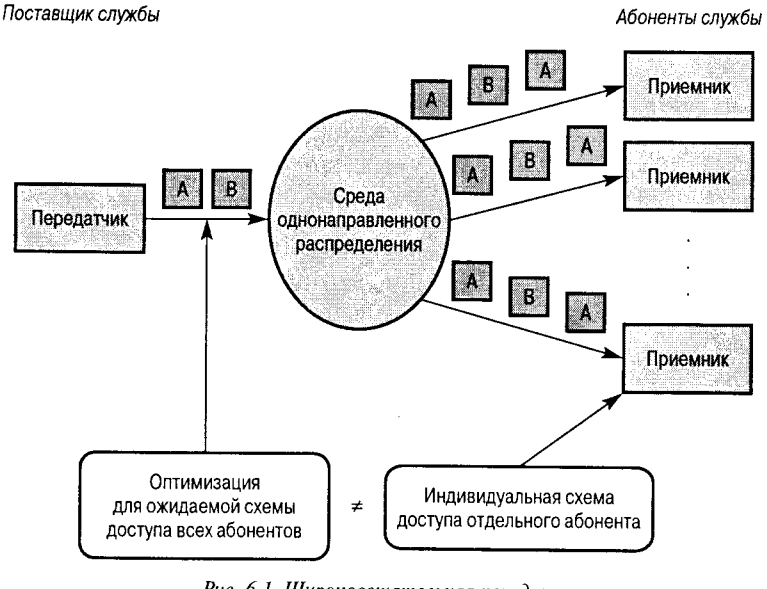

Рис. 6.1. Широковещательная передача

Такие дополнительные механизмы необходимы для персонификации распределенных данных согласно индивидуальным нуждам и приложениям. Очень простым примером этого процесса может быть отсеивание ненужной информации с помощью пользовательского фильт-

ра. Например, радиоприемник внутри машины может лишь предоставлять информацию о движении локальному окружению, а блок памяти телевизора — сохранять только программу фильмов, отсеивая спортивные передачи.

Однако здесь передатчик сталкивается с проблемой — какую информацию в какой момент времени посылать? В следующем разделе приводится несколько решений этого вопроса.

6.2. Циклическое повторение данных

Широковещательному передатчику данных неизвестно, когда приемник начинает прослушивание передачи. Для радио и телевидения это не является проблемой (без прослушивания нельзя получить никакого сообщения). В то же время передачу другой важной информации, например, о состоянии дорожного движения или погодных условиях, необходимо повторять несколько раз. Это делается для того, чтобы приемники могли гарантированно получить такую информацию после прослушивания на протяжении некоторого времени (это похоже на ежечасную трансляцию новостей).

Циклическое повторение данных широковещания в соответствии с проектом [Acharya, 1995] часто называют **широковещательным диском** (broadcast disk). Возможны различные конфигурации такого повторения (на рис. 6.2 показано три примера). Пусть передатчик циклически повторяет три блока данных — A, B и C. При использовании **плоских дисков** блоки чередуются. На каждый блок отводятся равные промежутки времени. Среднее время ожидания блоков A, B и C одинаково. В **скошенных дисках** отдается предпочтение одному или нескольким блокам данных, которые повторяются несколько раз. Это увеличивает вероятность приема повторяемого блока (здесь это блок A), если в первом случае блок был поврежден. Наконец, в **мульти-дисках** блоки, повторяемые более часто, равномерно распределены по циклической конфигурации. Это минимизирует задержку, если пользователю нужен доступ, например, к блоку A.

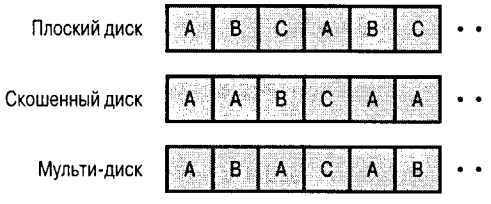

Рис. 6.2. Различные конфигурации широковещания

Описанные конфигурации можно оптимизировать только в том случае, если передатчику что-то известно о содержимом блоков данных и схемах доступа всех пользователей. Допустим, что широковещательным передатчиком является радиостанция, передающая информацию о состоянии дороги (блок A), погодных условиях (блок B) и последних городских событиях (блок C). Блок D, кроме музыки, будет содержать меню доступа ко всем предыдущим блокам. Допустим, что передатчик считает блок D самым главным, поскольку он обеспечивает доступ к остальной информации. Пусть следующим по важности является блок A, затем идет блок B, и наконец, блок C. Один из возможных широковещательных дисков в этом случае выглядит следующим образом:

DADBDADCDADBDADC...

Задача приемника — поместить блоки данных в кэш. Это делается для минимизации задержки доступа, когда пользователю вдруг потребуется определенная информация. Опять же, приемник может оптимизировать кэширование только в том случае, когда ему что-то известно о содержимом блоков данных. В приемнике могут храниться типичные схемы доступа пользователя. С их помощью можно установить, к каким блокам он обратится с большей вероятностью. **Кэширование** обычно определяет-

ся стратегией наименьших затрат, т. е. тем, какими будут затраты пользователя (вызванные временем ожидания), если требуемого блока данных нет в кэше.

В радиоприемниках будущего, возможно, будет учитываться, что событиями в городе пользователи обычно интересуются вечером, а состоянием дорог — утром. Соответственно утром в кэш будет помещаться, скажем, блок А, а вечером — блок С. Такой механизм позволяет уменьшить время ожидания пользователя, если тот постоянно применяет описанную схему доступа.

6.3. Цифровое аудиовещание (DAB)

В основу современных аналоговых радиосистем положен принцип частотной модуляции, изобретенный в 1933 году. В дополнение к передаче звука, радиопрограмма может сопровождаться лишь очень ограниченной информацией (например, идентификацией станции). Качество передачи может изменяться из-за помех и эффектов многолучевого распространения. В полностью цифровой системе **DAB** качество звука сравнимо с качеством записей на компакт-дисках. Кроме того, эта система практически невосприимчива к помехам и эффектам многолучевого распространения [ETSI, 1997a], [WorldDAB, 1999].

В системах DAB можно использовать **сети единой частоты (SFN)**. При этом все передатчики одной радиопрограммы работают на одинаковой частоте. В настоящее время различные передатчики для избежания взаимных помех должны действовать на разных частотах, даже если они транслируют одну и ту же радиопрограмму. Применение сетей SFN повышает эффективность использования частот, поскольку в этом случае для каждой радиостанции в пределах одной страны нужна лишь одна частота. Система DAB работает в диапазонах ОВЧ и УВЧ (например, наземные телеканалы с пятого по двенадцатый используют частоты 174–230 МГц), а также в диапазоне СВЧ (1452–1492 МГц). Для модуляции применяется схема **DQPSK**. Уплотнение в системе DAB происходит согласно схеме **COFDM** (см. главу 2), использующей от 192 до 1536 поднесущих частот (также называемых **ансамблем**) внутри канала DAB шириной 1,5 МГц. Кроме того, для уменьшения количества ошибок применяется схема **FEC**, а между отдельными символами при передаче вводятся **защитные интервалы**. Схема COFDM и применение защитных интервалов сводят к минимуму межсимвольную интерференцию. В системе DAB можно даже выиграть от многолучевого распространения, объединяя сигналы, распространявшиеся по различным путям.

В каждом блоке частот шириной 1,5 МГц система DAB может транслировать до шести стереозвуковых программ со скоростью 192 Кбит/с для каждой из них. В качестве альтернативы доступна служба передачи данных со скоростью до 1,5 Мбит/с. В системе передачи DAB звук является всего лишь одним из возможных типов данных (если не принимать во внимание различные схемы кодирования). Технология DAB использует два основных транспортных механизма.

- **Канал главной службы (main service channel — MSC)**. Через канал MSC передаются все пользовательские данные, в частности, звук и мультимедийные данные. Каналы MSC состоят из **общих чередующихся кадров (common interleaved frame — CIF)** — полей данных с размерами 55 296 бит. Они передаются каждые 24 мс (этот промежуток зависит от режима передачи [ETSI, 1997a]). При этом обеспечивается скорость передачи 2,304 Мбит/с. Кадр CIF состоит из **модулей емкости (capacity unit — CU)** размером по 64 бита. Они являются наименьшими адресуемыми составляющими системы DAB.
- **Канал быстрой информации (fast information channel — FIC)**. Канал FIC содержит **блоки быстрой информации (fast information block — FIB)**, каждый из которых имеет размер 256 бит (16 бит отводится для контрольной суммы). Через канал FIC передается вся управляющая информация, необходимая для конфигурации и интерпретации содержимого канала MSC.

Для каналов MSC определено два транспортных режима. В **потоковом режиме** сквозная передача данных от отправителя к получателю производится через подканал с постоянной скоростью.

Подканал является частью канала MSC и содержит несколько модулей емкости в каждом кадре CIF. Постоянная скорость передачи получается кратной 8 Кбит/с. В **пакетном режиме** данные передаются в виде блоков (пакетов), имеющих адреса. Эти блоки используются для транспорта данных канала MSC через подканалы.

В системе DAB определено много структур обслуживающей информации, сопровождающей поток аудиоданных. Такие **ассоциированные с программой данные** (**PAD**) могут включать информацию о программе, управляющую информацию, изображения для индикации на ЖК-дисплеях, главные меню и т.п. Для кодирования аудиоданных используется импульсно-кодовая модуляция с частотой 48 кГц, а также сжатие MPEG. Скорость передачи потока сжатых аудиоданных может составлять от 8 Кбит/с до 384 Кбит/с. Для повышения устойчивости к резкому увеличению интенсивности информационного обмена производится чередование аудиоданных.

На рис. 6.3 показана общая структура кадров системы DAB. Каждый кадр имеет продолжительность T_F. В зависимости от режима передачи она может равняться 24, 48 или 96 мс. В системе DAB определено четыре различных режима передачи. У каждого из них есть определенные сильные стороны, повышающие эффективность передачи через кабеля, наземные сети или спутники [ETSI, 1997a]. В каждом кадре передается 76 или 153 символа. При этом схема COFDM может использовать 192, 384, 768 или 1 536 различных поднесущих частот. Защитные интервалы T_d для каждого символа могут составлять 31, 62, 123 или 246 мкс.

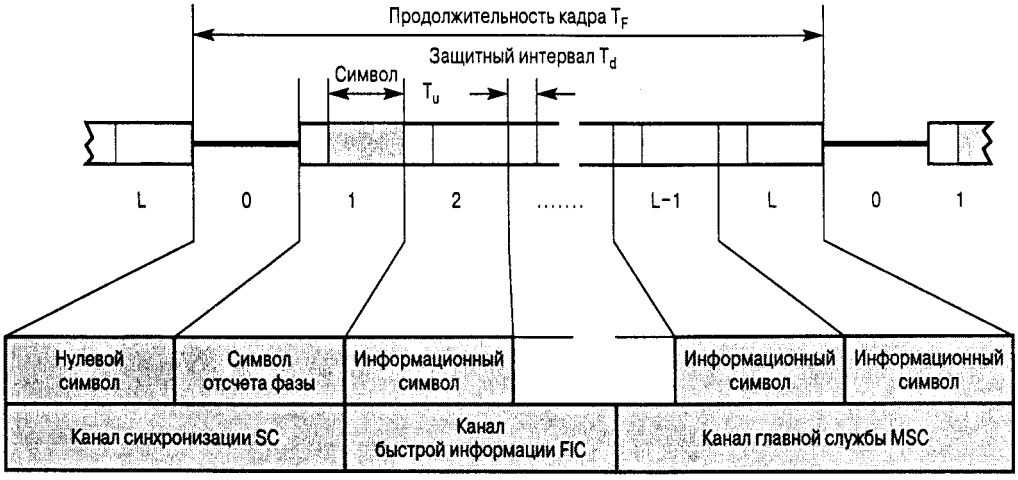

Рис. 6.3. Структура кадров системы DAB

Каждый кадр состоит из трех частей. **Канал синхронизации** (**synchronization channel — SC**) обозначает начало кадра. Он содержит нулевой символ и символ отсчета фазы, которые необходимы для синхронизации приемника. Затем следует **канал быстрой информации** (**FIC**), содержащий управляющие данные в блоках FIB. Наконец, через **канал главной службы** (**MSC**) передается аудиоинформация и другие данные.

На рис. 6.4 приводится упрощенная структурная схема передатчика системы DAB. До передачи аудиоинформация сжимается (по алгоритму MPEG) и кодируется (по схеме FEC). Остальные данные уплотняются и также кодируются. Мультиплексор канала MSC объединяет все потоки пользовательских данных и направляет их на мультиплексор передачи. В этом блоке с помощью чередования каналов FIC создается кадровая структура. В заключение производится кодирование по схеме OFDM, и затем сигнал передается.

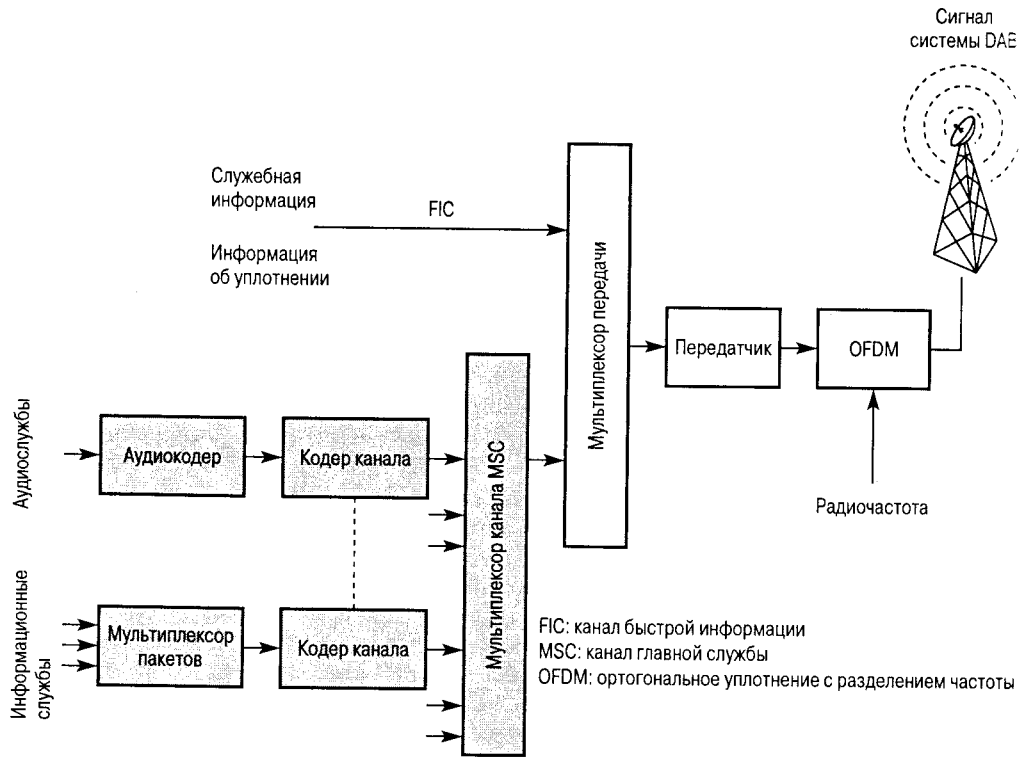

Рис. 6.4. Упрощенная структурная схема передатчика системы DAB

Система DAB не требует фиксированного распределения каналов с определенными свойствами между различными службами. Рис. 6.5 иллюстрирует возможности динамической реконфигурации, производящейся во время передачи. Пусть изначально система DAB транслирует шесть аудиопрограмм различного качества и девять служб данных. С каждой аудиопрограммой связаны соответствующие данные PAD. В приведенном примере программы 1, 2 и 3 характеризуются высоким качеством, программы 4 и 5 — более низким, а программа 6 — самым низким качеством. Такая ситуация может возникнуть, например, когда программы 1–3 представляют собой высококачественную передачу данных, а программа 6 — передачу речи (новости и т.п.). Радиостанция может решить, например, что после начала трансляции новостей программе 3 достаточно скорости 128 Кбит/с. Новости начинают передаваться в режиме моно или стерео с худшим качеством, но зато становятся доступными дополнительные данные (здесь это блоки D10 и D11 — заголовки, иллюстрации и т.д.). Мультиплексор DAB динамически чередует данные из различных источников. Чтобы проинформировать приемник о настоящей конфигурации канала MSC, передающего различные потоки данных, в канал FIC помещается **информация о конфигурации** мультиплексора (**MCI**).

6.3.1. Протокол передачи мультимедийных объектов

При использовании таких технологий как DAB возникает проблема, обусловленная широким диапазоном характеристик приемников. Приемники могут быть простыми устройствами, воспринимающими лишь аудиоинформацию и обладающими однострочными текстовыми дисплеями. Также могут применяться более сложные устройства с дополнительными цветными

графическими дисплеями. В мультимедийных компьютерах приемники DAB можно использовать в качестве адаптеров. Однако приемники всех возможных типов должны, по крайней мере, различать ассоциированные с программой данные и данные, независимые от программы, а также обрабатывать некоторую часть этих данных.

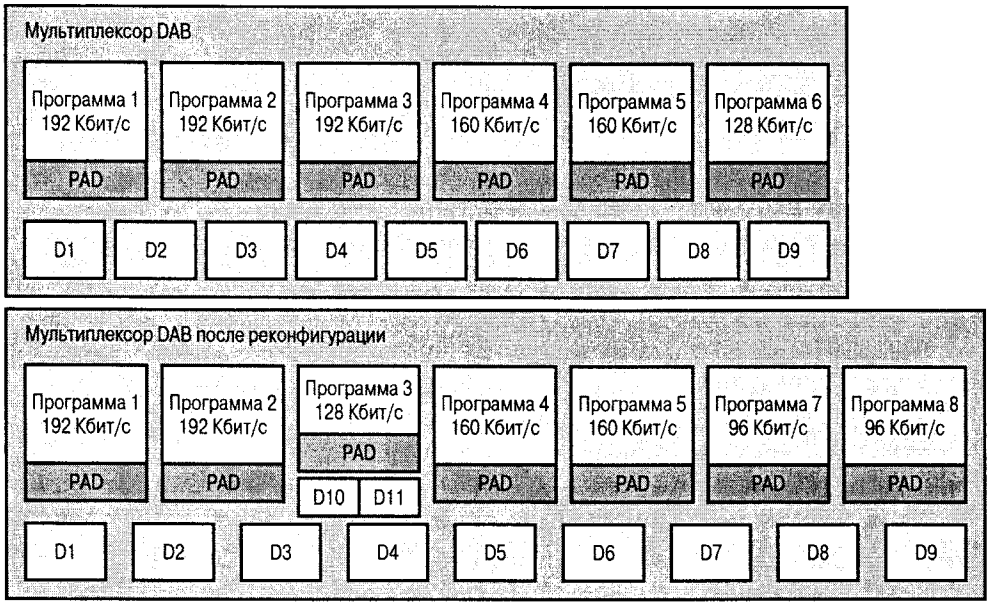

Рис. 6.5. Динамическая реконфигурация мультиплексора системы DAB

Для решения этой проблемы в технологии DAB определен общий стандарт передачи данных — протокол **передачи мультимедийных объектов (multimedia object transfer — MOT** [ETSI, 1998a]. Протокол MOT предназначен, главным образом, для поддержки форматов данных, использующихся в других мультимедийных системах (например, в online-службах, сети Internet, устройствах чтения компакт-дисков). Примеры таких форматов: язык Java, американский стандартный код обмена информацией (ASCII), язык гипертекстовой разметки (HTML) и протокол передачи гипертекстовых файлов (HTTP). Сюда же можно отнести форматы MHEG (группа экспертов по кодированию мультимедийной и гипермедийной информации), JPEG (объединенная группа экспертов в области фотографии) MPEG (экспертная группа по вопросам движущегося изображения), BMP (битовое отображение) и GIF (формат графического обмена).

Данные MOT передаются в виде объектов MOT, состоящих из ядра заголовка, расширения заголовка и тела (рис. 6.6):

Рис. 6.6. Структура объекта MOT

- **Ядро заголовка**. Это поле имеет размер 7 байт. В нем указываются размеры заголовка и тела, а также тип объекта. С помощью этой заглавной информации приемник определяет, доста-

точно ли у него ресурсов (памяти, мощности процессора, возможностей дисплея и т.п.) для декодирования и дальнейшей обработки объекта.

- **Расширение заголовка**. Поле расширения имеет переменный размер и содержит дополнительную информацию об объекте, например, приоритет данных, сведения о сегментации и периоды повторения, необходимые для поддержки расширенных методик кэширования (см. раздел 6.2). Информация о приоритете нужна, чтобы определить, какие данные помещать в кэш. К примеру, страница HTML с алфавитным указателем может иметь больший приоритет, чем произвольная страница.
- **Тело**. Тело также имеет переменный размер и используется для передачи произвольных данных согласно полям заголовка.

Более крупные объекты MOT разбиваются на сегменты. В технологии DAB к объектам и сегментам могут применяться различные схемы чередования и повторения.

- **Повторение объектов**. При передаче согласно технологии DAB один объект может повторяться несколько раз. Если объект A состоит из четырех сегментов (A_1, A_2, A_3 и A_4), то соответствующей схемой повторения будет последовательность $A_1A_2A_3A_4A_1A_2A_3A_4A_1A_2A_3A_4$... .
- **Чередующиеся объекты**. Для избежания проблем, вызванных пакетами ошибок, можно чередовать сегменты различных объектов. Чередование объектов A, B и C может давать последовательность $A_1B_1C_1A_2B_2C_2A_3B_3C_3$... .
- **Повторение сегментов**. Если некоторые сегменты являются более важными, технология DAB позволяет повторять их чаще (например, $A_1A_1A_2A_2A_2A_3A_4A_4$...).
- **Повторение заголовков**. Если приемнику неизвестен заголовок объекта MOT, он не сможет декодировать сам объект. Поэтому может оказаться полезной повторная передача заголовка. После получения заголовка приемник синхронизируется с потоком данных и начинает декодирование. Соответствующая последовательность выглядит следующим образом: $HA_1A_2HA_3A_4HA_5A_6H$..., где H — заголовок объекта A.

Ясно, что все эти схемы чередования и повторения при передаче согласно технологии DAB можно использовать одновременно.

6.4. Цифровое видеовещание (DVB)

После применения цифровых технологий в радиовещании естественно проделать то же самое и с традиционной телевизионной системой. Используемая в настоящее время аналоговая система на протяжении десятилетий оставалась неизменной. Единственным изменением, о котором стоит здесь упомянуть, стало освоение массового рынка цветным телевидением, происходившее в 60-е годы. Телевидение все еще использует низкое разрешение — 768 × 576 в европейской системе PAL и 720 × 460 в американской системе NTSC. Дисплеи в этих системах чередуют, соответственно, 25 и 30 кадров в секунду. Современные компьютерные дисплеи имеют разрешение 1 280×1 024, частоту смены кадров 75 Гц, чередование в них не происходит. В сравнении с ними характеристики телевидения выглядят удручающе.

Предпринималось много попыток изменить эту ситуацию и ввести цифровое телевидение с более высоким разрешением, более качественным звуком и дополнительными возможностями. Однако ни одна из таких попыток не увенчалась успехом. Одной из причин этого является огромное количество установленных старых систем, которые не заменяются так же быстро, как компьютеры (самый новый фильм можно посмотреть по старому телевизору, однако новое программное обеспечение на старом компьютере запустить

нельзя). Кроме того, появлению общего стандарта цифрового телевидения препятствуют изменяющиеся политические и экономические интересы. Один из подходов к установлению такого стандарта представлен в следующих разделах.

После того, как попытки внедрения цифрового телевидения в нескольких государствах потерпели неудачу, в 1991 году была основана так называемая Европейская Группа Запуска (European Launching Group). Ее целью была разработка общей европейской цифровой телевизионной системы. В 1993 году эти общие усилия привели к созданию технологии **цифровой широковещательной передачи видеоинформации (digital video broadcasting — DVB)** [Reimers, 1998], [DVB, 1999]. Несмотря на то, что сами названия подчеркивают некоторое сходство систем DAB и DVB, между ними существуют фундаментальные различия в технологии передачи, частотах, модуляции и т.п. Целью технологии DVB является установление цифрового телевещания, использующего спутниковую передачу (DVB-S, [ETSI, 1997в]), кабельную технологию (DVB-C, [ETSI, 1998б]), а на более поздних этапах — и наземную передачу (DVB-T, [ETSI, 1997б]).

На рис. 6.7 показаны компоненты, интегрированные в структуру системы DVB. В центре находится объединенный приемник/декодер (приставка), соединенный с монитором высокого разрешения. Эта приставка принимает сигналы DVB через спутники, наземные местные/региональные передатчики (многоточечные системы распределения, наземные приемники), кабеля, сети B-ISDN и ADSL. В будущем возможно применение и других технологий. При использовании кабеля или соединений через сети ADSL и B-ISDN предоставляется и обратный канал, т.е. пользователь может посылать некоторые данные (например, выбор канала, идентификационную информацию или список покупок). Кроме того, потоки аудио- и видеоданных могут записываться, обрабатываться и воспроизводиться с помощью **универсальных цифровых дисков (DVD)**, **цифровых видеомагнитофонов (DVTR)** или мультимедийных компьютеров. Предусмотрены различные уровни качества: **телевидение стандартной четкости (SDTV)**, **телевидение улучшенной четкости (EDTV)** и **телевидение высокой четкости (HDTV)**. Разрешение может достигать 1 920×1 080 пикселей.

В системе DVB, подобно системе DAB, для передачи данных используются гибкие контейнеры. В основу этих контейнеров положены кадры MPEG-2 без ограничений на тип информации. Частью потока данных DVB является обслуживающая информация, описывающая содержимое контейнеров. Определены следующие типы содержимого.

- **Таблица сетевой информации (NIT)**. Таблица NIT содержит список всех услуг поставщика и дополнительную информацию для приставок.
- **Таблица описания служб (SDT)**. В таблице SDT приводятся названия и параметры всех служб уплотненного канала MPEG.
- **Таблица информации о событиях (EIT)**. В таблице EIT находится информация о состоянии текущей передачи и некоторая дополнительная информация для приставок.
- **Таблица времени и дат (TDT)**. Наконец, в таблице TDT содержится информация об обновлении для приставок.

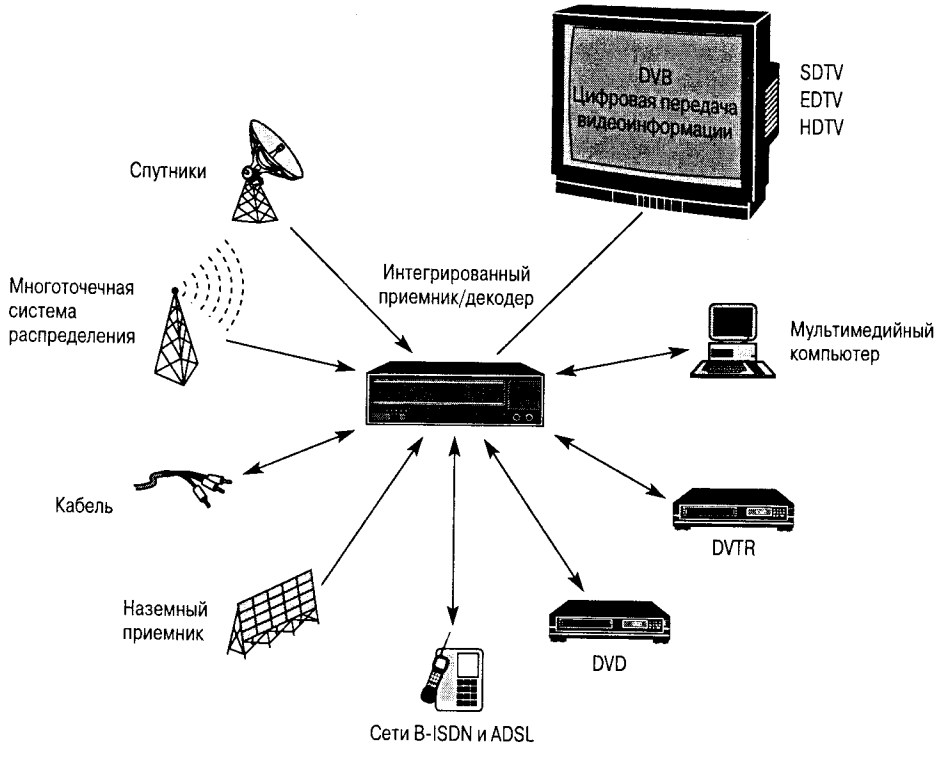

Рис. 6.7. Схема цифровой передачи видеоинформации

В контейнере MPEG-2/DVB могут храниться различные типы данных (рис. 6.8). Контейнер может содержать один канал уровня HDTV, несколько каналов уровня EDTV или SDTV, или произвольные мультимедийные данные (передача данных).

Систему DVB можно использовать для высокоскоростного асимметричного доступа к Internet, даже не принимая во внимание описанную выше передачу мультимедийных данных. Типичная схема такого доступа выглядит следующим образом (см. рис. 6.9): поставщик информации (например, видеомагазин) с помощью поставщика услуг предлагает свои данные потенциальным покупателям. Если покупателю нужно загрузить информацию большого объема, поставщик информации передает эти данные поставщику спутниковой связи через поставщика услуг. В стационарных сетях для этого обычно применяются выделенные линии, так как они обеспечивают высокую скорость передачи и предоставляют гарантированное качество обслуживания. Поставщик спутниковой связи уплотняет полученный поток данных вместе с остальными цифровыми телеканалами и через спутник и спутниковый приемник передает их покупателю. Покупатель получает запрошенную информацию с помощью адаптера DVB, встроенного в мультимедийный компьютер. В большинстве случаев информация кодируется. Это делается, чтобы информацию могли использовать лишь оплатившие ее покупатели. В качестве обратного канала для запросов можно использовать обычное (TCP/IP)-соединение через Internet, поскольку для этого канала достаточно и низкой скорости передачи.

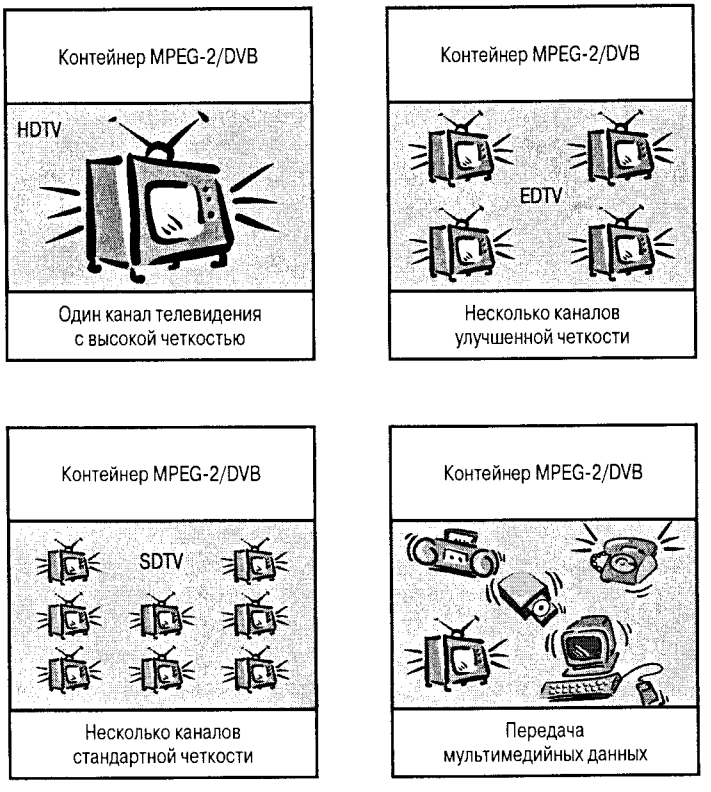

Рис. 6.8. Различные типы содержимого контейнеров MPEG-2/DVB

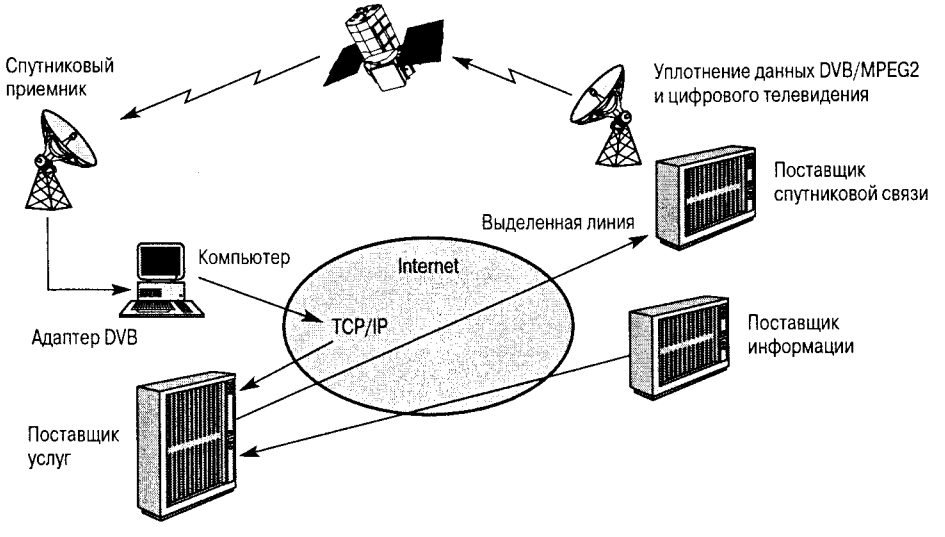

Рис. 6.9. Высокоскоростной доступ к Internet с использованием технологии DVB

Планируется, что скорость передачи для нисходящей линии связи через спутник в расчете на одного пользователя будет составлять 6–38 Мбит/с. Скорость передачи в обратном канале при использовании обычного модема равна 33 Кбит/с. Для сети ISDN эта величина составляет 64 Кбит/с, а для сети ADSL она может достигать нескольких сотен Кбит/с. Преимуществом рассмотренного подхода является то, что информация передается вместе с телепрограммами. Для этого используется свободное пространство в передаваемом потоке данных. Таким образом, этот подход не требует наличия дополнительных устройств и линий связи. Этот фактор особенно важен для удаленных областей и развивающихся стран, где высокоскоростной проводный доступ (например, через сеть ADSL) недоступен. Очевидным недостатком описанного подхода является наличие общей среды — "спутника". Если данные через систему DVB запрашивают много пользователей, пропускная способность спутника разделяется между ними. Поэтому без существенного увеличения стоимости в этой системе нельзя обеспечить жесткие гарантии качества обслуживания всем пользователям.

6.5. Резюме

В этой главе рассматривались два примера широковещательных технологий. Они выделяются среди остальных технологий мобильной связи, представленных в этой книге. Системы DAB и DVB являются наиболее вероятными преемниками обычного радио и телевидения во многих странах (не везде, из-за меняющихся политических и экономических интересов). Кроме передачи потоков аудио- и видеоинформации, в этих системах возможна передача потоков мультимедийных данных. Несмотря на то, что обе технологии поддерживают лишь однонаправленную связь, они являются неотъемлемой частью будущих мобильных коммуникаций. Системы DAB и DVB будут рентабельно использоваться для распределения массовых данных. В тех случаях, когда необходим обратный канал, будут применяться другие беспроводные технологии с более низкой скоростью передачи. Технологии DAB и DVB, объединяя радио/телесигналы с персонализированными потоками мультимедийных данных, поддерживают грядущее объединение компьютерной индустрии с индустрией средств связи и развлечений. Можно представить себе ситуацию, когда фильмы передаются всем пользователям, а, например, реклама является индивидуализированной и зависит от интересов пользователя. Приставка объединяет оба потока данных, и пользователь наблюдает, например, за футбольным матчем с полностью индивидуализированными рекламными щитами. Еще одна черта технологии DAB делает ее особенно привлекательной для мобильной связи. Система DAB является единственной коммерческой радиосистемой, применимой при высоких скоростях: пропускная способность достигает 1,5 Мбит/с даже при скорости 900 км/ч! Это дает возможность устанавливать телевизоры в поездах и других средствах передвижения, подвергающихся влиянию многолучевого распространения при использовании других технологий. То, что технология DVB не подходит для быстродвижущихся приемников, компенсируется способностью системы DAB передавать видеоинформацию.

6.6. Упражнения

1. Технологии GSM и GSM следующего поколения могут применяться для передачи данных. Сравните эти подходы с технологией DAB и назовите преимущества и недостатки ее использования.
2. Какие Web-страницы целесообразно распределять через систему DAB или DVB?
3. Как можно объединить глобальную систему навигации и определения положения (GPS) с технологией DAB?

6.7. Рекомендуемая литература

Acharya, S.; Franklin, M.; Zdonik, S. (1995) "Dissemination-based data delivery using broadcast discs", *IEEE Personal Communications*, 2, (6).

DVB Project Office (1999), `http://www.dvb.org/`.

ETSI (1997a) *Digital Audio Broadcasting (DAB) to mobile, portable, and fixed receivers*, European Telecommunications Standards Institute, ETS 300 401.

ETSI (1997б) *Framing structure, channel coding and modulation for digital terrestrial television*, European Telecommunications Standards Institute, ETS 300 744.

ETSI (1997в) *Framing structure, channel coding and modulation for 11/12 GHz satellite services*, European Telecommunications Standards Institute, ETS 300 421.

ETSI (1998a) *DAB Multimedia Object Transfer (MOT) protocol*, European Telecommunications Standards Institute, ETS 301 234.

ETSI (1998б) *Framing structure, channel coding and modulation for cable systems*, European Telecommunications Standards Institute, ETS 300 429.

ETSI (1999) *Digital Audio and Video Broadcasting*, European Telecommunications Standards Institute, `http://www.etsi.org/`.

Reimers, U. (1998) "Digital Video Broadcasting", *IEEE Communications Magazine*, 36, (6).

World DAB Forum (1999) `http://www.worlddab.org/`.

Беспроводные локальные сети 7

В этой главе рассматриваются некоторые технологии построения беспроводных локальных сетей (wireless local area network — WLAN, БЛВС), составляющих быстро развивающийся рынок, который позволяет использовать гибкий беспроводный доступ дома, в офисе или на производстве. В отличие от технологий, описанных в главах 4–6, радиус охвата БЛВС ограничен размерами отдельной комнаты, здания, студенческого городка и т. п., а предоставляются эти сети отдельными, не глобальными сетевыми поставщиками. Основной целью беспроводных локальных сетей является замена существующих кабельных локальных сетей и, дополнительно, введение высокой гибкости в эпизодическую связь, например, групповые встречи. Ниже перечислены основные достоинства и недостатки беспроводных локальных сетей по сравнению с их проводными аналогами.

Достоинства беспроводных локальных сетей.

- **Гибкость**. Радиоволновая физическая природа беспроводных локальных сетей обеспечивает большую степень свободы пользователей сети. Для радиосетей нет ограничений, характерных для кабельных сетей. Радиоволны беспрепятственно проникают сквозь видимые преграды, приемники и передатчики могут располагаться где угодно, например, скрыты в стенах и т. д. Для кабельных сетей требования пожарной безопасности в зданиях и помещениях ограничивают возможности быстрого проектирования и разворачивания сети.
- **Планирование**. Поддерживая в своих устройствах режим эпизодической связи, можно построить разветвленные сети с самой различной топологией соединений. Для проводных сетей в первую очередь необходимо проектирование кабельной инфраструктуры; возможны проблемы с правильным подключением, иногда требуется установка коммутаторов.
- **Проектирование**. Современные устройства связи и обработки информации компактны и необременительны (PDA, электронные записные книжки). Беспроводные сети с их возможностями скрытого развертывания — гармоничное дополнение с точки зрения дизайна и удобств эксплуатации
- **Устойчивость**. Беспроводные сети легче проводных переносят стихийные бедствия и разрушения, обеспечивая связь. Кабельные сети в таких случаях отказывают полностью вследствие обрывов.

Слабые стороны беспроводных локальных сетей.

- **Качество обслуживания**. Параметры связи, которые обеспечивают беспроводные локальные сети, обычно хуже, чем параметры их проводных аналогов. В первую очередь это связано с меньшей пропускной способностью радиоканала, только 1–10 Мбит/с. Слабая помехозащищенность при-

водит к увеличению частоты появления ошибок передачи (10^{-4} по сравнению с 10^{-10} для оптоволоконных сетей), большим задержкам/увеличению дисперсии задержек.

- **Высокая стоимость**. Скоростные Ethernet-адаптеры стоят около 10', тогда как цена на адаптеры беспроводных локальных сетей, например, PC-карты, доходит до 100'.
- **Частные решения**. Вследствие медленного прохождения процедур стандартизации в области связи многие компании используют свои собственные разработки и модификации при создании сетей (например, увеличение пропускной способности канала за счет использования различных методов сжатия данных). Однако, эти модификации работают только в однородной среде, и все сетевые адаптеры в узлах сети должны быть одного производителя.
- **Ограничения**. Регулирование в области беспроводных технологий связи производится государственными органами. В мире существуют правительственные и неправительственные институты, занимающиеся стандартами, ограничивающими и регулирующими деятельность в области радиосвязи, распределением частотного диапазона для уменьшения интерференции. В рамках национальных программ процессы становления занимают длительное время, пример — IMT-2000. Поэтому беспроводные локальные сети ограничиваются маломощными передатчиками, и для них выделяется нелицензируемый диапазон частот.
- **Безопасность и защита**. Передача данных с использованием радиочастот подвержена помехам со стороны технологичных приборов и устройств. Поэтому должны быть предусмотрены соответствующие меры предосторожности. Кроме того, открытый радиоинтерфейс делает возможным несанкционированный доступ к данным, тогда как "прослушивание" оптоволоконных сетей сложнее.

Разные и порой противоречивые идеи были заложены в основу беспроводных локальных сетей с целью обеспечения их коммерческого успеха.

- **Распространение во всем мире**. Беспроводные локальные сети продаются по всему миру, и поэтому возникают вопросы государственного и межгосударственного частотного регулирования. В отличие от беспроводных глобальных сетей оборудование для малых сетей может передаваться из одной страны в другую, поэтому эксплуатация сетей должна быть узаконена и согласована.
- **Энергосбережение**: Устройства, связывающиеся через беспроводные локальные сети, — это обычно мобильные беспроводные аппараты, питающиеся от батарей. С учетом этого разработчики локальных сетей проектируют энергосберегающие варианты, модели с функциями управления энергопотреблением. Устройства беспроводной связи лишь в некоторых случаях используют стационарное сетевое питание от электророзеток, когда нет необходимости заново проводить электропитание в зданиях. Однако, будущее принадлежит носимым устройствам без ограничивающих проводов.
- **Безлицензионная эксплуатация**. Операторы беспроводных локальных сетей не спешат получать лицензию на использование радиочастот для своего оборудования. Поэтому пока используется нелицензируемая во всем мире частота 2,4 ГГц (диапазон ISM — "для промышленного, научного и медицинского применения").
- **Помехоустойчивые технологии связи**. По сравнению со своими проводными аналогами беспроводные локальные сети эксплуатируются в более сложных условиях. Так как сети используют радиосвязь, возможны помехи, наводимые другими электрическими приборами. Кроме того, применяемые в БЛВС ненаправленные антенны не всегда обеспечивают надежную связь.

- **Самоорганизующаяся сеть со случайными связями**. Используется на практике, когда нужно быстро развернуть сеть, полагаясь на вероятную возможность связи между узлами. Такие БЛВС не требуют сложных процедур инициализации и запускаются сразу при включении, однако, они не обеспечат всех возможностей эпизодических сетей.
- **Простота в использовании**. В отличие от больших территориальных сетей локальные сети разрабатывались для отдельных пользователей. Для них не требуется централизованное управление, и они работают по принципу "подключай и работай".
- **Защита инвестиций**. В кабельные сети уже вложены большие средства. Новые БЛВС могли бы сохранить сделанные ранее вложения, взаимодействуя с проводными сетями посредством простых мостов и поддерживая стандарты существующих сетей.
- **Безопасность**. БЛВС должны быть безопасны в эксплуатации, особенно это касается уменьшения вредного электромагнитного излучения. Важным условием эксплуатации сети является обеспечение надежной защиты от несанкционированного доступа или нарушения работы в результате прослушивания среды.
- **Незаметность для приложений**. Существующие приложения должны работать в БЛВС с той лишь разницей, что возможны задержки при трансляции данных, уменьшение пропускной способности каналов. Факт беспроводного доступа должен быть скрытым, но сеть должна поддерживать приложения, локализующие пользователя.

Следующий раздел — это введение в два основных способа беспроводной передачи данных: радиочастотный и в инфракрасном диапазоне спектра. Также рассматриваются два основных класса БЛВС — сети с явно выраженной инфраструктурой и эпизодические сети. Три основных раздела настоящей главы посвящены, соответственно, стандарту IEEE 802.11, европейскому стандарту ETSI для высокоскоростных БЛВС, именуемому HIPERLAN1, и промышленному подходу к беспроводным персональным сетям, т.е. БЛВС малого охвата, Bluetooth.

7.1. Инфракрасный диапазон против радиопередачи

В настоящее время в БЛВС используются два основных способа передачи данных. Это технология с использованием инфракрасного излучения (например, с длиной волны 900 нм) и связь с помощью радиоволн в гигагерцевом диапазоне (например, нелицензируемый ISM диапазон 2,4 ГГц). Обе эти технологии поддерживают организацию эпизодических сетей. Например, обеспечивают информационный обмен между настольным ПК и удаленным принтером, поддерживают мобильность пользователей сети небольшого радиуса действия.

Возможны два способа **связи в инфракрасном диапазоне**. В одном случае приемник и передатчик находятся в зоне прямой видимости и используют направленное излучение либо устанавливаются адаптеры так называемого диффузного типа, передатчики которых излучают сигналы в полусферической области. Такие ИК-лучи претерпевают многократное отражение от потолка, стен и прочих поверхностей, что позволяет переносить мобильные ПК в пределах помещения, не теряя контакта с ИК-устройством. В качестве излучателя ИК-лучей могут использоваться обычный светодиод либо лазерный диод. Приемником сигнала является фотодиод. Подробное описание технологии (принципы модуляции, факторы ухудшения качества) можно найти в работах [Wesel, 1998] и [Santamaria, 1994].

- К основным **преимуществам** ИК-технологии относится низкая стоимость приемников и передатчиков инфракрасного излучения. Большинство портативных ПК, ноутбуков, PDA, мобильных телефонов снабжены адаптерами ИК-лучей. Версия 1.0 промышленного стандарта ИК-технологии устанавливает максимальную скорость передачи данных в 115 Кбит/с. По версии 1.1 пропускная способность ИК-каналов увеличена до 1,152 Мбит/с

и 4 Мбит/с. Использование ИК-диапазона не лицензируется, передача данных в этом диапазоне мало подвержена помехам от электроприборов.

- К **недостаткам** сетей ИК-диапазона относится небольшая пропускная способность каналов. Ограничение скорости передачи 4 Мбит/с существенно по сравнению с проводными локальными сетями. Причем наибольшая скорость и качество связи достигаются в зоне прямой видимости передатчика и приемника, любой предмет на пути ИК-лучей является препятствием для их прохождения.

Почти все БЛВС, рассматриваемые в настоящей книге, используют радиоволновый способ передачи данных. Аналогами в мобильной беспроводной телефонной связи в радиоволновом диапазоне являются стандарт GSM 900, 1 800 и 1 900 Мгц, а также стандарт DECT 1 880 Мгц.

- К **достоинствам** радиопередачи относится большой опыт разработок в радиосвязи и передаче данных мобильных сотовых систем и территориально-распределенных сетей. Радиоволновый способ передачи имеет меньше ограничений при распространении сигнала и большую пропускную способность (10 Мбит/с) по сравнению с ИК-лучами.
- В тоже время слабая помехозащищенность радиосигнала от электрических возмущений и межсетевых влияний, а также необходимость получения лицензии на использование радиочастот для своего оборудования являются **недостатками** использования радиоволнового диапазона.

Из трех технологий БЛВС, рассматриваемых в данной главе, только стандарт IEEE 802.11 наряду с радиопередачей описывает сети на основе ИК-лучей. Сети HIPERLAN и Bluetooth, изучаемые в этой главе, работают в радиоволновом диапазоне. Слабая проникающая способность ИК-лучей, необходимость прямой видимости для приемников и передатчиков ограничивают применение ИК-диапазона пределами небольшого помещения без видимых преград. Будущее мобильных коммуникаций принадлежит малым переносным устройствам, легко помещающимся в кармане или кейсе. Использование ИК-лучей в этом случае вряд ли возможно.

7.2. Сети с инфраструктурой и эпизодические сети

Беспроводная сеть с явно выраженной структурой исключает непосредственную связь периферийных станций (узлов). Связь между ними осуществляется через точки доступа (рис. 7.1). Точки доступа могут играть роль мостов, связывающих несколько сетей, беспроводных или кабельных. Несколько физически разделенных сетей, объединенных таким образом, могут составить одну сеть на логическом уровне. Рис. 7.1 иллюстрирует объединение трех беспроводных сетей посредством их точек доступа и проводной сети. Радиус действия такой сети гораздо больше, чем физическое радиопокрытие каждой БЛВС в отдельности.

Прототипом БЛВС с инфраструктурой являются локальные кабельные сети Ethernet, имеющие звездообразную топологию, обеспечивающую полноценное управление сетью. Все функции управления в сети возлагаются на базовую станцию. Точка доступа регулирует информационные потоки, определяет качество связи и пропускную способность каждого узла. Сети с инфраструктурой поддерживают протоколы множественного доступа с обнаружением несущей и избежанием конфликтов, т.е. ситуаций, когда две периферийные станции пытаются начать передачу одновременно. Постоянная структура с определенной топологией характерна не только для кабельных сетей. Системы мобильной сотовой телефонной связи, спутниковой телефонной связи построены по такому же принципу. Статическая природа сетей с инфраструктурой не позволяет проявиться всем возможностям, которые могли бы обеспечить беспроводные сети.

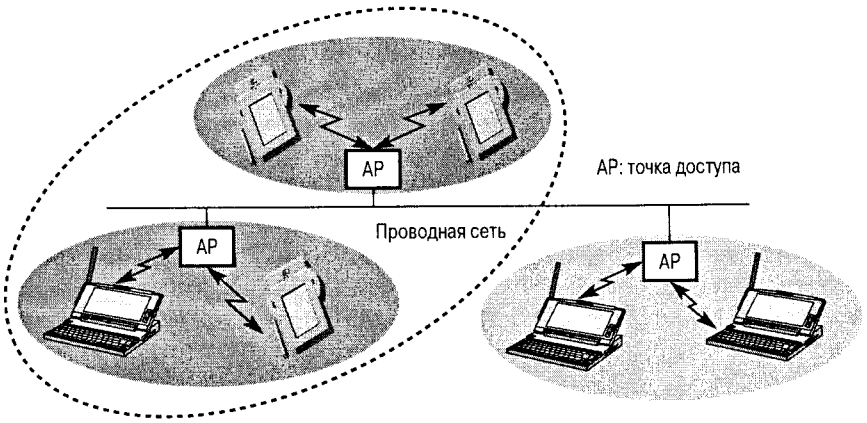

Рис. 7.1. Пример трех беспроводных сетей с инфраструктурой

Беспроводные эпизодические сети позволяют отказаться от использования точки доступа — базовой станции, управляющей механизмом доступа к среде обмена. Узлы сети могут устанавливать непосредственную связь между собой. С помощью такого режима можно построить разветвленную беспроводную сеть с различной топологией, гибко изменять структуру, быстро перенастраивать сценарии связи и обмена информацией (например, быстро организовать конференц-связь максимально большого числа узлов сети). На рис. 7.2 изображены две эпизодических сети, каждая из них состоит из трех узлов. Если узлы находятся в пределах радиовидимости, возникает среда обмена информацией. Связь между узлами разных сетей невозможна, так как узлы одной сети находятся вне досягаемости узлов другой.

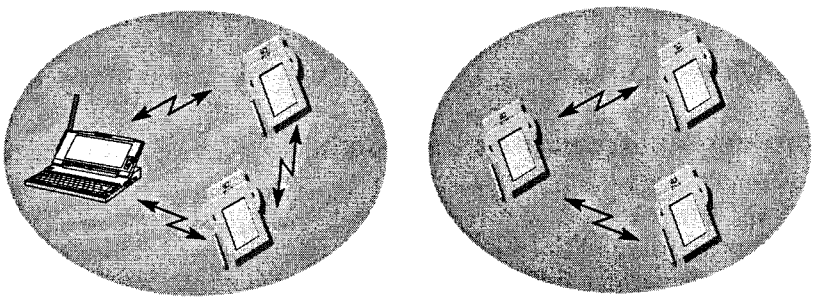

Рис. 7.2. Пример двух эпизодических беспроводных сетей

Адаптеры беспроводной связи в узлах эпизодической сети гораздо сложнее с функциональной точки зрения, так как должны обеспечить доступ к среде передачи данных, приемлемый уровень качества связи, предупреждение возможности возникновения изолированных узлов. На практике сети рассматриваемых классов редко применяются в чистом виде. Обычно действующие сети являются смешанными и имеют свойства, характерные для каждого класса сетей. Базовые станции, являющиеся точками доступа, управляют средой обмена. В то же время возможна прямая связь узлов сети между собой. Некоторые узлы эпизодических сетей могут выполнять функции ретрансляторов. Остальные узлы связываются с ними, когда необходим обмен с узлами, находящимися за пределами досягаемости радиосвязи. Из рассматриваемых в данной главе БЛВС сети стандарта IEEE 802.11 (рис. 7.3) и HIPERLAN (рис. 7.4) являются типичными сетями с выраженной структурой, но в то же время они поддерживают эпизодическое соединение.

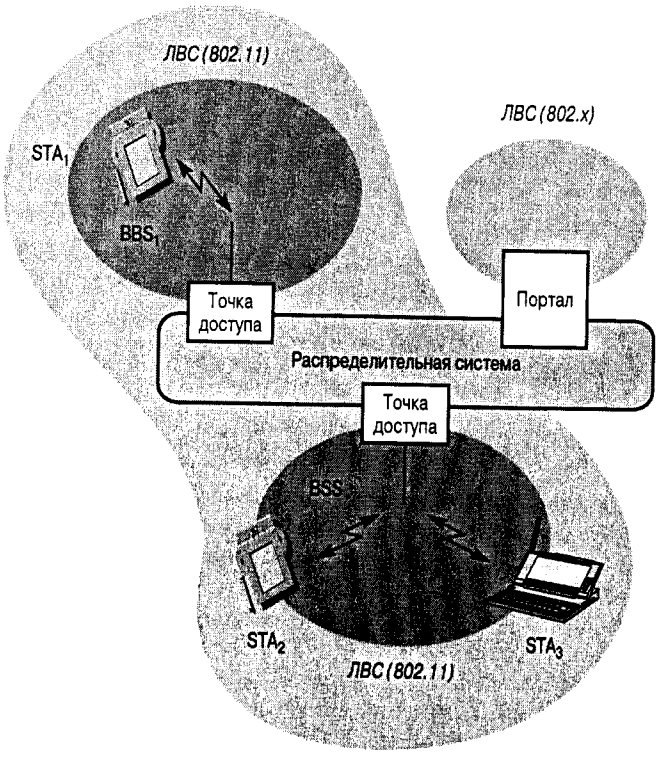

Рис. 7.3. Архитектура сети IEEE 802.11 с инфраструктурой

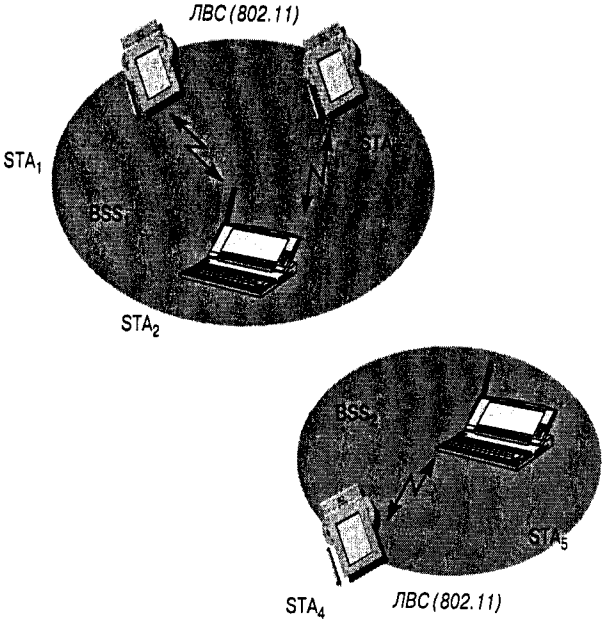

Рис. 7.4. Архитектура беспроводных эпизодических сетей IEEE 802.11

Сети стандарта HIPERLAN нельзя однозначно отнести к сетям с инфраструктурой, но на практике они обычно имеют явно выраженную топологию. Сеть Bluetooth (рис. 7.5), рассматриваемая в данной главе, является типичной эпизодической сетью без элементов централизованного управления.

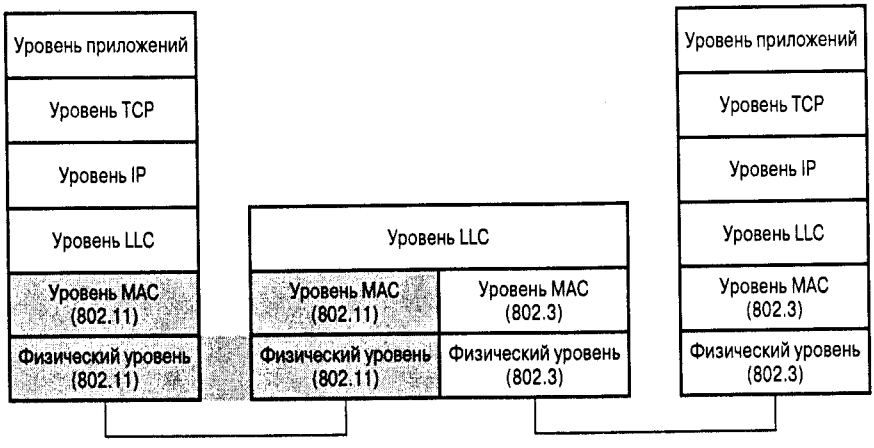

Рис. 7.5. Протокольная архитектура и мосты IEEE 802.11

7.3. Стандарт IEEE 802.11

Стандарт IEEE 802.11, утвержденный в 1997 году, входит в серию стандартов IEEE 802.x, относящихся к сетям и коммуникациям. Эта серия включает стандарт локальных проводных сетей 802.3 Ethernet, 802.5 Token Ring. Стандарт IEEE 802.11 определяет компоненты и характеристики сети на физическом уровне передачи данных и на уровне доступа к среде (medium access control — MAC) с учетом беспроводного способа передачи данных и возможности взаимодействия с существующими сетями.

Для передачи данных на физическом уровне стандарт определяет инфракрасный диапазон спектра и радиочастотный сигнал широкополосного спектра. Необходимый диапазон частот 2,4 ГГц относится к безлицензионному диапазону ISM, в большинстве стран мира выделенному для использования в промышленных, научных и медицинских целях. Максимально достигаемая скорость передачи данных от 1 Мбит/с до 2 Мбит/с. Стандарт устанавливает повышенные требования к энергосберегающим функциям элементов БЛВС и идентификации изолированных узлов. Главное назначение стандарта —определение спецификаций надежных и простых БЛВС, поддерживающих приложения срочной связи. На уровне MAC стандарт требует обеспечения доступа к протоколам более высокого уровня и физической среде передачи данных независимо от природы используемой среды.

7.3.1. Архитектура беспроводных сетей

Как уже было сказано в разделе 7.2, архитектура БЛВС делится на два основных класса — сети с явно выраженной инфраструктурой и эпизодические сети. Стандарт IEEE 802.11 определяет два компонента архитектуры беспроводных сетей с инфраструктурой. Это выделенная **точка доступа** (базовая станция) и STA_i — периферийные **станции** (узлы сети). Станции — это терминальное оборудование с возможностью входа в беспроводную среду сети посредством связи с точкой доступа. AP и периферийные станции, находящиеся в области радиопокрытия AP, составляют BSS_j — **базисный набор служб** (Basic Service Set) сети. Стандарт предполагает наличие нескольких базисных наборов служб, объединенных в одну единую **систему распределения**

(Distribution System — DS). DS соединяет несколько базисных наборов служб посредством точек доступа в одну сеть, таким образом расширяя область охвата сети. БЛВС, включающая различные наборы служб, их точки доступа и систему распределения, называется **расширенным набором служб** (Extended Service Set — ESS). Кроме того, стандарт определяет понятие **портала**. Это устройство, соединяющее БЛВС и любую другую локальную сеть через систему распределения и точки доступа. Пример на рис.7.3 иллюстрирует две базисные службы, BSS_1 и BSS_2, объединенные в сеть системой распределения.

Система распределения подробно не специфицируется в стандарте, она может включать сетевые мосты, беспроводные каналы связи и т. д.

Периферийные станции обычно связаны с одной точкой доступа. Стандарт определяет механизм роуминга, позволяющий мобильной станции перемещаться в пределах досягаемости нескольких AP, с которыми возможна связь по разным каналам. На AP возложены функции синхронизации станций, поддержки приложений, ограниченных временными интервалами. Эти и другие характеристики системной архитектуры сетей рассматриваются в следующих разделах.

Стандарт предусматривает возможность построения эпизодических сетей. В примере на рис. 7.4 изображены две системы BSS без точек доступа. Каждая из ячеек включает группу станций, работающих на одной радиочастоте. STA_1, STA_2 и STA_3 относятся к BSS_1, STA_4 и STA_5 и входят в BSS_2. В пределах одной системы возможна прямая связь станций, но связь между станциями разных ячеек невозможна. Чтобы области действия нескольких BSS не перекрывались, они должны быть разнесены на достаточное расстояние либо работать на разных несущих частотах. IEEE 802.11 не поддерживает создание специальных узлов, поддерживающих маршрутизацию, перенаправление информации или изменение топологии сети, как это возможно в HIPERLAN (см. раздел 7.4).

7.3.2. Протокольная архитектура

Стандарт IEEE 802.11 входит в серию стандартов IEEE 802.x, связанных с сетями и коммуникациями [Halsall, 1996]. Рис. 7.5 иллюстрирует наиболее общую и часто встречающуюся схему взаимодействия беспроводных сетей IEEE 802.11 и кабельных локальных сетей Ethernet IEEE 802.3 с помощью мостов. Работа сетевых приложений, протоколов передачи TCP, IP не зависит от физической среды обмена, взаимодействие узлов обеспечивает проводная или беспроводная сеть. Возможно только изменение скорости передачи, связанное с различной пропускной способностью каналов сетей различной природы. Различают два подуровня уровня управления каналом передачи данных (Data Link Control — DLC) — верхний, LLC (Logical Link Control), подуровень управления логическим каналом, и нижний, MAC (Medium Access Control), подуровень управления доступом к среде передачи. Подуровень LLC находится непосредственно над подуровнем MAC и является посредником между протоколами сетевого уровня и разнообразными методами доступа к среде передачи данных. В настоящее время еще не выработался однозначный подход к построению LLC. Дополнительные сведения, касающиеся существующих протоколов разных уровней, Ethertype, SNAP (Sub-network Access Protocol — протокол доступа к подсети) и технологии использования мостов можно найти в работе [Perlman, 1992].

Как и все стандарты серии 802.x, 802.11 определяет только физический уровень (PHY) и уровень доступа к среде MAC. В свою очередь физический уровень подразделяется на два подуровня — PLCP (Physical Layer Convergence Protocol — конвергентный протокол физического уровня), и PMD (Physical Medium Dependent — зависимый от физического носителя) (см. рис. 7.6). Основными функциями уровня MAC являются доступ к физической среде обмена, фрагментация и шифрование данных. В функции подуровня PLCP входит формирование сигнала обнаружения несущей. Эта функция называется CCA (Clear Channel Assessment — оценка занятости канала). В задачи подуровня входит организация точек доступа к службам (Service Access Point — SAP) для взаимодействия между элементами различных уровней независимо от

выбранной физической среды обмена. Подуровень PMD выполняет процедуры модуляции и кодирования/декодирования сигнала. Более подробно физический уровень организации сети изучается в следующих разделах.

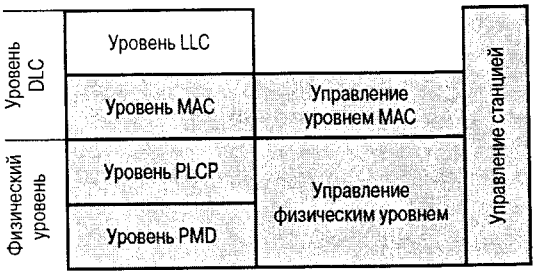

Рис. 7.6. Подробная протокольная архитектура и управление IEEE 802.11

В ведении уровня MAC находится функция доступа станций к BSS. На уровне MAC происходит управление роумингом — процессом перемещения станции от одной точки доступа к другой без потери связи. Кроме того, здесь происходит управление синхронизацией станций с точкой доступа, аутентификацией, шифрованием и энергосбережением. Управление уровнем MAC также включает поддержание **базы управляющей информации** (management information base — MIB) уровня MAC

Главными задачами **управления физическим уровнем** являются настройка каналов и поддержка базы MIB физического уровня. В задачи управления периферийных станций входит взаимодействие с сетью на физическом уровне и уровне MAC.

7.3.3. Физический уровень

На физическом уровне IEEE 802.11 предусматривает три способа беспроводной передачи данных, два из которых предполагают использование радиопередачи, а третий основан на инфракрасном излучении. Для радиочастотной передачи выбран безлицензионный диапазон ISM на частоте 2,4 ГГц. На физическом уровне формируется сигнал оценки занятости канала (clear channel assessment — CCA), который используется на уровне MAC механизмами управления доступа к среде и указывает, свободна ли среда в данный момент. Как в действительности получается данный сигнал — определяет технология передачи (которая рассматривается позже).

Кроме того, физический уровень предлагает точку доступа к службе (service access point — SAP), имеющую скорость передачи на уровень MAC 1 или 2 Мбит/с. Далее в разделе подробно описываются три версии физического уровня передачи данных, как они определены в стандарте IEEE 802.11.

7.3.3.1. Расширение спектра методом перескока частоты (frequency hopping spread spectrum — FHSS)

FHSS — это метод расширения спектра, позволяющий многим передатчикам работать в одной полосе радиоспектра, не мешая друг другу, чему способствуют разные последовательности перескоков частоты. Стандарт определяет 79 каналов перескока для Северной Америки и Европы и 23 — для Японии (каждый с шириной полосы 1 МГц в диапазоне ISM 2,4 ГГц). Порядок чередования каналов определяется псевдослучайной последовательностью. Прочие параметры передачи, такие как мощность передатчиков, устанавливаются каждой страной в отдельности. В США, например, мощность EIRP (equivalent isotropically radiated power — эквивалентная изотропно-излучаемая мощность) передатчика ограничена 1 Вт, в Европе — 100 мВт.

Для модуляции сигнала стандарт определяет гауссову частотную манипуляцию (Gaussian FSK — GFSK). Для скорости 1 Мбит/с стандарт задает использование двухуровневой GFSK (т.е. один бит отображается в одну частоту, см. главу 2). Для скорости 2 Мбит/с стандарт задает использование четырехуровневой GFSK (т.е. два бита отображаются в одну частоту). Скорость передачи 1 Мбит/с является обязательной, а 2 Мбит/с — возможной для более дорогих устройств с расширенными возможностями.

На рис. 7.7 представлен кадр физического уровня, используемый с методом FHSS. Кадр состоит из двух основных частей — части PLPC (предварительная последовательность битов и заголовок) и части полезной нагрузки. Первая служебная часть кадра всегда передается на скорости 1 Мбит/с, полезная нагрузка (т.е. данные MAC) может использовать скорость 1 или 2 Мбит/с. Помимо этого данные MAC скремблируются с использованием полинома $s(z) = z^7 + z^4 + 1$ для удаления из спектра постоянной составляющей и приближения его к спектру белого шума. Поля кадра выполняют следующие функции.

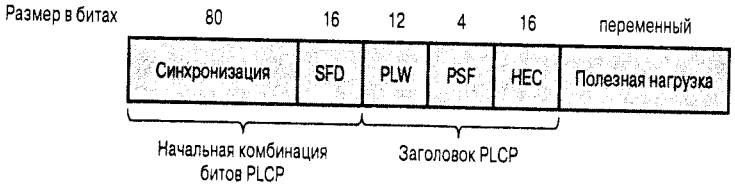

Рис. 7.7. Формат кадра физического уровня IEEE 802.11 при использовании схемы FHSS

- **Синхронизация**. Предварительная последовательность PLCP начинается с 80 бит синхронизации, представляющих собой последовательность 010101.... Данная последовательность используется для синхронизации потенциальных приемников и обнаружения сигнала CCA.
- **Стартовый разделитель кадра (start frame delimiter — SFD)**. Следующие 16 бит указывают начало кадра, таким образом обеспечивая кадровую синхронизацию. Последовательность SFD имеет вид 0000110010111101.
- **Длина слова PLCP_PDU (PLCP_PDU length word — PLW)**. Первое поле заголовка PLCP — это длина в байтах информационного блока, включая 32-битовый код CRC в конце полезной нагрузки. Поле может иметь размер от 0 до 4 095.
- **Сигнальное поле PLCP (PLCP signaling field — PSF)**. В настоящее время в этом четырехбитовом поле задан один бит, указывающий скорость передачи следующей за ним информации (1 или 2 Мбит/с).
- **Проверка на наличие ошибок в заголовке (Header Error Check — HEC)**. Контрольная сумма заголовка (поле длиной 16 бит), используемая для защиты заголовка PLCP. Для вычисления применяется стандартный полином ITU-T $G(x) = x^{16} + x^{12} + x^5 + 1$.

7.3.3.2. Расширение спектра методом прямой последовательности (direct sequence spread spectrum — DSSS)

DSSS — это альтернативный метод расширения спектра с использованием разделения по коду, а не по частоте. При передаче информации по методу DSSS в сети IEEE 802.11 расширение достигается с помощью последовательности (+1, −1, −1, +1, −1, +1, +1, +1, −1, −1, −1), называемой кодом Баркера (Barker code). Ключевой особенностью данного метода является устойчивость к помехам и нечувствительность к многолучевому распространению.

Стандарт IEEE 802.11 также устанавливает для схемы DSSS использование диапазона ISM на частоте 2,4 ГГц и предлагает скорость передачи 1 или 2 Мбит/с. Для передачи со скоростью 1

Мбит/с используется дифференциальная двоичная фазовая манипуляция (differential binary phase shift keying — DBPSK), а для передачи со скоростью 2 Мбит/с — дифференциальная квадратурная фазовая манипуляция (differential quadrature phase shift keying — DQPSK). Как и ранее, максимальная мощность EIRP составляет 1 Вт для США и 100 мВт для Европы. Скорость передачи символов 1 МГц дает скорость передачи элементарных сигналов 11 МГц. Для удаления из спектра постоянной составляющей и приближения его к спектру белого шума все данные, передаваемые физическим уровнем с использованием схемы DSSS, скремблируются с использованием полинома $s(z) = z^7 + z^4 + 1$.

На рис 7.8 изображен кадр физического уровня при использовании схемы DSSS. Кадр состоит из частей — части PLPC (предварительная комбинация битов и заголовок) и части полезной нагрузки. Для передачи PLPC всегда используется скорость 1 Мбит/с, полезная нагрузка (т.е. данные MAC) может передаваться на скорости 1 или 2 Мбит/с. Поля кадра имеют следующие функции.

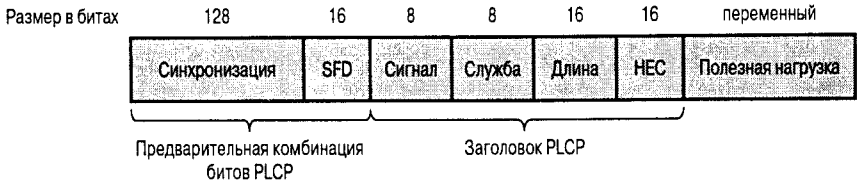

Рис. 7.8. Формат кадра физического уровня IEEE 802.11 при использовании схемы DSSS

- **Синхронизация**. Первые 128 бит используются не только для синхронизации, но и для регулировки усиления, обнаружения энергии (для CCA) и компенсации сдвига частоты.
- **Стартовый разделитель кадра (start frame delimiter — SFD)**. Данное 16-битовое поле используется для синхронизации в начале кадра и включает фиксированный код 1111001110100000.
- **Сигнал**. В настоящее время определены два значения для этого поля, указывающих скорость передачи полезной нагрузки. Значение 0x0A указывает на скорость 1 Мбит/с (а следовательно, модуляцию DBPSK), 0x14 указывает, что будет использована скорость 2 Мбит/с (и модуляция DQPSK). Другие значения зарезервированы для будущего использования, т.е. более высоких скоростей передачи данных.
- **Служба**. Данное поле зарезервировано для будущего использования, впрочем, значение 0x00 указывает, что передается пакет стандарта IEEE 802.11.
- **Длина**. Как и в других системах, 16 бит используется для указания размера поля данных.
- **Проверка на наличие ошибок в заголовке (Header Error Check — HEC)**. Поля сигнала, службы и длины защищены данной контрольной суммой, использующей стандартный полином ITU-T CRC-16.

7.3.3.3. Передача в инфракрасном диапазоне

Инфракрасная передача — способ беспроводной передачи в инфракрасном диапазоне с длиной волны 850–900 нм (около видимого светового спектра). Для организации ИК-канала не требуется оформление лицензии, но при применении лазерных диодов в качестве излучателей необходимо руководствоваться правилами безопасного их использования. Стандарт предусматривает два типа ИК-передачи — прямая передача, для которой приемник и передатчик находятся в зоне прямой видимости, и диффузная, основанная на отражении сигнала от различных поверхностей. Это позволит организовывать соединение точечной конфигурации с многоточеч-

ной. Для инфракрасного диапазона при отсутствии источника тепла и яркого солнечного освещения возможна скорость передачи до 10 Мбит/с. Обычно подобные сети используются только в зданиях, например, аудиториях, конференц-залах, и т. п. Повторное использование частоты достигается довольно простым способом — стены более чем достаточно для экранирования одной сети IEEE 802.11 с ИК-передачей от другой подобной. (См. также раздел 7.1, где сравниваются характеристики сетей ИК-диапазона и сетей, использующих радиопередачу, а также работу [Wesel, 1998].)

7.3.4. Уровень управления доступом к среде

Уровень управления доступом к среде (medium access control layer — MAC) должен выполнять несколько задач. Прежде всего это собственно контроль доступа к среде, но он также может предлагать поддержку роуминга, аутентификации и экономии мощности. Основной службой, предлагаемой уровнем MAC, является обязательная **служба асинхронной передачи** и возможная **служба срочных приложений**. Стандарт 802.11 предусматривает асинхронную передачу только в режиме эпизодических сетей, но обе указанные службы можно предложить, использовав сеть с инфраструктурой и доступ к среде, управляемый посредством точек доступа. Асинхронная служба поддерживает широковещание и многоадресную передачу, а также обмен пакетов по принципу "без обязательств", т.е. при передаче не устанавливается пределов задержки.

В стандарте IEEE 802.11 определены три основных механизма доступа: обязательный основной метод на основе разновидности CSMA/CA, альтернативный метод, позволяющий избежать проблемы изолированного терминала, и бесконфликтный метод опроса для срочных служб. Первые два метода иногда объединяются под общим названием **распределенная координационная функция** (Distributed Coordination Function — DCF), третий метод называется **точечной координационной функцией** (Point Coordination Function — PCF). Схема DCF предлагает только асинхронную службу, тогда как PCF предлагает и асинхронную, и срочную службу, но требует точки доступа для управления доступом к среде и предупреждения конфликтов. Механизмы MAC также называются **распределенным беспроводным управлением доступом к среде** (distributed foundation wireless medium access control — DFWMAC).

Для всех методов доступа важными являются несколько параметров контроля времени ожидания перед доступом к среде. На рис. 7.9 показаны три различных параметра, определяющих приоритеты доступа к среде. Как показано, среда передачи может быть либо свободна, либо занята обменом пакетами (что определяется с помощью CCA). Если среда занята, то это могут быть информационные кадры или другие кадры управления. В течение фазы состязания доступ к среде пытаются получить несколько узлов.

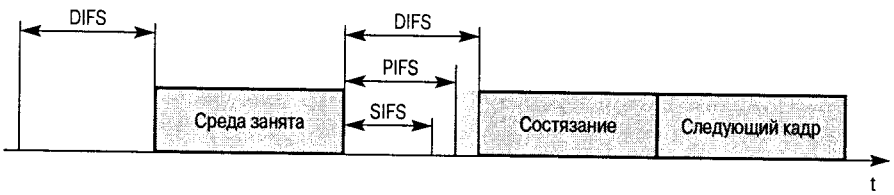

Рис. 7.9. Доступ к среде и межкадровые промежутки

- **Межкадровый промежуток DCF (DCF inter-frame spacing — DIFS)**. Указывает наибольшее время ожидания, а следовательно, самый низший приоритет для доступа к среде. Это время ожидания используется асинхронной службой в течение периода состязания (данный параметр и основной метод доступа подробно рассматриваются в разделе 7.3.4.1).

- **Межкадровый промежуток PCF (PCF inter-frame spacing — PIFS).** Время ожидания между DIFS и SIFS (следовательно — средний приоритет) используется для срочных служб, т.е. точка доступа, опрашивающая другие узлы, должна для доступа к среде выждать только время PIFS (подробнее см. раздел 7.3.4.3).
- **Короткий межкадровый промежуток (Short inter-frame spacing — SIFS).** Интервал ожидания доступа к среде, определенный для кратких управляющих сообщений, таких как подтверждения приема информационных пакетов или отклики на опрос. Это самый короткий интервал времени ожидания, поэтому станции, применяющие SIFS, имеют высший приоритет доступа. Подробное описание использования параметра SIFS представлено в разделах 7.3.4.1–7.3.4.3.

7.3.4.1. Основная служба DFWMAC-DCF с использованием CSMA/CA

Основным механизмом доступа в стандарте 802.11 является метод множественного доступа с обнаружением несущей и предотвращением конфликтов (carrier sense multiple access with collision avoidance — CSMA/CA), который использует случайный откат и изображен на рис. 7.10. Если обнаруживается, что среда свободна, по крайней мере, в течение промежутка DIFS (с помощью сигнала CCA физического уровня), узел может получать доступ к среде. Такой подход допускает краткую задержку доступа при небольшой загрузке среды. Но если доступ к среде пытаются получить все больше и больше станций, возникает потребность в дополнительных механизмах.

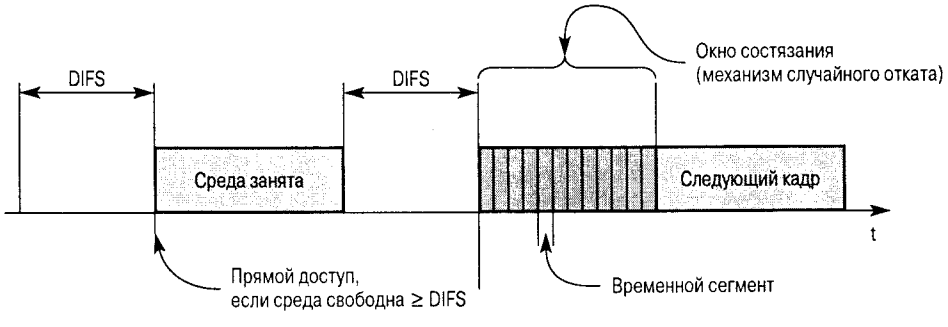

Рис. 7.10. Окно состязания и время ожидания

Если среда занята, узел должен ожидать, по крайней мере, время DIFS, после чего он вступает в фазу состязания. Каждый ожидающий узел теперь выбирает **случайное время отката** (в пределах **окна состязания**) и ожидает освобождения среды в течение этого промежутка. Если канал занят, узел должен пропустить данный цикл доступа и подождать следующей возможности, т.е. пока среда не будет свободна, по крайней мере, время DIFS. Дополнительное время ожидания каждого узла измеряется числом временны*х* слотов. Это время может исчисляться с учетом задержки распространения в среде, задержки передатчика и других физических параметров среды.

Ясно, что базовый механизм CSMA/CA доступа нельзя назвать справедливым. Независимо от общего времени ожидания в новом цикле состязания каждый узел имеет одинаковый шанс его выиграть, т. е. получить доступ к среде. Для обеспечения равноправия стандарт IEEE 802.11 предусматривает использование **таймера отката** (backoff timer). Как и ранее, узел выбирает случайный интервал времени ожидания в пределах окна состязания. Если станция не получила доступ к каналу в течение первого цикла, она останавливает таймер, ожидает, пока канал не освободится, и пытается получить доступ к среде. Это означает, что таймер отката уменьшается только тогда, когда канал свободен в течение, по крайней мере, времени DIFS, и замораживается, когда канал занят. При обнулении таймера станция может начинать передачу. Таким образом, станция, ожидающая доступ к каналу большее время, статистически получит его раньше, чем станция, включившая таймер отката после нее.

На рис. 7.11 объясняется механизм доступа стандарта IEEE 802.11 на примере пяти станций, пытающихся передать пакет в указанные моменты времени. У станции 3 первой появляется пакет данных, который необходимо передать. Она прослушивает канал, определяет, что он свободен, и по истечении промежутка DIFS получает доступ к среде, т.е. передает пакет. Станциям 1, 2 и 5 также требуется доступ к среде. Теперь они должны ждать, пока станция 3 закончит передачу, и канал будет оставаться незанятым в течение промежутка DIFS. Согласно алгоритму станции выбирают случайные интервалы в пределах окна состязания и включают таймеры отката.

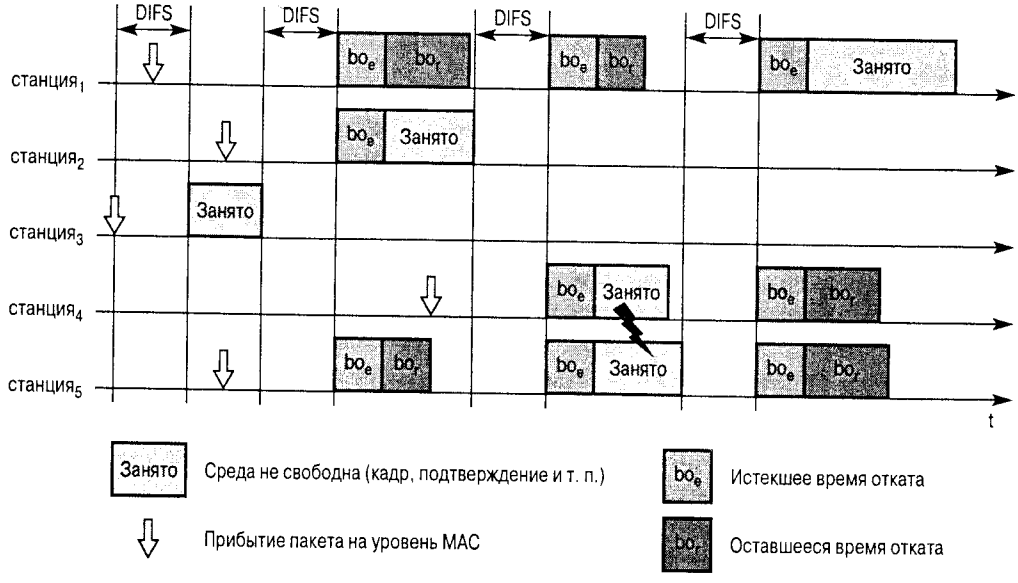

Рис. 7.11. Схема DFWMAC-DCF с несколькими состязающимися отправителями

На рис. 7.11 случайное время отката станции 1 показано как сумма времени bo_e (истекшего времени отката) и времени bo_r (остаточного времени отката). Аналогичным образом время отката определено для станции 5. В приведенном примере у станции 2 общее время отката равно всего bo_e, она первой получает доступ к среде. Таким образом, остаточное время отката для станции 2 не показано. Таймеры отката для станций 1 и 5 останавливаются, и станции сохраняют значения своих времен отката. Значит, когда новая станция включится в состязание за доступ к среде, станции 1 и 5 будут иметь статистическое преимущество, так как ожидать им необходимо только оставшееся время bo_r, тогда как новая станция должна случайным образом выбирать время отката из полного окна состязания.

Допустим, станции 4 также потребовался доступ к среде. Она тоже ожидает интервал DIFS и включает таймер отката. Теперь три станции ожидают, пока обнулятся их таймеры. Может случиться так, что таймеры двух станций обнулятся одновременно (например, у станций 4 и 5, как показано на рис. 7.11). Станции одновременно начинают передавать пакеты данных, и происходит конфликт. Теперь эти станции должны заново инициализировать таймеры и повторить процедуру доступа. В это время станция 1 снова останавливала таймер, но в последнем показанном цикле состязания она все-таки получает доступ к среде, пока станции 4 и 5 должны были ожидать, поскольку конфликт инициирует повторную передачу с заново выбранным временем отката.

Описанная схема доступа неэффективна при малой или значительной нагрузке. В зависимости от размера окна состязания (contention window — CW), в пределах которого выбираются интервалы отката, случайные значения могут либо быть слишком близкими друг к другу, что

приведет к большому числу конфликтов, либо слишком далекими, что приведет к ненужным задержкам. Таким образом, система пытается адаптироваться к текущему числу станций, пытающихся передать информацию.

Для окна состязания устанавливается первоначальный размер, например, $CW_{min} = 7$. Каждый раз, когда происходит конфликт (что свидетельствует о возрастающей нагрузке в среде доступа), длительность CW удваивается, пока не достигнет максимального значения, например, $CW_{max} = 255$ (CW может принимать значения 7, 15, 31, 63, 127 и 255). Чем больше окно состязания, тем больше разрешающая способность схемы с использованием случайных чисел, меньше вероятность выбрать одинаковые интервалы отката, а значит, меньше вероятность конфликта. В то же время, при малой нагрузке малое окно состязания обеспечивает меньшее время задержки доступа. Описанный алгоритм получил название **экспоненциального отката** (exponential backoff algorithm) и уже знаком пользователям подобной версии — CSMA/CD стандарта IEEE 802.3.

До сих пор рассматривался механизм доступа для широковещательного способа передачи. Для двухточечных соединений стандарт предусматривает дополнительные возможности. На рис. 7.12 изображена схема доступа с **подтверждением** получения пакета (acknowledgement — ACK), применяемая для двухточечной передачи. Станция отправляет пакет, принимающая станция проверяет контрольную сумму CRC полученного пакета и по истечении интервала времени SIFS отправляет кадр с подтверждением приема. Так как интервал SIFS самый короткий, то станция, применяющая его, имеет высший приоритет доступа к среде, и кадр ACK передается в первую очередь. Метод доступа с подтверждением приема обеспечивает контроль передачи пакетов, что имеет большое значение для беспроводной среды, подверженной помехам. Если отправитель не получил подтверждения, он будет повторно передавать пакет, пока не получит подтверждения, либо после определенного числа попыток прекратит передачу, при этом повторные попытки имеют обычный приоритет.

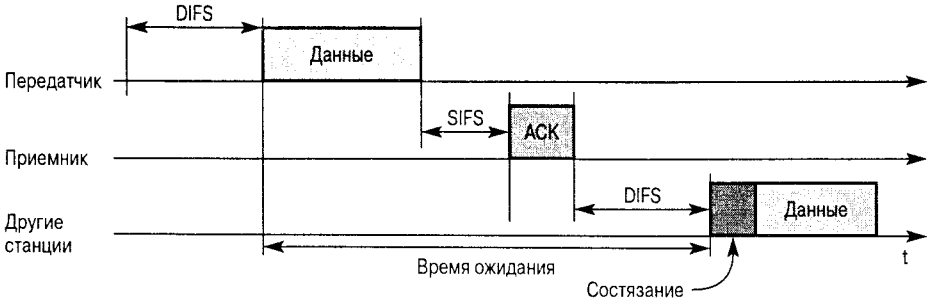

Рис. 7.12. Двухточечная передача IEEE 802.11

7.3.4.2. Схема DFWMAC-DCF с использованием пакетов RTS/CTS

В разделе 3.1 обсуждалась возможность появления скрытых узлов, что также возможно в сетях IEEE 802.11. Проблема возникает, если одна станция может принимать сообщения от двух других, которые не сообщаются между собой. В таком случае данные две станции могут обнаружить среду свободной, передать свои кадры и привести к конфликту на находящемся между ними приемнике. Для предотвращения этого стандарт определяет дополнительный механизм с использованием двух контрольных пакетов, RTS и CTS. Использование данного механизма не является обязательным, однако, на каждом узле 802.11 эта функция должна быть реализована для надлежащего приема контрольных пакетов RTS/CTS.

На рис. 7.13 иллюстрируется механизм доступа с применением RTS/CTS. После ожидания в течение времени DIFS (плюс случайное время отката, если требуется) отправитель посылает **запрос на передачу** (request to send — RTS). Приоритет пакета RTS не превышает приоритет дру-

гих информационных пакетов и содержит адрес получателя и полное время передачи, которое определяется как время, требуемое на передачу информационного кадра и получения подтверждения его приема. Все станции, получающие RTS, теперь должны установить собственные векторы распределения сети (Net Allocation Vector — NAV) согласно полю длительности. Вектор NAV задает наименьший интервал времени, который должен пройти, прежде чем станция снова попытается получить доступ к среде.

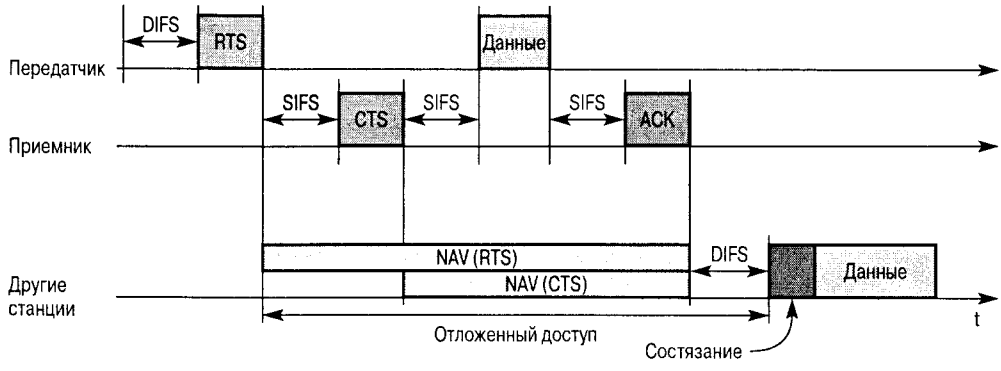

Рис. 7.13. Возможности, предлагаемые скрытым узлам при доступе без состязания

Принимающая станция по истечении SIFS отправляет кадр CTS (clear to send — "свободна для передачи"). Данный пакет CTS также содержит поле длительности, и все станции, получившие CTS, соответствующим образом изменяют свои значения NAV. (Данное множество приемников может не совпадать с множеством, получившим ранее пакет RTS.) Теперь все узлы, находящиеся в радиусе действия передатчиков отправителя и приемника пакетов RTS/CTS, информированы о том, что среда будет занята определенное время, и должны подождать перед попыткой доступа. Такой механизм резервирования среды исключительно для одного отправителя иногда называется схемой виртуального резервирования.

Получив разрешение на передачу, отправитель (по прошествии времени SIFS) посылает пакет данных. Адресат, проверив CRC переданного пакета, по прошествии времени SIFS отсылает отправителю подтверждение приема (пакет ACK). Теперь передача завершена, и значения NAV всех станций указывают, что среда свободна и стандартный цикл может начаться снова.

Рассмотренный сценарий резервирования канала решает проблему скрытого узла. Конфликт возможен только вначале, при передаче кадра RTS. Две или более станций могут начать передачу одновременно (пакеты RTS или информационные кадры). Использование механизма RTS/CTS может привести к значительным служебным издержкам, увеличивающим неэффективное использование пропускной способности и задержку доступа. Следовательно, нужно пороговое значение RTS, определяющее, когда использовать дополнительный механизм (обычно при *больших* пакетах), а когда — отключать (короткие кадры). В работах [Chhaya, 1996] и [Chhaya, 1997] представлен хороший обзор асинхронных служб IEEE 802.11, в частности, производительность при различной загрузке сети.

Частота появления ошибочных битов (Bit Error Rate — BER) в беспроводных системах передачи на несколько порядков выше, чем, например, для оптоволоконных каналов. Таким образом, для беспроводных каналов связи вероятность кадра с ошибкой выше, предполагая равный размер кадров. При уменьшении размера пакета коэффициент битовой ошибки не уменьшается, но количество правильно переданных пакетов возрастает, т.е. уменьшается частота появления ошибочных кадров (frame error rate — FER).

Стандарт 802.11 предусматривает единый механизм фрагментации — разбиения пакетов на более мелкие фрагменты, передающиеся по отдельности. Более того, на уровне MAC существует возможность изменения размеров пакета в зависимости от коэффициента ошибок в среде передачи. Таким образом, стандарт IEEE 802.11 задает режим **фрагментации** (см. рис. 7.14). Как и в рассмотренной выше схеме, станция-отправитель резервирует канал, отсылая пакет RTS, в котором указана продолжительность передачи первого фрагмента и соответствующего ему подтверждения ACK. Узлы, принявшие RTS, устанавливают время ожидания в NAV. Адресат пакета RTS дает разрешение на передачу, отсылая кадр CTS, также содержащий длительность передачи первого фрагмента, включая время на подтверждение. Узлы (возможно уже иные), принявшие CTS, корректируют время ожидания в NAV.

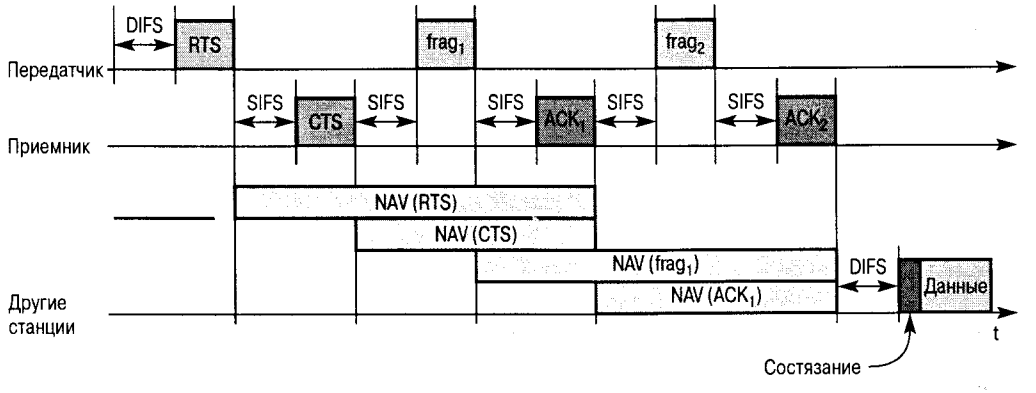

Рис. 7.14. Фрагментация пользовательских данных в IEEE 802.11

Как показано на рис. 7.14, отправитель может передать первый фрагмент данных $frag_1$. Отличием фрагментарной передачи пакета является изменение времени объявленной длительности на время занятия канала следующим фрагментом. Эта информация передается в текущем фрагменте, $frag_1$. Возможна ситуация, когда не все узлы, получившие пакет RTS с первоначально зарезервированным временем, получат пакеты $frag_1$ с изменениями и отобразят их в своих NAV. Это может произойти либо вследствие изменения параметров среды, либо из-за перемещения узлов сети.

Адресат, получив фрагмент пакета $frag_1$, по истечении времени SIFS отсылает станции-отправителю подтверждение ACK_1, содержащее время, необходимое для передачи следующего фрагмента. Группа станций, "услышавших" ACK_1, четвертая по счету в данном примере, корректирует свои NAV. В приведенном примере $frag_2$ является последним фрагментом пакета, поэтому он больше не резервирует время, так же как и соответствующий ему кадр ACK_2. Следовательно, после промежутка DIFS среда свободна, и узлы могут начать следующий цикл состязания.

7.3.4.3. DFWMAC-PCF с опросом узлов

Рассмотренные методы доступа не всегда обеспечивают оптимальные параметры передачи — минимальную задержку ожидания доступа и максимальную пропускную способность канала. Для срочных приложений стандарт задает использование поверх стандартного механизма DCF **точечной координационной функции** (point coordination function — PCF). В противоположность распределенному управлению, в методе PCF для доступа к каналу передачи требуется точка доступа, которая производит опрос отдельных узлов. Применение PCF для эпизодических сетей невозможно, поэтому эти сети не обеспечивают качество обслуживания, а предлагают работу по принципу "без обязательств" в беспроводных локальных сетях IEEE 802.11.

Точечный координатор в точке доступа разбивает время доступа на последовательные промежутки, называемые суперкадрами (см. рис. 7.15). Каждый **суперкадр**, в свою очередь, разделяется на два интервала — **период без состязания** и **период состязания**. В течение последнего периода могут использоваться два описанных выше метода доступа. На рисунке также изображены несколько беспроводных станций (использующих один канал) и векторы NAV этих станций.

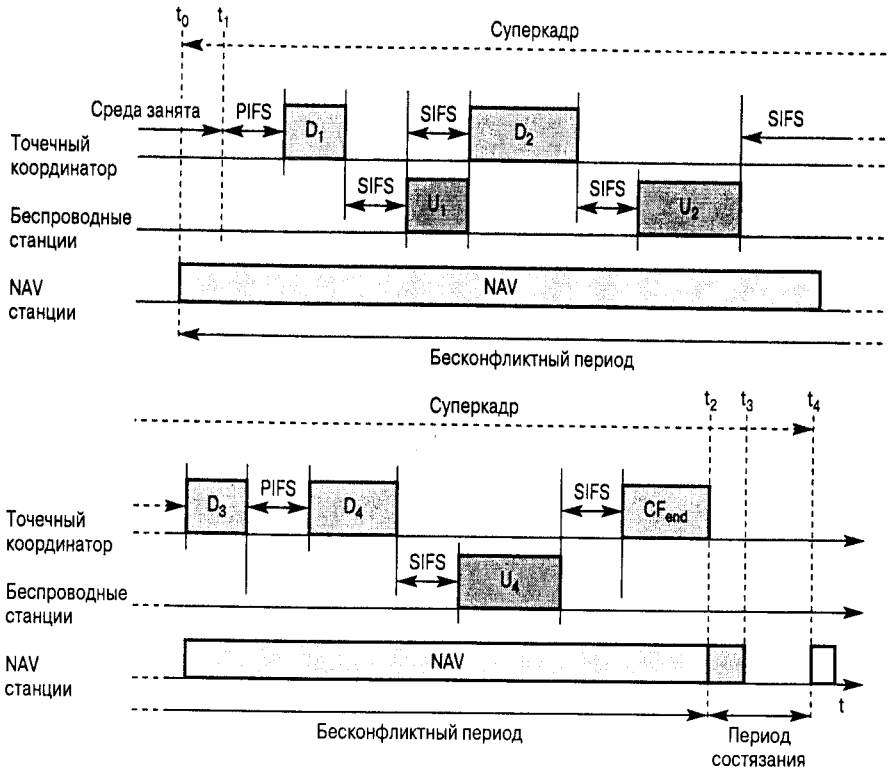

Рис. 7.15. Доступ без состязания с использованием механизмов опроса (PCF)

В момент времени t_0 теоретически должен был начаться период без состязания, но среда оказалась занятой, и начало передачи отложено. После того, как среда освободилась (время t_1), точечный координатор должен выждать время PIFS перед доступом к среде. Поскольку промежуток PIFS меньше DIFS, ни одна станция не начнет передавать раньше.

Теперь точечный координатор отправляет пакет данных D_1 ближайшей беспроводной станции. Через интервал SIFS узел может отправлять ответ (см. рис. 7.15). Снова, выждав время SIFS, точечный координатор может опросить вторую станцию, передав пакет D_2. Вторая станция отправляет данные U_2. Далее производится опрос третьего узла, который не может ответить данными, поэтому точечный координатор не получит пакета по прошествии времени SIFS.

После ожидания в течение времени PIFS координатор может возобновить опрос станций. Наконец, точечный координатор может выпустить маркер конца (end marker) CF_{end}, указывающий, что может возобновиться период состязания. Использование PCF автоматически настраивает значения NAV, запрещая передачу других станций. В примере на рис. 7.15 изначально планировался период без состязания от t_0 до t_3, но координатор закончил опрос раньше, в момент t_2. В момент t_4 начинается очередной суперкадр.

Эффективность работы беспроводной сети при использовании PCF полностью определяется возможностями опроса точки доступа. Обычно время периода без состязаний равномерно распределяется между всеми узлами сети, таким образом обеспечивая равноправный доступ. Эта схема подобна множественному доступу с временным разделением (time division multiple access — TDMA) в системах передачи данных. Для методов временного разделения характерны служебные издержки, если узлы не имеют информации, готовой к передаче, а точка доступа постоянно их опрашивает. В работе [Anastasi, 1998] приводится пример голосовой передачи с использованием 48-байтных пакетов. В этом случае использование PCF вводит служебные издержки размером 75 байт.

7.3.4.4. Кадры уровня MAC

На рис. 7.16 изображен пакет уровня MAC в стандарте IEEE 802.11. Структура и назначение полей пакета приводятся ниже.

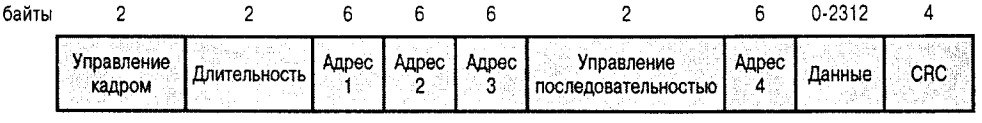

Рис. 7.16. Структура пакета IEEE 802.11

- **Управление кадром**. Поле длиной 2 байта состоит из нескольких подполей, несущих информацию о версии протокола, типе кадра (управляющий, проверочный, данные), о фрагментации передаваемого пакета, информации о конфиденциальности и 2-битового поля системы распределения (distribution system — DS), указывающего значение четырех адресных полей кадра.
- **Идентификатор длительности**. Поле используется для схемы виртуального резервирования канала с применением RTS/CTS и содержит значение, указывающее период планируемой занятия среды.
- **Адресные поля**. Четыре адресных поля длиной 48 бит каждое. Имеют то же назначение, что и адресные поля MAC уровня в других стандартах серии IEEE 802.x. Подробное описание дается в отдельном разделе.
- **Управление очередностью**. Может произойти так, что кадр будет продублирован (вследствие использования механизма подтверждения приема). Таким образом, данное поле предназначено для выявления и отсева дублирующихся фрагментов.
- **Поле данных**. Кадр MAC может содержать произвольные данные (до 2312 байт), передаваемые от отправителя получателю (получателям).
- **Контрольная сумма (CRC)**. 32-битовый код CRC для проверки безошибочной передачи пакета (обычная процедура для всех сетей IEEE 802.x).

Кадры уровня MAC могут передаваться между мобильными станциями, между мобильными станциями и между точкой доступа и точками доступа (см. рис. 7.3). Эти случаи различаются двумя битами в поле Frame Control, "**к DS**" и "**от DS**", таким образом, осуществляется управление значениями четырех адресных полей. В табл. 7.1 представлены четыре возможных значения битов DS и соответствующая интерпретация четырех адресных полей.

Каждая станция, точка доступа или беспроводный узел используют **адрес 1**, который идентифицирует физического получателя (получателей). Следующий **адрес 2** представляет физического отправителя кадра. Данная информация важна, поскольку данный конкретный отправитель также является получателем подтверждения уровня MAC. Два оставшихся адресных поля

используются для логического присвоения кадров (логический отправитель, идентификатор BSS, логический получатель).

Таблица 7.1. Интерпретация MAC-адресов в кадре MAC 802.11

К DS	От DS	Адрес 1	Адрес 2	Адрес 3	Адрес 4
0	0	DA	SA	BSSID	—
0	1	DA	BSSID	SA	—
1	0	BSSID	SA	DA	—
1	1	RA	TA	DA	SA

Для адресации возможны четыре сценария (см. табл. 7.1).

- **Эпизодическая сеть**. Если оба бита DS равны 0, кадр MAC включает пакет, передаваемый между двумя беспроводными узлами без участия в обмене системы распределения. В этом случае DA обозначает **адрес назначения пакета** (destination address), SA — **адрес отправителя** (source address) кадра, что соответствует физическим адресам приемника и передатчика. Третий адрес идентифицирует **базовый набор служб (basic service set — BBSID)**. Четвертый адрес не используется.
- **Инфраструктурная сеть (от AP)**. Если задан только бит "от DS", кадр происходит от точки доступа. DA — это физический и логический приемник, второй адрес определяет BSS, третий адрес — логического отправителя, адрес источника кадра MAC. Данный случай представляет передачу посредством точки доступа.
- **Инфраструктурная сеть (к AP)**. Если станция отправляет пакет другой станции посредством точки доступа, устанавливается только бит "к DS". В этом случае первый адрес представляет физический приемник кадра, точку доступа. Второй адрес — это физический и логический отправитель кадра, третий — логический получатель.
- **Инфраструктурная сеть (в пределах DS)**. Последний случай — пакеты, пересылаемые точками доступа друг другу в пределах системы распределения. Устанавливаются оба бита. Первый, **адрес приемника** (receiver address — RA), представляет MAC-адрес принимающей точки доступа. Подобным образом второй адрес, **адрес передатчика** (transmitter address — TA), определяет отправляющую точку доступа в пределах системы распределения. Теперь требуется еще два адреса — для определения исходного назначения кадра DA и исходного источника SA. Без этих дополнительных адресов для передачи кадров MAC потребуется некоторый механизм инкапсуляции.

7.3.5. Управление MAC

Управление MAC играет центральную роль на станции IEEE 802.11, поскольку оно в большей или меньшей степени контролирует все функции, связанные с интеграцией системы, т.е. функции интеграции беспроводной станции в BSS, формирования ESS, синхронизации станции и т.д. Определены следующие функциональные группы (подробнее рассматриваются в следующих разделах).

- **Синхронизация**. К ней относится функция поддержки поиска беспроводной локальной сети, синхронизации локальных таймеров узлов и генерации сигналов-маяков.
- **Управление питанием**. Функции, контролирующие активность передатчика с целью энергосбережения, например, использование буферизации пакетов, периодическое переключение "спящего" и рабочего режимов без потери кадра.

- **Роуминг**. Функции для присоединения к сети (ассоциации), изменения точек доступа и поиска точек доступа.
- **База управляющей информации (management information base — MIB)**. Все параметры, представляющие текущее состояние беспроводной станции и точки доступа, хранятся в базе MIB для внутреннего и внешнего доступа. Доступ к базе возможен через стандартные протоколы, например, SNMP (Simple network management protocol — простой протокол сетевого управления).

7.3.5.1. Синхронизация

Все узлы сети оборудованы внутренними таймерами, поскольку для выполнения функций энергосбережения, синхронизации при перескоке частот в FHSS, координации опроса в PCF узлы должны быть синхронизированы друг с другом либо с точкой доступа. Для управления процессами обмена в сети IEEE 802.11 задана **синхронизирующая функция** (timing synchronization function — TSF). При использовании PCF каждый узел, руководствуясь сигналами внутреннего таймера, может предсказать начало суперкадра, т.е. периоды без состязания и периоды состязания. При использовании схемы FHSS физическим уровням требуются одинаковые последовательности перескоков для всех узлов, чтобы все узлы могли сообщаться с BSS.

В пределах базовой системы BSS с инфраструктурой синхронизация производится следующим образом. Точка доступа (квази)периодически отправляет **сигнальные кадры** (beacon frame). Эти кадры включают временные метки (timestamp) и другую информацию о настройках точки доступа (например, идентификация BSS), необходимые для энергосбережения, роуминга и т. д. Принимающие станции используют временные метки для настройки своих таймеров. В силу того, что среда не всегда свободна, передача сигналов квазипериодична. Для поддержания синхронизации в сети узлам достаточно время от времени сверять свои внутренние таймеры и перестраиваться в случае необходимости.

На рис. 7.17 представлена простейшая схема синхронизации для сетей с инфраструктурой. Временные метки должны передаваться через определенные временные интервалы (beacon interval), устанавливаемые точкой доступа. Поскольку своевременная передача возможна не всегда (из-за занятости среды), сигнальные интервалы не могут быть ориентиром для узлов сети. В то же время точки доступа всегда пытаются согласовать передачи с ожидаемыми сигнальными интервалами (**целевое сигнальное время передачи**), т.е. сигнальные интервалы не смещаются при пропуске одного из них. Реальными временными точками отсчета являются временные метки, передаваемые в сигнальных кадрах, но не время по графику.

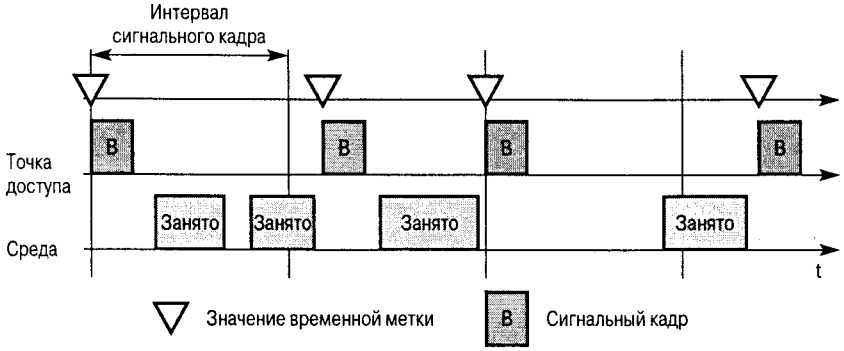

Рис. 7.17. Сигнальная передача в занятой инфраструктурной сети IEEE 802.11

В эпизодических сетях ситуация несколько отличается, поскольку здесь нет точки доступа, с которой могли бы синхронизироваться все узлы сети. В данном случае каждый узел поддер-

живает собственный таймер синхронизации и начинает передачу сигнального кадра по прошествии сигнального интервала. На рис. 7.18 приведен пример организации передачи сигнальных кадров несколькими станциями. При этом к сигнальным кадрам применяется стандартный алгоритм случайного откат, следовательно, "выигравшим" постоянно является один такой кадр. Теперь все станции настраивают свои локальные таймеры, используя принятый сигнальный кадр и подавляя передачу своих сигнальных кадров для данного цикла. При возникновении конфликта сигнальный кадр теряется. В данном сценарии интервалы передачи сигнальных кадров могут несколько смещаться во времени, поскольку все локальные таймеры могут по-разному отсчитывать время, следовательно, время начала интервала передачи сигнального кадра может меняться с точки зрения узла. Впрочем, после синхронизации все узлы снова обладают согласованным представлением о системе.

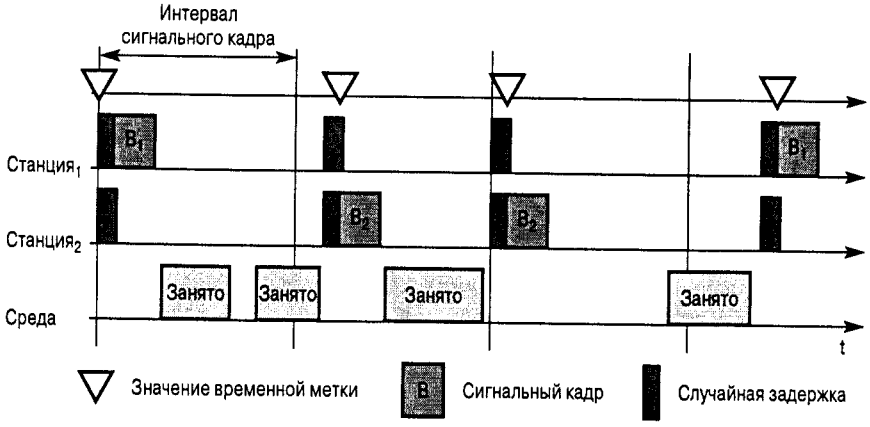

Рис. 7.18. Сигнальная передача в занятой эпизодической сети IEEE 802.11

7.3.5.2. Управление питанием

Беспроводные устройства питаются от батарей (обычных или солнечных). Следовательно, для успешного использования подобных устройств нужно обращать внимание на механизмы энергосбережения. Стандартные протоколы локальных сетей подразумевают, что станции всегда готовы к приему данных, хотя большую часть времени в слабо загруженных сетях приемники работают вхолостую. В то же время постоянная готовность приемопередающих устройств требует повышенного энергопотребления, уменьшает срок службы батарей мобильных узлов, поскольку ток приемника должен достигать 100 мА [Woesner, 1998].

Основной принцип управления питанием в сетях IEEE 802.11 заключается в отключении трансиверов, когда в них нет необходимости. Поскольку при проектировании схемы управления питанием нельзя наверняка предсказать, когда трансивер должен активизироваться для приема/передачи пакета, должно предусматриваться периодическое "пробуждение" трансивера. Отключение трансивера должно быть незаметным для существующих протоколов и достаточно гибким для поддержки различных приложений. В то же время срок службы батареи можно увеличить, пожертвовав пропускной способностью. Длительные периоды отключения экономят время "жизни" батареи, но снижают среднюю пропускную способность и наоборот.

Реализация основной идеи энергосбережения включает использование для станций двух состояний — **состояния ожидания** (sleep state) и **активного состояния** (awake state) — и буферизацию данных отправителя. При попытке отправителя связаться со станцией, реализующей схему энергосбережения, он должен буферизировать данные, если станция находится в состоянии

ожидания. Ожидающие станции, с другой стороны, должны периодически на определенное время переходить в активное состояние. В течение этих периодов все отправители могут объявлять адресатов буферизированных информационных кадров. Если станция обнаруживает, что входит в список адресатов, она должна оставаться в активном состоянии до завершения передачи. Выход из состояния ожидания в нужный момент времени требует **синхронизирующей функции** (timing synchronization function — TSF), введенной в разделе 7.3.5.1. Периоды активных режимов должны совпадать для всех станций.

Управление питанием в **инфраструктурных** сетях происходит значительно проще, чем в эпизодических. Точка доступа постоянно отслеживает информацию, приходящую для узлов, работающих в энергосберегающем режиме, и буферизирует пакеты, адресованные станциям. В каждом сигнальном пакете, отправляемом точкой доступа, содержится **карта индикации транспортного потока** (traffic indication map — TIM). Карта TIM включает список станций, для которых в точке доступа занесены в буфер информационные кадры.

Функция TSF обеспечивает своевременный переход узлов в рабочий режим для приема сигнальных кадров и карт TIM. Если карта TIM указывает, что одноадресный кадр буферизован для данной станции, станция переходит в активный режим. При многоадресной/широковещательной передаче станции всегда должны находиться в активном режиме. В рабочий режим переходят также станции, которым необходимо отправить информацию точке доступа. Следует также отметить, что таймеры TSF не выключаются при переходе станции в состояние ожидания.

На рис 7.19 изображена одна точка доступа и одна станция с указанием состояния среды. Как и ранее, через определенный интервал точка доступа передает сигнальный кадр. Теперь данный интервал равен интервалу TIM. Кроме того, точка доступа поддерживает **карту индикации целевого транспортного потока** (delivery traffic indication map — DTIM), используемую для многоадресной/широковещательной передачи. Интервал DTIM всегда кратен интервалу TIM.

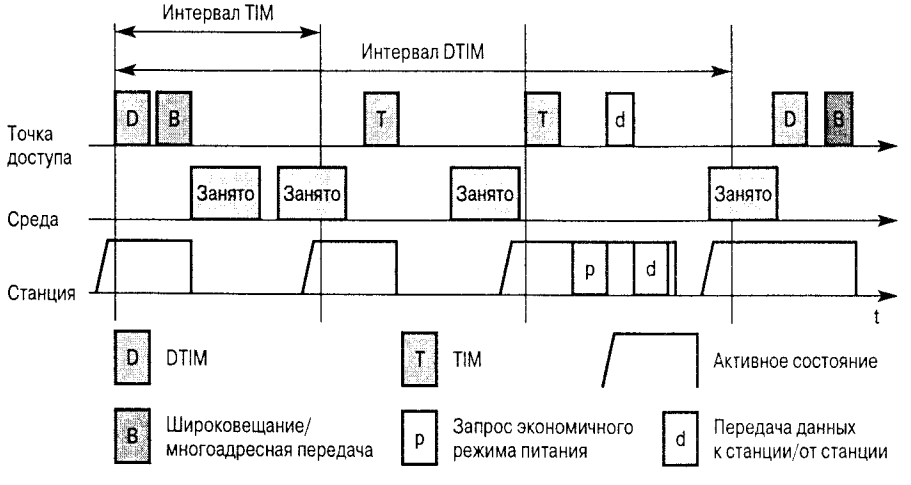

Рис. 7.19. Управление питанием в инфраструктурных сетях IEEE 802.11

Все станции (в данном примере показана только одна) входят в активный режим до момента трансляции таблицы TIM или DTIM. В первом случае точка доступа должна передать широковещательный кадр; таким образом, станция находится в активном режиме для приема этого кадра. После получения широковещательного кадра станции возвращаются в режим ожидания. В активный режим станция перейдет только перед приемом очередного кадра TIM. В приведенном примере кадр TIM задержан вследствие занятости среды, следовательно, станция оста-

ется в активном режиме. Точка доступа не имеет готовой к передаче информации, и станция возвращается в режим ожидания.

В течение следующего интервала TIM точка доступа обнаруживает, что станция является адресатом буферизованных кадров. Теперь станция отвечает **запросом PS** (запрос энергосберегающего режима) и остается активной для приема данных. Затем точка доступа передает данные станции, станция подтверждает прием и может в ответ передать некоторую информацию (как показано в примере), подтверждение приема, которое будет отправлено точкой доступа (не показано на рисунке). Впоследствии станция снова переходит в режим ожидания.

В заключение точка доступа желает передать некоторую широковещательную информацию в течение следующего интервала DTIM, что снова откладывается из-за занятости среды. В зависимости от внутреннего порога станция может находиться в активном режиме, если ожидается слишком короткий период ожидания. Данный механизм ярко показывает компромисс между короткими задержками доступа станции и экономией батареи. Чем короче интервал TIM, тем короче время ожидания доступа и тем меньше эффект энергосбережения.

Для эпизодических сетей управление питанием значительно сложнее. В данном случае нет точки доступа, занимающейся буферизацией данных, но если станции желают сообщаться со станцией, работающей в экономичном режиме, они должны обладать возможностями буферизации. Теперь все станции объявляют список буферизованных кадров в течение периода активности всех станций. Адресаты кадров объявляются с использованием **карты индикации эпизодического адресного потока** (ad hoc traffic indication map — ATIM) — период объявления называется **окном ATIM**.

На рис. 7.20 приведена простая эпизодическая сеть, состоящая из двух станций. Как и ранее, сигнальные интервалы определяются распределенной функцией (сигнальные кадры могут посылать различные станции). В то же время вследствие данной синхронизации все станции эпизодической сети активизируются одновременно. Все станции остаются активными в течение интервала ATIM, как показано для первых двух этапов, и переходят в режим ожидания, если для них нет кадров в буферах. На третьем этапе станция 1 содержит данные, буферизированные для станции 2 (указывается передачей ATIM). Станция 2 подтверждает данный кадр ATIM и остается активной в течение передачи. После завершения окна ATIM станция 1 может передавать информационный кадр и станция 2 подтвердит его прием. В этом случае станции активны до начала следующего сигнального интервала.

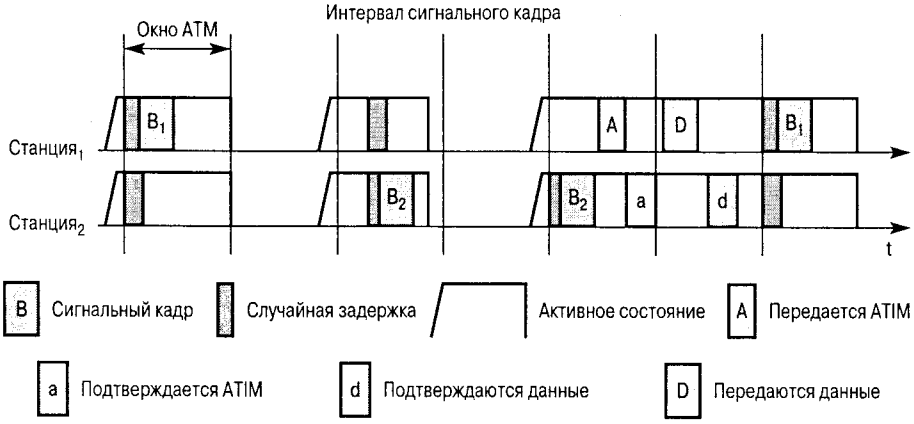

Рис. 7.20. Управление питанием в эпизодических сетях IEEE 802.11

Одной из проблем данного подхода является масштабируемость. Если в пределах эпизодической сети в режиме энергосбережения работает много станций, осуществить передачу в пределах окна ATIM могут пожелать также большое число станций. Чем больше передач ATIM,

тем больше возможностей для конфликта, и тем больше станций будут вынуждены отложить свою передачу.

7.3.5.3. Роуминг

Типичные беспроводные сети в пределах здания требуют более одной точки доступа для охвата всех помещений. В зависимости от характеристик стен диапазон передачи одной точки доступа составляет порядка 10-20 м, если желательна передача приемлемого качества. Если пользователи (а с ними и беспроводная станция) перемещаются, станция должна переключаться с одной точки доступа на другую с целью обеспечения непрерывного обслуживания. Перемещение между точками доступа получило название роуминг (roaming).

Перечислим этапы роуминга между двумя точками доступа.

- Если по каким-либо причинам произошло ослабление сигнала и ухудшилось качество связи с точкой доступа AP_1, станция начинает **сканирование** сети в поисках другой точки доступа.
- Сканирование включает активный поиск другой ячейки BSS и может использоваться для настройки новой BSS в эпизодических сетях. Стандарт IEEE 802.11 задает сканирование одного или нескольких каналов (если это допускает физический уровень) и различает пассивное и активное сканирование. **Пассивное сканирование** означает просто прослушивание среды с целью поиска новых сетей, т.е. получение сигнальных кадров, выпущенных другими сетями (за это отвечает синхронизирующая функция точки доступа). При **активном сканировании** станция определяет местоположение точки доступа, передавая в каждый канал пробный импульс и ожидая ответа на него от точки доступа. Сигнальный кадр и отклик на пробный импульс содержат всю информацию, требуемую для вхождения в новую ячейку BSS.
- Находясь в зоне действия нескольких точек доступа, станция на основе, например, мощности сигнала выбирает одну из них, AP_2 и передает запрос на ассоциацию к этой выбранной точке.
- Новая точка доступа AP_2 отвечает **откликом ассоциации**. Если отклик удовлетворителен, станция переключается на новую точку доступа AP_2. В противном случае станция продолжает сканирование и поиск новой точки доступа.
- Точка доступа, принявшая запрос на ассоциацию, указывает системе распределения (distribution system — DS) новую станцию в своей BSS. Затем система распределения обновляет свою базу данных, содержащую информацию о текущем положении беспроводных станций. Эта база данных необходима для передачи кадров между различными BSS, т.е. между различными точками доступа, контролирующими BSS, входящие в расширенную систему ESS (см. рис. 7.3). Кроме того, система распределения может информировать AP_1 о том, что станция больше не входит в контролируемую ею ячейку сети.

7.3.6. Дальнейшее развитие

По мере роста продуктов стандарта IEEE 802.11 в комитете IEEE появлялись новые группы, темами обсуждения в которых стало улучшение стандарта и новые приложения.

Один из таких проектов (группа **802.11a**) связан с беспроводными локальными сетями IEEE 802.11, т.е. с совместимыми уровнями MAC, но использующими полосу частот 5 ГГц. Данная высокая частота позволяет получить скорость передачи до 20 Мбит/с. Эта группа тесно сотрудничает с группой широковещательных сетей радиодоступа (broadband radio access network — BRAN) института ETSI. Целью такого сотрудничества является согласование общего физического уровня. Группа BRAN уже специфицировала беспроводную локальную сеть со скоростью 23,5 Мбит/с в диапазоне 5 ГГц (см. раздел 7.4).

Участники группы **802.11b** заняты предоставлением более высоких скоростей передачи данных на частоте 2,4 ГГц. Как обсуждалось выше, структура заголовков PLCP уже позволяет

поддерживать различные скорости передачи данных — специальное поле указывает скорость следующей далее полезной нагрузки. Исследовательская группа в рамках этого проекта рассматривает полностью совместимые пакеты с заголовками на уже знакомой скорости 1 Мбит/с, но с полезной нагрузкой на более высоких скоростях — 3 Мбит/с при использовании схемы FHSS и 11 Мбит/с при использовании схемы DSSS. Следует отметить, что единичные частные продукты уже поддерживают скорость 10 Мбит/с на частоте 2,4 ГГц при использовании специальных методов кодирования и высокопроизводительных устройств обработки сигналов.

Последняя исследовательская группа беспроводных частных сетей (wireless personal area networks — WPAN), входящая в комитет IEEE 802.11, занимается рассмотрением беспроводных сетей с точки зрения следующих пяти критериев.

- **Рыночный потенциал**. Сколько прикладных систем, устройств, поставщиков, заказчиков доступны при использовании данной технологии.
- **Совместимость**. Одно из необходимых требований — совместимость с IEEE 802.
- **Своеобразие**. Исследовательская группа не желает установить второй стандарт IEEE 802.11. В то же время в стандарте 802.11 не рассмотрены многие темы (такие как низкая стоимость, малая мощность или форм-фактор).
- **Техническая осуществимость**. Для дальнейшего обсуждения требуются прототипы, так что исследовательская группа не будет полагаться только на "бумажную работу".
- **Экономическая осуществимость**. Все, что разрабатывается данной группой, должно быть дешевле других решений и допускать масштабное промышленное производство.

В разделе 7.5 рассмотрена технология, согласующаяся с предложенным выше определением: Bluetooth, возникшая в результате сотрудничества исследовательской группы WPAN и консорциума Bluetooth.

7.4. HIPERLAN

В 1996 году институт ETSI стандартизовал HIPERLAN 1 (High Performance Local Area Network — высокопроизводительная ЛВС), локальную сеть, разрешающую мобильность и поддерживающую эпизодические и инфраструктурные топологии [ETSI, 1996]. Проект охватывает четыре спецификации — HIPERLAN 1, 2, 3, 4, поскольку было решено стандартизовать разные сети для разных целей. Ключевым отличием сетей HIPERLAN является интеграция срочных служб передачи данных. В табл. 7.2 представлено сравнительное описание спецификаций, которые в настоящее время называются HIPERLAN 2, HIPERACCESS и HIPERLINK. Три последние стандарта, а также беспроводная технология ATM и сети BRAN обсуждаются в главе 8.

Таблица 7.2. Исходные спецификации семейства беспроводных сетей HIPERLAN

	HIPERLAN (тип 1)	HIPERLAN (тип 2)	HIPERLAN (тип 3)	HIPERLAN (тип 4)
Применение	Беспроводная ЛВС	Доступ к стационарным сетям ATM	Беспроводные местные линии связи	Двухточечные беспроводные соединения ATM
Частота	5,1–5,3 ГГц		17,2–17,3 ГГц	
Топология	Децентрализованная эпизодическая/инфраструктурная	Сотовая, централизованная	Многоточечная	Двухточечная
Антенна	Ненаправленная		Направленная	

Окончание табл. 7.2

	HIPERLAN (тип 1)	HIPERLAN (тип 2)	HIPERLAN (тип 3)	HIPERLAN (тип 4)
Диапазон	50 м	50–100 м	5000 м	150 м
QoS	Статистическое	Классы информационного обмена ATM (VBR, CBR, ABR, UBR)		
Мобильность	< 10м/с	< 10 м/с	Стационарные узлы	Стационарные узлы
Интерфейс	Обычные ЛВС		Сети ATM	
Скорость передачи данных	23,5 Мбит/с	> 20 Мбит/с	155 Мбит/с	
Энергосбережение	Да	Да	Не обязательно	

Далее в этой главе HIPERLAN 1 будет называться просто HIPERLAN. Текущая версия стандарта [ETSI, 1998] описывает беспроводную локальную сеть, поддерживающую распределение приоритетов и времен "жизни" пакетов для передачи данных со скоростью 23,5 Мбит/с, включая механизмы передачи, изучения топологии, шифрования пользовательских данных, идентификации сети и энергосбережения. Сети HIPERLAN работают в диапазоне 5,1–5,3 ГГц и при уровне излучаемой мощности 1 Вт обеспечивают в помещения и зданиях область охвата до 50 метров.

Службы, предлагаемые HIPERLAN, совместимы со стандартными службами MAC, известными из ЛВС серии IEEE 802.x. Таким образом, адресация узлов основана на использовании 48-битовых адресов MAC. Специальная схема идентификации HIPERLAN разрешает параллельную работу двух или большего числа физически перекрывающихся сетей, не смешивая их сеансы связи. Для обеспечения конфиденциальности предусматривается шифрование информационного обмена с использованием алгоритма шифрования/дешифрования, требующего идентичных ключей и векторов инициализации для успешной дешифровки потока данных, зашифрованного отправителем.

К нововведениям HIPERLAN, не имеющим аналогов в других сетях, относится возможность передачи информационных пакетов с использованием нескольких ретрансляторов. Ретрансляторы могут расширять соединения уровня MAC за пределы диапазона радиопередачи. Узлы сети могут использовать специальную схему запуска после режима ожидания. Данная схема определяет, в какой момент времени узел готов принимать, так что все остальное время узел может отключать свой приемник и экономить энергию. Данным узлам требуются партнеры, содержащие информацию о схеме запуска всех подобных узлов, за которые они отвечают. Подобные обслуживающие станции только перенаправляют данные узлам в момент перехода последних в активное состояние. Для данного действия также требуется механизм буферизации пакетов, предназначенных узлам в режиме ожидания.

В следующих разделах описывается протокольная архитектура HIPERLAN, принятые обозначения, три подуровня, базы данных, формируемые в узлах сети для поддержки маршрутизации в условиях динамической топологии.

7.4.1. Протокольная архитектура

На рис. 7.21 представлено сравнительное описание уровней эталонных моделей HIPERLAN, OSI и IEEE 802.x. Стандарты серии 802.x специфицируют уровень управления логическим каналом LLC и уровень управления доступом к среде MAC, а также физический уровень. В эталонной модели OSI уровни LLC и MAC в сумме формируют канальный уровень DLL.

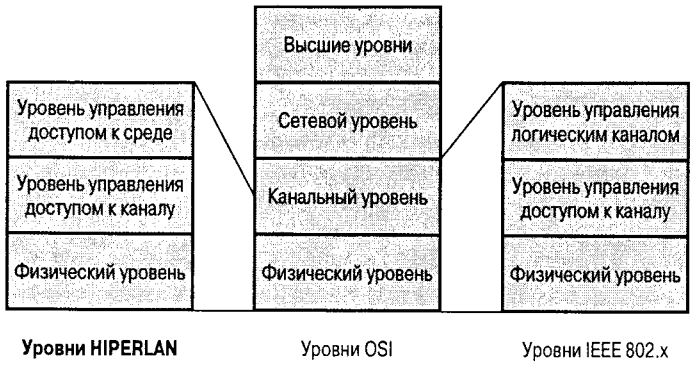

Рис. 7.21. Стандарт HIPERLAN в сравнении со стандартами OSI и IEEE

Стандарт HIPERLAN охватывает только часть канального уровня и физического уровня эталонной модели OSI. Службы MAC уровня HIPERLAN совместимы с определением служб ISO 15802-1 [ISO, 1995] и архитектурой IEEE 802.11a [IEEE, 1990]. Уровень LLC в настоящее время является частью стандартизации HIPERLAN. В то же время в этом стандарте уровень MAC, как он известен из стандартов IEEE, разделен на подуровни собственно **управления доступом к среде** (MAC) и **управления доступом к каналу** (CAC). Значение **физического уровня** (PHY) одинаково во всех названных эталонных моделях.

Ниже приводится пример, объясняющий задачи, модули данных, протоколы и службы всех трех уровней, специфицированных в стандарте HIPERLAN (см. рис 7.22). Уровень MAC предлагает свою службу следующему высшему уровню (обычно — LLC) в точке доступа к службе MSAP. Данная служба включает службу соединения по одной сети HIPERLAN, транспорт модуля данных протокола MAC (MSDU) и изучение доступных сетей HIPERLAN на предмет динамического доступа. Объекты уровня MAC используют протокол MAC для обмена PDU уровня MAC (HMPDU) и предоставления службы MAC. Изучение уровня MAC будет продолжено в разделе 7.4.4.

Уровень CAC специфицирован специально для обеспечения приоритетного доступа к каналу и снятия с высших уровней необходимости слежения за характеристиками беспроводных каналов связи. Служба CAC принимает SDU уровня CAC (HCSDU) в точке HCSAP, и объекты HC обмениваются модулями HCPDU, используя протокол CAC (более подробно этот процесс описан в разделе 7.4.3). Данный уровень содержит схему доступа EY-NPMA, уникальную для сетей HIPERLAN, а следовательно, заслуживающую отдельного рассмотрения.

Наконец, физический уровень (PHY) предлагает механизмы передачи потоков битов и задает радиоканалы, методы модуляции, схемы энергопотребления. Стандарт HIPERLAN не дает подробного описания точки SAP физического уровня. Информация передается в форме пакетов различной величины и с разной скоростью (см. раздел 7.4.2).

7.4.2. Физический уровень

Самый нижний, физический уровень сети стандарта HIPERLAN отвечает за модуляцию/демодуляцию радионесущей потока битов, синхронизацию передатчиков и приемников, а также предоставление механизмов прямой коррекции ошибок, измерение интенсивности сигнала и прослушивание канала.

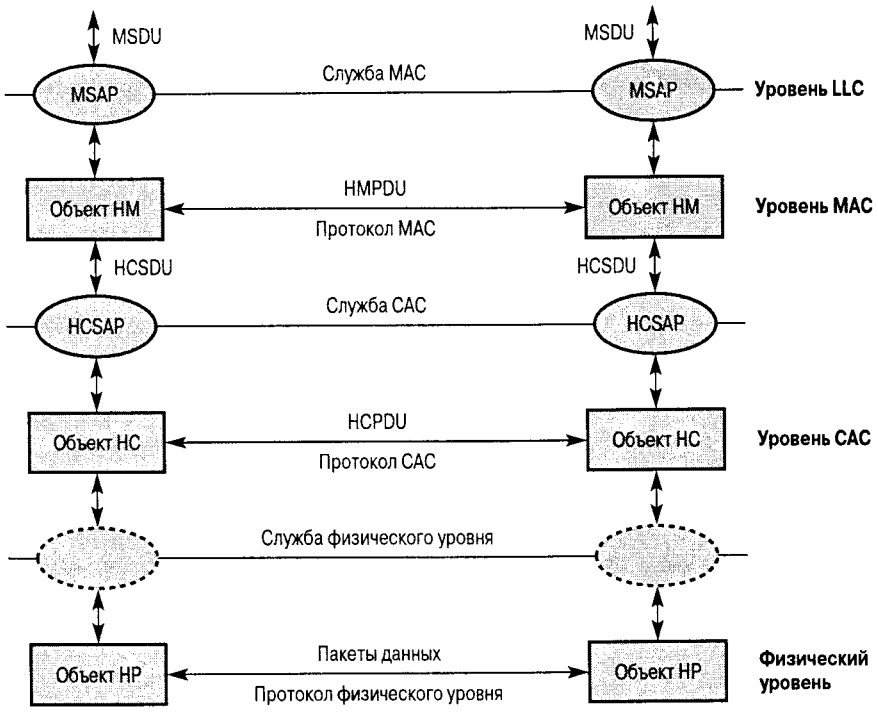

Рис. 7.22. HIPERLAN: протоколы, службы и модули данных

Стандарт задает три обязательных несущих и две возможных. **Обязательные несущие** центрированы на следующих частотах.

- $F(0) = 5{,}1764680$ ГГц
- $F(1) = 5{,}1999974$ ГГц
- $F(2) = 5{,}2235268$ ГГц

Две дополнительные частоты, которые доступны не во всех странах.

- $F(3) = 5{,}2470562$ ГГц
- $F(4) = 5{,}2705856$ ГГц

Ширина полосы при использовании данных несущих составляет 23 МГц.

Постоянное поддержание высокой скорости передачи 23,5294 Мбит/с (±235 бит/с) не всегда оправдано и целесообразно для мобильных клиентов, так как требует повышенного энергопотребления. В связи с этим HIPERLAN задает два типа информационных пакетов (см. рис. 7.23).

Рис. 7.23. Пакет LBR-HBR

- **Информационный пакет LBR-HBR**. Состоит из **низкоскоростного** (LBR) потока, синхронизирующей и настроечной последовательностей и нескольких **высокоскоростных** (HBR) ин-

формационных блоков. Разрешено от одного до 47 информационных блоков, содержащих по 496 бит каждый. В то же время для высоких скоростей беспроводного терминала максимальное число информационных блоков нужно снижать. Синхронизирующая и настроечная последовательности включают по 450 бит и строго определены в стандарте HIPERLAN.

- **Информационный пакет LBR**. Блоки данных этого типа состоят исключительно из **низкоскоростной** (LBR) части, содержащей идентификатор получателя. Часть LBR используется только при скорости передачи 1,4705875 Мбит/с (±15 бит/с).

Высокоскоростные блоки в пакете LBR-HBR кодируются с использованием схемы прямой коррекции ошибок, кода БХЧ (Bose-Chadhuri-Hocquenghem, BCH, код Боуза-Чоудхури-Хоквенгема), чередования битов, периодической инверсии битов и дифференциального предварительного кодирования [ETSI, 1998].

Низкоскоростная часть модулируется с использованием простой частотной манипуляции (FSK), причем частота $F(i) - 386$ кГц обозначает логический нуль, а частота $F(i) + 386$ кГц — логическую единицу. $F(i)$ — центральная частота, определенная стандартом. Для высокоскоростной части используется гауссова манипуляция с минимальным сдвигом (GMSK). Мощность передатчиков HIPERLAN может принимать значения 1 Вт, 100 мВт либо 10 мВт.

7.4.3. Подуровень управления доступом к каналу

Подуровень управления доступом к каналу (CAC) предлагает вышестоящему уровню MAC службу передачи данных без установления соединения. Передаваемые модули данных называются **модулями данных службы уровня CAC** (HCSDU). Передача начинается в **точке доступа к службе уровня CAC** (HCSAP) источника и заканчивается в HCSAP адресата (или нескольких адресатов). При передаче данных максимальная длина блока составляет 2 422 байта. Кроме того, уровень MAC, использующий данную службу, может задать приоритет, **качество обслуживания CAC** (HCQoS), определяющий приоритет доступа к каналу, подобно тому, как это реализовано в механизме EY-NPMA, который описывается далее.

Адрес HCSAP может быть индивидуальным, определяющим одну точку HCSAP, или групповым, определяющим группу точек HCSAP. Отдельные адреса HCSAP могут использоваться как исходные или целевые, тогда как групповой адрес может использоваться только как целевой адрес (точнее, адреса). Данные адреса соответствуют спецификации 48-битовых MAC-адресов сетей IEEE 802.x.

Для передачи данных подуровень CAC использует два примитива службы.

- **HC-UNITDATA.req** (адрес HCSAP отправителя, адрес HCSAP назначения, модуль HCSDU, идентификатор сети HIPERLAN, приоритет доступа к каналу). Этот примитив используется для передачи данных.
- **HC-UNITDATA.ind** (адрес HCSAP отправителя, адрес HCSAP назначения, модуль HCSDU, идентификатор сети HIPERLAN). Данный примитив службы указывает входящие данные.

Идентификатор сети HIPERLAN служит для разграничения нескольких сетей. Это значение может быть установлено на 0 (указывая на любую сеть HIPERLAN) и определяется уровнем MAC.

Название **EY-NPMA** (elimination-yield non-preemptive priority multiple access — исключение, порождающее приоритетный множественный доступ) — это не просто сложный акроним, но также и суть доступа к каналу с использованием схемы приоритетов и разных схем доступа. EY-NPMA делит доступ к среде различных состязающихся узлов на три фазы.

- **Установление приоритетов**. Определение наивысшего приоритета пакета данных, готового к передаче на состязающихся узлах.

- **Состязание**. Устранение всех конкурентов, за исключением одного (если наивысший приоритет имеет более одного узла).
- **Передача**. Собственно передача блока данных оставшимся узлом.

Если за право доступа к среде состязаются несколько узлов, необходимыми являются все три этапа (это называется "доступ к каналу в **синхронизированном состоянии**"). Если же канал не занят, по крайней мере, 2 000 периодов высокоскоростной передачи плюс некоторый динамический промежуток времени, требуется только третья фаза, т.е. передача (это называется "доступ к каналу **при свободном канале**"). Упомянутый динамический промежуток времени случайно (и с равной вероятностью) выбирается из диапазона от 0 до 3 периодов высокоскоростной передачи 200 бит. Данное расширение времени далее минимизирует вероятность конфликтов при доступе к свободному каналу станций, синхронизированных на высших уровнях и пытающихся одновременно получить доступ к свободному каналу. Помимо двух названных состояний HIPERLAN поддерживает "доступ к каналу в **состоянии устранения скрытых терминалов**", название которого говорит само за себя [ETSI, 1998].

Фаза состязания в свою очередь подразделяется на два этапа — **фазу исключения** и **фазу выпуска**. Целью фазы исключения является отсеивание как можно большего числа (но, разумеется, не всех) узлов из состязающихся за доступ к каналу. Результатом фазы исключения является более или менее постоянное число узлов, почти независимое от исходного числа состязающихся узлов. Фаза выпуска завершает фазу исключения таким образом, чтобы остался всего один "победивший" узел.

На рис 7.24 представлены три основные фазы и некоторые подробности, которые рассматриваются в следующих разделах. Для каждого узла, готового передавать данные, цикл доступа начинается с синхронизации текущего отправителя. Следует первая фаза, установление приоритетов. После этого идут фазы исключения и выпуска (в сумме дающие фазу состязания). В завершение оставшийся узел может передавать свои данные. Каждая фаза имеет определенную продолжительность, измеряемую числом слотов и обозначенную переменными I_{PS}, I_{PA}, I_{ES}, I_{ESV} и I_{YS}.

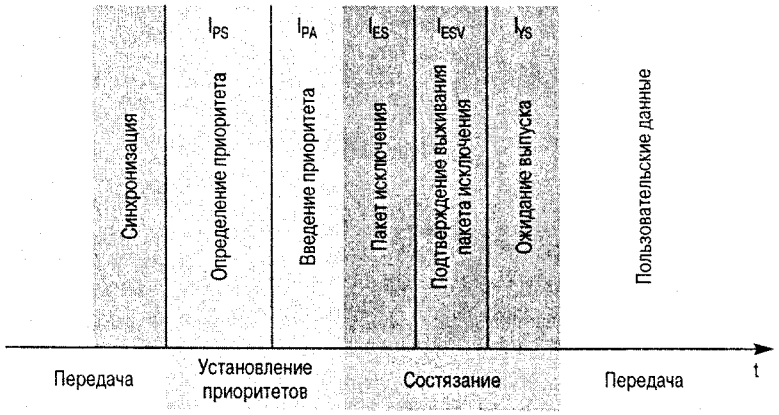

Рис. 7.24. Фазы схемы доступа EY-NPMA

7.4.3.1. Фаза установления приоритетов

HIPERLAN предусматривает пять уровней приоритета для пакетов, готовых к передаче. После завершения очередного сеанса передачи за право отправления пакета могут состязаться несколько узлов. Цели фазы установления приоритетов — гарантировать узлам своевременный доступ к каналу в соответствии с назначенными приоритетами. Данный механизм всегда дает доступ к среде узлам с высоким приоритетом вне зависимости от того, насколько велика загрузка на узлах с более низкими приоритетами.

На первом этапе фазы, выявлении приоритетов, время делится на пять слотов: от слота 0 (наивысший приоритет) до 4 (нижайший приоритет). Длительность каждого слота $I_{PS} = 168$ периодов HBR. Узел, имеющий уровень приоритета p, прослушивает среду в течение p слотов. Если в это время среда оставалась свободной, узел немедленно занимает канал, отправляя пакет продолжительностью $I_{PA} = 168$ периодов HBR (установление приоритета). Пакет состоит из следующей последовательности (передаваемой на скорости HBR), повторяемой столько раз, сколько требуется для заполнения всего пакета.

11111010100010011100000110010110

Если узел обнаруживает активность в среде, он прекращает попытки передачи данных в данном цикле и ожидает следующего цикла. Фаза установления приоритета считается законченной, как только какой-либо из состязающихся узлов установит свой приоритет (путем передачи пакета). Таким образом, продолжительность фазы не является фиксированной величиной, а определяется текущим наивысшим приоритетом.

Рассмотрим пример, в котором за доступ к среде состязаются три узла сети. Пакеты узла 1 и 2 имеют приоритет 2, а пакеты узла 3 — приоритет 4. Узлы 1, 2 и 3 прослушивают среду и обнаруживают незанятыми слоты 0 и 1. Затем в слоте 2 узлы 1 и 2 посылают пакет в виде установления приоритета. После этого узел 3 прекращает свои попытки передачи. В данном примере длительность фазы установления приоритетов равна трем слотам.

После описанной первой фазы побеждает, по крайней мере, один узел. Возможны варианты "выживания" нескольких узлов, имеющих одинаковый наивысший приоритет.

7.4.3.2. Фаза исключения

Как уже отмечалось, после фазы установления приоритетов может остаться несколько узлов. Алгоритм фазы исключения предусматривает подготовку и передачу узлами сети специального пакета — пакета исключения. На этапе исключения время ожидания разбивается на слоты длительностью $I_{ES} = 212$ периодов HBR. Продолжительность передачи каждого отдельного пакета исключения может составлять от 0 до 12 слотов I_{ES} в зависимости от длины пакета. Вероятность пакетной передачи в течение слота равна 0,5. Вероятность $P_E(n)$ того, что длительность пакета исключения будет равна n слотам исключения, равна

- $P_E(n) = 0{,}5^{n+1}$, если $0 < n < 12$
- $P_E(n) = 0{,}5^{12}$, если $n = 12$

Состязание в фазе исключения разрешается посредством передачи пакетов исключения и определения "выживших" пакетов. Каждый узел, претендующий на доступ к каналу, с определенной вероятностью отправляет пакет исключения длиной n слотов, а затем прослушивает канал в течение времени $I_{ESV} = 256$ периодов HBR. Посылаемый пакет идентичен используемому для введения приоритета. Состязающийся узел успешно проходит фазу исключения тогда и только тогда, когда после передачи пакета канал не занимается в течение времени I_{ESV}. В противном случае узел выбывает из состязания и прекращает попытки передачи пакета до следующего цикла.

Общая продолжительность фазы исключения определяется продолжительностью самого длинного пакета исключения из отсылаемых узлами плюс время I_{ESV}. Окончательно получаем один или несколько узлов, прошедших фазу исключения, которые затем переходят к следующей фазе.

7.4.3.3. Фаза выпуска

В течение следующей фазы, фазы выпуска, оставшиеся узлы только прослушивают канал, не передавая новых пакетов. Как и ранее, время этапа разбивается на слоты продолжительностью $I_{YS} = 168$ периодов HBR. Время прослушивания канала каждым узлом может с равной вероят-

ностью составлять от 0 до 9 слотов. Таким образом, вероятность $P_Y(n)$ того, что время прослушивания фазы выпуска составит n слотов, равна 0,1 для всех n, $0 < n \leq 9$.

Теперь каждый узел прослушивает среду в течение выбранного промежутка времени. Если в течение этого интервала канал остается свободным, фаза выпуска для узла заканчивается, и узел может начинать передачу данных. В противном случае узел отказывается от попыток передачи до завершения текущего цикла передачи. Теперь продолжительность фазы утверждения определяется длительностью наименьшего периода прослушивания, имеющегося среди состязающихся узлов. По окончании данной фазы будет, по меньшей мере, один узел, который начнет передачу своих данных. Эту передачу обнаружат другие узлы, имеющие большее время прослушивания. Важно отметить, что на этот момент может существовать более одного выжившего узла, так что описанный механизм доступа не исключает возникновения конфликтов.

7.4.3.4. Фаза передачи

Узел, прошедший фазы установления приоритетов и состязания, получает возможность передавать свои данные, которые называются модулями данных протокола (PDU) CAC стандарта HIPERLAN с низкой и высокой скоростями передачи (LBR-HBR HCPDU). Данные PDU могут применяться как для одно-, так и для многоадресной передачи. В первом случае отправитель ожидает получения немедленного подтверждения со стороны адресата, которое называется подтверждающим HCPDU (AK-HCPDU) и представляет собой блок, содержащий только часть LBR.

7.4.3.5. Формат HCPDU

Как отмечалось в разделе 7.4.2, кадры, которыми обмениваются узлы сети HIPERLAN, могут состоять из **низкоскоростной** (LBR) части, периода синхронизации и **высокоскоростной** (HBR) части. Такая структура отражена в модуле LBR-HBR HCPDU. В следующем ниже разделе представляется структура информационного PDU подуровня CAC, называющегося DT-HCPDU (рис. 7.25).

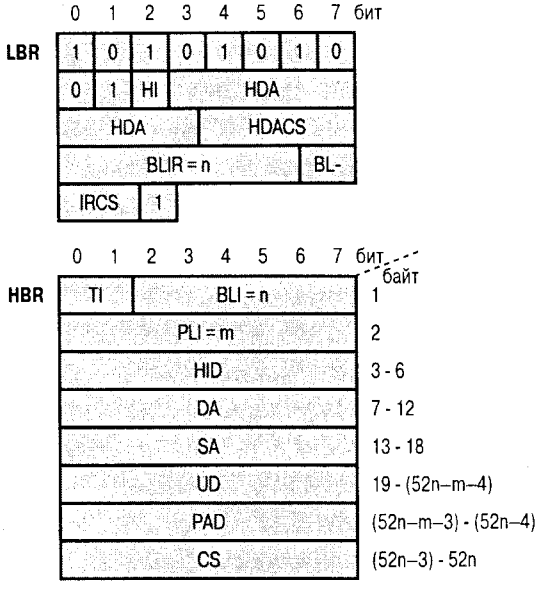

Рис. 7.25. Информационный PDU подуровня CAC (DT-HCPDU)

Все модули HCPDU начинаются с низкоскоростной (**LBR**) части. Этот блок имеет следующую структуру. Первые 10 бит — фиксированная последовательность 1010101001. Следующее

поле, индикатор высокоскоростной части (**HI**), указывает на наличие данных HBR. Поскольку модуль DT-HCPDU всегда включает скоростные блоки, значение поля HI здесь равно 1. Следующие 9 бит представляют **хешированный адрес целевой точки HCSAP** (поле **HDA**). Для контроля безошибочной передачи HDA передается 4-битовая контрольная сумма **HDACS**, вычисляемая с использованием полинома $G(x) = x^4 + x + 1$. Следующее 6-битовое поле **BLIR** (копия индикатора длины блока) указывает число блоков части HBR и здесь является копией поля **BLI**. Блок части HBR состоит из 52 байт, всего может быть от 1 до 47 таких блоков. Поле BLIR также защищено контрольной суммой **BLIRCS**, причем использован тот же полином $G(x)$. Последнее поле блока LBR — всегда 1.

Между блоками LBR и HBR физический уровень вводит синхронизирующую последовательность (см. раздел 7.4.2). Блок **HBR** начинается с идентификатора типа **TI** — 2-битового поля, равного 1 для информационного PDU. Как упоминалось выше, блок HBR содержит поле **индикатора длины блока** (**BLI**), указывающее число блоков в части HBR. Следующий байт — это **индикатор длины заполнения** (**PLI**), который содержит число байт, использованных в поле заполнения PAD.

Следующие три поля идентифицируют отправителя, получателя и сеть HIPERLAN. Первое 4-байтовое поле, **идентификатор сети HIPERLAN** (**HID**), служит для определения сети HIPERLAN (в одной территориальной области может существовать несколько различных сетей HIPERLAN). **Адрес целевой HCSAP** (поле **DA**) — это 48-битовый MAC-адрес ЛВС, подобный используемому в сетях IEEE 802.x и определяющий SAP назначения. Подобным образом адрес исходной HCSAP (поле **SA**) представляет исходную точку доступа к службе (SAP).

Поле **пользовательских данных** (**UD**) содержит от 1 до 2 422 байт пользовательской информации. При необходимости PDU может дополняться **полем заполнения** (**PAD**) длиной от 0 до 51 байта. Последним полем является контрольная сумма **CS**, защищающая блок HBR. Данная контрольная сумма генерируется с использованием полинома $H(x) = x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$.

Подтверждающий модуль PDU (AK-HCPDU) состоит только из низкоскоростной части (так называемой LBR HCPDU), как показано на рис. 7.26. Таким образом, после фиксированного кода 10101010001 следует поле HI (теперь — 0). Далее идет **идентификатор подтверждения** (поле **AID**), который также защищается с использованием полинома $G(x)$ (соответствующая контрольная сумма обозначена как AIDCS).

7.4.4. Подуровень управления доступом к среде

Подуровень MAC предлагает службы без организации соединения для передачи MSDU от исходной MSAP к одной целевой MSAP или группе целевых MSAP. Модули MSDU передаются без ограничений на содержимое. Кроме того, данная передача может ограничиваться по времени или же иметь специальный уровень приоритета или время жизни, определяемые пользователем службы. Адреса, использованные для идентификации точек MSAP идентичны адресам HCSAP, т.е. являются обычными 48-битовыми MAC-адресами.

Качество обслуживания, предлагаемое уровнем MAC, основывается на трех параметрах (параметры **HMQoS**). Пользователь может установить **приоритет** данных: высокий приоритет — 0, низкий — 1. Более того, пользователь может определить время жизни MSDU для указания на срочную доставку. **Время жизни MSDU** определяется как предельное время, которое может пройти между отправлением и приемом блока прежде, чем данные устареют и передача потеряет смысл. Время жизни MSDU составляет порядка 0–16 000 мс. Наконец, **остаточное время жизни MSDU** указывает оставшееся время жизни пакета.

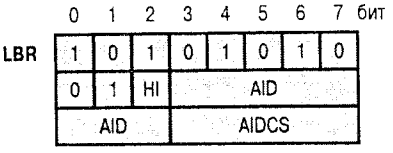

Рис. 7.26. Подтверждающий PDU подуровня CAC (AK-HCPDU)

Для передачи данных уровень MAC предлагает два примитива службы.

- **HM-UNITDATA.req** (адрес источника, адрес назначения, MSDU, приоритет данных, время жизни MSDU). Этот примитив может использоваться для передачи MSDU с определенным приоритетом одному или нескольким адресатам.
- **HM-UNITDATA.ind** (адрес источника, адрес назначения, MSDU, приоритет данных, время жизни MSDU, остаточное время жизни MSDU). Этот примитив обозначает принятый MSDU и дополнительно включает остаточное время жизни модуля данных.

Помимо передачи данных уровень MAC предлагает функции поиска других сетей HIPERLAN, находящихся в диапазоне радиоохвата, а также специальные функции энергосбережения. **Энергосбережение** достигается за счет настройки определенных периодических схем перехода в активный режим получения вместо постоянного нахождения в нем. Для разрешения многоадресной передачи все узлы, входящие в группу многоадресной передачи, должны быть готовы принимать, когда отправитель посылает сообщение, адресованное данной группе.

MAC-уровень HIPERLAN также предлагает возможность **шифрования** и **дешифрования** информации с использованием простой схемы исключающего ИЛИ плюс случайные числа. Ключ шифрования выбирается из набора ключей с помощью идентификатора ключа KID и используется совместно с вектором инициализации IV для активации генератора псевдослучайных чисел. Шифрование заключается в применении операции исключающего ИЛИ к пользовательским данным с целью генерации шифрованных данных. Дешифрование пользовательских данных производится подобным образом с использованием аналогичной последовательности случайных чисел. Поскольку описанная схема не является особо сильным шифрованием, на высших уровнях применяются более сложные схемы.

Рассмотрим, как на уровне MAC выбирается следующий PDU для передачи при наличии нескольких PDU и как время ожидания передачи сказывается на приоритете PDU. Итак, поскольку служба является срочной, выбор должен отражать пользовательский приоритет (0 или 1) и учитывать остаточное время жизни. Уровень MAC должен отобразить эту информацию в приоритет доступа к каналу, используемый уровнем CAC, состязаясь с другими узлами за право передачи.

Прежде всего, уровень MAC определяет **нормированное остаточное время жизни (NRL)** — остаточное время жизни, деленное на приблизительное число маршрутных участков, которые должен пройти PDU. Данный параметр включает как время ожидания PDU в узле, так и расстояние, а следовательно, дополнительные промежутки ожидания в других узлах. Затем уровень MAC вычисляет приоритет доступа к каналу для каждого PDU, используя схему отображения, приведенную в табл. 7.3.

Окончательный выбор важнейшего HMPDU выполняется следующим образом.

- Выбирается HMPDU с наивысшим приоритетом.
- Из них выбираются все HMPDU с наименьшим NRL.
- Из полученных HMPDU произвольным образом выбирается один.

Таблица 7.3. Отображение нормированного остаточного времени жизни (NRL) в приоритет CAC

NRL	Приоритет MSDU = 0	Приоритет MSDU = 1
NRL < 10 мс	0	1
10 мс ≤ NRL < 20 мс	1	2
20 мс ≤ NRL < 40 мс	2	3
40 мс ≤ NRL < 80 мс	3	4
80 мс ≤ NRL	4	4

На рис. 7.27 представлена структура информационного PDU (DT-HMPDU). Первые 2 байта, **индикатор длины LI**, представляют размер HMPDU в байтах. Поле **индикатор типа TI** содержит указание типа HMPDU (т.е. 1 для DT-HMPDU). Другие значения TI зарезервированы для управления топологией, поиска адресов, схем перехода в активный режим, и т.д. [ETSI, 1998]. Следующие два байта указывают **остаточное время жизни модуля RL** и используются для исчисления приоритета.

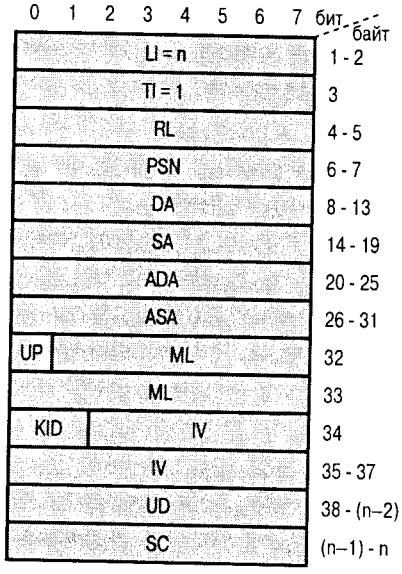

Рис. 7.27. Информационный PDU подуровня MAC (DT-HMPDU)

Поле **порядкового номера PDU (PSN)** используется для различения и определения относительного порядка PDU, сгенерированных одним объектом HM. Следует отметить, что порядковые номера присваиваются модулям PDU и не являются счетчиками байтов. PSN очередного PDU вычисляется как [(предыдущий порядковый номер + 1) по модулю 65536]. Если PDU A и PDU B имеют порядковые номера a и b, соответственно, A следует за B, если [(65536 + (a – b)) по модулю 65536] < 32768. В противном случае модуль B следует за A.

Следующие два поля, **DA** и **SA**, представляют **целевой адрес MSAP** и **исходный адрес MSAP** соответственно. Оба адреса совместимы с 48-битовыми адресами MAC. Следующие два поля также содержат адреса: **целевой альтернативный адрес MSAP** и **исходный альтернативный адрес MSAP**. Эти адреса также согласуются со стандартом 48-битовых адресов MAC. Альтернативные адреса используются для идентификации точек MSAP вне сети и описываются в разделе 7.4.5.

Далее идет поле пользовательского приоритета **UP**. Следующее поле **ML** имеет значение времени жизни модуля данных. По умолчанию для MPDU устанавливается время 500 мс (значение поля — 500), максимальное значение составляет 16000 (возможно увеличение до 32767). Поля **идентификатор ключа KID** и **вектор инициализации IV** используются для описанного выше дешифрования данных. **Данные пользователя UD** начинаются с 38-го байта блока. Поскольку служба CAC ограничивает размер HCSDU величиной 2 482, служба MAC ограничивает пользовательские данные величиной 2 383 байт (далее могут быть и другие ограничения). Последнее поле блока данных, **санитарная проверка SC**, является не строгой контрольной суммой, а особой суммой нешифрованных октетов HMPDU.

7.4.5. Информационные базы и организация сетей

В HIPERLAN реализована возможность перенаправления информационного обмена через несколько других беспроводных узлов, что особенно важно в беспроводных и эпизодических сетях, не имеющих инфраструктуры. Механизм перенаправления может также использоваться, когда удаленные узлы не имеют прямой связи с точкой доступа.

На рис. 7.28 рассматривается пример эпизодической сети HIPERLAN из шести узлов. Узлы, выполняющие функции транзитной передачи пакетов, называются **ретрансляторами**. Для направления собственных пакетов или перенаправления чужих узлы содержат **базу данных маршрутизации (RIB)**. Позиции этой базы являются триплетами (R_{dest}, R_{next}, R_{dist}), где R_{dest} — адрес MSAP приемника, R_{next} — адрес HCSAP следующего транзитного узла на пути к пункту назначения, R_{dist} —предполагаемое число маршрутных участков до целевого узла.

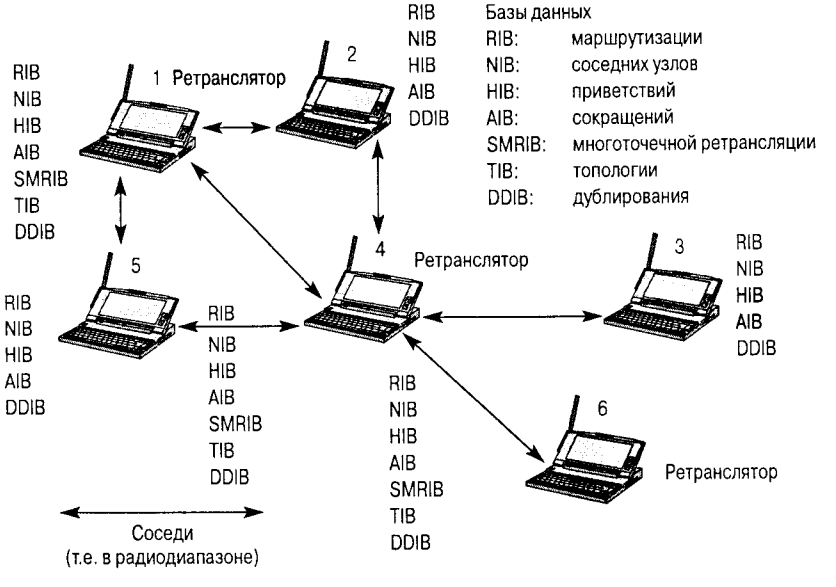

Рис. 7.28. Пример двух эпизодических сетей

В каждом узле сети формируется **база данных соседей NIB** с информацией о каждом канале связи узла с его соседями. С интервалом в 10 с (или меньше) каждый узел объявляет свою базу NIB всем соседним узлам, используя модуль HO-HMPDU ("HellO-HMPDU"). Записи (N_{nbor}, N_{status}) описывают текущее состояние канала связи N_{status} относительно соседнего узла N_{nbor} (который указывается адресом HCSAP). Каждая запись хранится 20 с, после чего удаляется. N_{status} может принимать следующие значения.

- **N_Asym**, если с узлом-соседом установлена асимметричная связь.
- **N_Sym**, если связь с соседним узлом симметрична.
- **N_MultiRelay**, если связь с соседним узлом симметрична и узел-сосед выбран в качестве многоточечного ретранслятора.

База данных приветствий HIB хранит дополнительную информацию об узлах-соседях по HIPERLAN. Запись представляет собой триплеты (H_{dest}, H_{status}, H_{next}). H_{dest} — адрес HCSAP назначения, H_{next} — адрес HCSAP соседнего узла, посредством которого можно достичь H_{dest}. H_{status} несет информацию о состоянии объекта HM по адресу HCSAP H_{next} и может принимать следующие значения.

- **H_NeighborNF**. Узел H_{dest} является соседним узлом, но не является ретранслятором, поэтому $H_{dest} = H_{next}$.
- **H_NeighborF**. Узел H_{dest} является соседним узлом-ретранслятором, поэтому $H_{dest} = H_{next}$.
- **H_TwoHop**. H_{dest} находится в двух маршрутных участках от данного узла, H_{next} — это адрес HCSAP соседнего ретранслятора.

Узлы HIPERLAN могут сообщаться с внешними (по отношению к данной сети) узлами на уровне MAC, т.е. модуль MSDU должен доставляться целевой MSAP, отличной от точек MSAP сети. Следовательно, узел HIPERLAN может хранить таблицу отображения внешних MSAP в так называемые альтернативные MSAP, принадлежащие данной сети HIPERLAN. Это позволяет применить механизм маршрутизации HIPERLAN и базы данных RIB. Альтернативные имена для внешних узлов хранятся в парах (A_{ori}, A_{alias}) в **базе данных альтернативных имен AIB**. A_{ori} — это исходный адрес MSAP объекта вне HIPERLAN, а A_{alias} — связанный с этим объектом псевдоним. Данный механизм допускает независимое назначение адресов MSAP в сети HIPERLAN и вне ее и предлагает простой механизм выполнения отображения. Все позиции базы данных имеют определенное время поддержки, по истечении которого удаляются.

Узлы HIPERLAN, исполняющие роль ретрансляторов, дополнительно требуют двух баз данных — информационную базу исходных многоточечных ретрансляторов SMRIB и информационную базу топологии TIB (узлы 1, 4 и 6 в примере на рис. 7.28). Каждый узел выбирает соседний ретранслятор в качестве многоточечного ретранслятора, который становится так называемым исходным многоточечным ретранслятором данного узла (например, узел 5 в качестве своего ретранслятора может выбрать узел 1). Теперь ретранслятор формирует **информационную базу исходных многоточечных ретрансляторов (SMRIB)**. Позиции этой базы данных представляют собой пары (S_{SMR}, S_{seq}), где S_{SMR} задает узел (по точке HCSAP), а S_{seq} — порядковый номер узла во множестве многоточечных ретрансляторов. Все узлы-ретрансляторы обмениваются друг с другом базами SMRIB, обновление происходит каждые 20 с с использованием пакетов TC-HMPDU (Topology Control-HMPDU), время хранения записей SMRIB — 40 с.

Узлы множества многоточечных ретрансляторов используются для перенаправления HMPDU от локального узла ко всем узлам данной сети HIPERLAN. Выбирая это множество для каждого конкретного случая передачи, можно оптимизировать передачу данных в соответствии с каким-либо критерием. Каждый узел выбирает для себя множество многоточечных ретрансляторов на основе информации баз данных NIB и HIB. Соседний узел может входить в это множество, если он является ретранслятором и его параметр N_{status} имеет значение N_Sym или N_MultiRelay. Кроме того, узел должен проверить в базе HIB каждый объект HM с H_{status} = H_TwoHop (т.е. с H_{dest} в двух маршрутных участках от данного узла), чтобы множество многоточечных ретрансляторов содержало, по крайней мере, один ретранслятор, обозначенный H_{next}. Другими словами, узлы множества многоточечных ретрансляторов должны быть ретрансляторами, локальный узел должен иметь симметричный канал связи с узлами и доступ к узлам, указанным в своей базе HIB, посредством

этих ретрансляторов. Множество многоточечных ретрансляторов меняется со временем, поэтому с каждым множеством соотносится порядковый номер.

Каждый ретранслятор создает **информационную базу топологии TIB**, в которой отражена текущая топология сети HIPERLAN, основываясь на данных о ретрансляторах и их исходных многоточечных ретрансляторах. Каждая запись базы является триплетом (T_{dest}, T_{last}, T_{seq}), несущим информацию о том, что узел с адресом T_{dest} точки HCSAP выбрал в качестве многоточечного ретранслятора узел с адресом T_{last} с порядковым номером T_{seq}. Такая запись показывает, как достичь узла T_{dest} за один маршрутный участок — т.е. посредством ретранслятора T_{last}.

Информационная база обнаружения копий (DDIB) является независимой от маршрутизации, ретрансляции и топологии. В беспроводных сетях существует множество возможностей дублирования пакетов при передаче. Например, иногда единственный способ передать одноадресный пакет — это использовать многоадресную или широковещательную передачу через несколько промежуточных узлов, если точный маршрут к целевому узлу неизвестен. Для исключения возникающего дублирования при приеме в узлах сети создаются базы DDIB. Чтобы избежать избыточной обработки, каждый узел хранит информацию о принимаемых пакетах в виде записи (D_{src}, D_{seq}). D_{src} — это адрес HCSAP источника данных, а D_{seq} — порядковый номер HMPDU. Записи базы DDIB хранятся определенное время и затем обновляются.

Для поддержания энергосберегающего режима работы узлов сети ретрансляторы формируют базы данных с информацией о схемах перехода в активный режим узлов, которым отправляет информацию данный ретранслятор [ETSI, 1998].

Рассмотрим примеры баз данных узлов 1 и 5 с рис. 7.28. Предположим, что узлы 5 и 2 выбрали узел 1 в качестве ретранслятора данных. Тогда узел 1 в своей базе данных исходных многоточечных ретрансляторов фиксирует записи, относящиеся к узлам 5 и 2 (номера узлов в данном случае играют роль адресов MSAP и HCSAP).

Ниже приводятся примеры баз данных узла 5.

- **RIB**. База данных RIB всегда является мгновенным снимком текущей топологии, и направления к другим узлам могут меняться со временем. Текущую топологию можно представить следующими записями. $\{(R_{dest}, R_{next}, R_{dist})\} = \{(1, 1, 1), (2, 1, 2), (3, 1, 3), (4, 1, 2), (6, 1, 3)\}$.
- **NIB**. Данная база может иметь следующий вид: $\{(N_{nbor}, N_{status})\}$ = {(4, N_Asym), (1, N_MultiRelay)}.
- **HIB**. Информация в этой базе может иметь следующий вид: $\{(H_{dest}, H_{status}, H_{next})\}$ = {(1, H_NeighborF, 1), (4, H_NeighborF, 4), (2, H_TwoHop, 1), (4, H_TwoHop, 1), (1, H_TwoHop, 4), (2, H_TwoHop, 4), (6, H_TwoHop, 4), (3, H_TwoHop, 4)}.

Ретранслятор 1 содержит следующие базы данных.

- **RIB**: $\{(R_{dest}, R_{next}, R_{dist})\} = \{(5, 5, 1), (2, 2, 1), (4, 4, 1), (3, 4, 2), (6, 4, 2)\}$.
- **NIB**: $\{(N_{nbor}, N_{status})\} = \{(5, N_Sym), (2, N_Sym), (4, N_MultiRelay)\}$.
- **HIB**: $\{(H_{dest}, H_{status}, H_{next})\}$ = {(5, H_NeighborNF, 5), (4, H_NeighborF, 4), (2, H_NeighborNF, 2), (3, H_TwoHop, 4), (6, H_TwoHop, 4)}.
- **SMRIB**: $\{(S_{SMR}, S)\} = \{(5, 4711), (2, 1234), (4, 30967)$. Порядковые номера, показанные для исходных многоточечных ретрансляторов, выбираются S_{SMR}.
- **TIB**: $\{(T_{dest}, T_{last}, T_{seq})\} = \{(2, 1, 1234), (5, 1, 4711), (4, 1, 30967), (3, 4, 6415), (6, 4, 22058)\}$.

7.5. Bluetooth

По сравнению с технологиями беспроводных ЛВС, рассмотренными в разделах 7.3 и 7.4, в основу Bluetooth была положена идея создания беспроводной эпизодической сети малого радиуса

действия в радиочастотном диапазоне, объединяющей небольшие устройства без требования строгой инфраструктуры. Соединение различных малых устройств без дорогих кабелей или необходимости инфраструктуры требует сетей различных типов [Bisdikian, 1998]. Многие современные устройства предлагают интерфейс IrDA со скоростями передачи, например, около 115 Кбит/с. Проблемами интерфейса IrDA является крайне малая дальность действия (обычно 2 м для встроенных интерфейсов), необходимость наличия прямой видимости между передатчиком и приемником и возможность связи только двух устройств, т.е. поддержка только двухточечных соединений. IrDA не допускает организации сетей, доступа к среде или других расширенных механизмов связи. Большим достоинством IrDA является низкая стоимость; таким образом, данный интерфейс используется практически во всех мобильных устройствах (переносные компьютеры, PDA, сотовые телефоны).

После нескольких лет исследовательских работ в области многоточечной связи, начатых фирмой Ericsson в 1994 году [Haartsen], весной 1998 года пять компаний (Ericsson, Intel, IBM, Nokia, Toshiba) организовали консорциум Bluetooth[1]. Созданная новая технология — это цельная, дешевая, беспроводная сетевая технология. Многие другие компании и исследовательские учреждения объединили свои усилия, в результате чего была создана рабочая группа Special Interest Group (1998) с целью внедрения технологии Bluetooth к концу 1999 года. Это означает, что до настоящего времени Bluetooth — это не утвержденный стандарт, подобный IEEE 802.11 или HIPERLAN 1, а технология, которая вскоре станет стандартом де-факто, как пророчит консорциум Bluetooth.

7.5.1. Пользовательские сценарии

Для беспроводных пикосетей можно представить несколько потенциальных областей применения.

- **Соединение периферийных устройств**. В настоящее время с настольным компьютером соединяются многие устройства (клавиатура, мышь, джойстик, наушники, динамики). Проводное соединение этих устройств является неудобным и требует места, различных разъемов и кабелей. Беспроводный способ освобождает от путаницы проводов, но требует батарейного питания периферии.
- **Поддержка организации эпизодических сетей**. Например, несколько людей собрались вместе, чтобы обсудить некоторые вопросы, обменяться информацией (планы, графики продаж и т.п.); лекция, когда учебный материал централизованно распространяется с ПК преподавателя на PDA студентов. Взаимодействие такого типа легко организовать с использованием беспроводных сетей; в то же время беспроводные устройства могут не иметь адаптеров БЛВС, согласующихся со стандартом IEEE 802.11; проще использовать чипы Bluetooth, встроенные в мобильные телефоны, PDA и настольные ПК.
- **Выполнение роли межсетевого интерфейса**. Мобильный телефон со встроенным чипом Bluetooth может исполнить роль моста между сетью Bluetooth и сетью GSM (см. рис. 7.29). Например, мобильный телефон принимает электронную почту посредством системы GSM и переадресовывает данные на переносной компьютер, находящийся в кейсе. Пока служащий фирмы поднимается в свой офис, PDA или переносной компьютер, находящийся в кармане или кейсе, синхронизируется с файловым сервером в здании и принимает деловую информацию.

[1] Bluetooth (англ. "Синий зуб") — это буквальный перевод имени средневекового датского короля викингов Blatand.

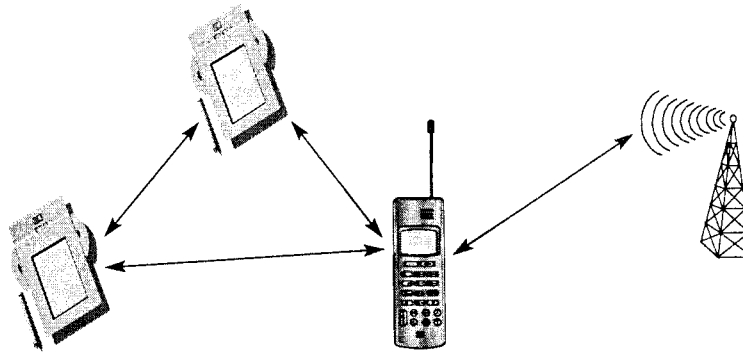

Рис. 7.29. Пример конфигурации с пикосетями на основе Bluetooth

Следует отметить, что при сравнении беспроводных технологий выбор Bluetooth предпочтителен с точки зрения стоимости устройств. С технической стороны технологии, описанные ранее, проектировались для больших территорий охвата и скоростей передачи, поэтому они дороже.

7.5.2. Физический уровень

При разработке физического уровня Bluetooth учитывались несколько ограничений. Устройства Bluetooth должны интегрироваться в мобильные приборы, питающиеся от батарей, т.е. это должны быть компактные, недорогие, с низким энергопотреблением чипы. Кроме того, эти устройства должны работать в частотном диапазоне, доступном во всем мире. Последнее требование — дизайн должен отражать возможность комбинированного использования данных и речи, т.е. технология Bluetooth должна поддерживать мультимедийные данные.

В качестве среды передачи выбран нелицензируемый во всем мире диапазон 2,4 ГГц. Для передачи выбрана дуплексная схема перескока частот с временным разделением, скорость переключения частот равна 1600 раз в секунду. Время между двумя перескоками частоты называется слотом, продолжительность слота составляет 625 мкс; таким образом, каждый слот использует другую частоту. В странах, где доступная полоса составляет не менее 80 МГц (США, большая часть Европы), Bluetoth использует 79 поднесущих с шагом 1 МГц. В Японии, Франции и Испании допустимо использование только 23 частот. В среднем последовательность перескоков частот "посещает" каждую поднесущую с равной вероятностью. Все устройства, использующие одинаковую последовательность чередования частот, называются **пикосетью**.

При излучаемой мощности передатчиков Bluetooth до 100 мВт устанавливается связь в пределах от 10 до 100 метров. Так как основными устройствами Bluetooth являются устройства, питающиеся от батарей, предусмотрено несколько режимов работы с пониженным энергопотреблением (рис. 7.30). Устройства, в данный момент не имеющие соединений с другими (но и не отключенные), находятся в **режиме ожидания**. В этом режиме устройства периодически, через каждые 2 048 слотов (1,28 с), прослушивают канал на предмет наличия сигналов персонального вызова. Устройство ищет подмножество частот перескока, так называемые несущие запуска. Это подмножество из 32 (или 16, в указанных выше регионах) несущих запуска выбирается псевдослучайным образом на основе индивидуальных характеристик устройства. Как определяется несущими запуска, устройство посещает каждую частоту из несущих запуска, перемещаясь на следующую частоту через каждые 1,28 с. (Посещение несущей означает прослушивание этой частоты в течение 18 слотов или 11,25 мс.) Прослушивание одной несущей запуска позволяет устройству скоррелировать поступающий сигнал (если таковой имеется) с кодом доступа к пикосети, выведенным из ее индивидуальных характеристик. Если коррелятор активируется данным сигналом, активируется и собственно устройство.

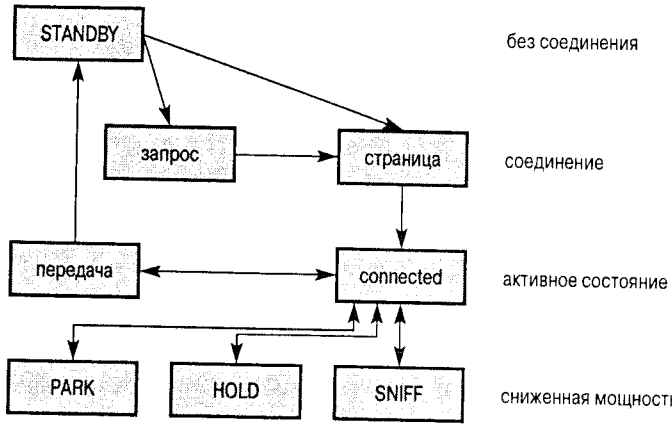

Рис. 7.30. Состояния устройства Bluetooth

Инициировать соединение (а следовательно, и становление пикосети) может любое устройство Bluetooth, которое становится **ведущим** (master). Это может производиться посредством передачи **сообщения персонального вызова**, если устройство знает адрес получателя, или **запросного сообщения**, за которым следует сообщение персонального вызова, если нужный адрес неизвестен. Ведущее устройство отсылает сообщение персонального вызова на 16 различных поднесущих, определенных для вызываемого устройства, которое теперь по определению является **подчиненным** (slave). Если искомое устройство не ответило, ведущее устройство использует для вызова оставшиеся 16 из 32 подчастот. Максимальная задержка такого метода вызова подчиненного устройства составляет 2,56 с, а среднестатистическая — 0,64 с.

Для увеличения срока службы батарей питания устройства могут находиться в одном из трех режимов пониженного энергопотребления.

- **Режим парковки**. В этом состоянии устройство имеет кратчайший рабочий цикл, следовательно, потребляет минимум энергии. Устройство остается синхронизированным в пикосети, но освобождает свой MAC-адрес, изредка прослушивая передачи ведущего устройства для выявления широковещательных сообщений.
- **Режим приостановки**. В этом состоянии устройство остается в пикосети, сохраняя свой MAC-адрес. Возможность отправления своей информации появляется сразу после выхода из режима приостановки.
- **Режим анализа**. Из трех режимов пониженного энергопотребления имеет наибольшее потребление. В этом состоянии устройство прослушивает среду с периодичностью, устанавливаемой программно в зависимости от характера использования устройства.

7.5.3. Уровень MAC

Управление доступом к среде в системе Bluetooth осуществляется за счет нескольких механизмов. Прежде всего в пикосети существует выделенное ведущее устройство, остальные устройства (до семи) являются подчиненными. В обязанности ведущего устройства входит синхронизация сети, т.е. установление последовательности переключения несущих, используя уникальный идентификатор устройства и фазу последовательности. Синхронизация в сети обеспечивается тактовыми генераторами устройств, имеющими фиксированную настройку. В каждой пикосети применяется своя последовательность перестройки частоты, что позволяет использовать коллективный доступ к каналу CDMA, не создавая взаимных помех. В пределах одной пикосети доступ к среде контролирует ведущее устройство, использующее механизм опроса и резервирования.

Все устройства Bluetooth равноправны, и любое может быть ведущим и подчиненным. Таким образом, различия между терминалами и базовыми станциями не существует, и два или более устройств могут образовать пикосеть. Устройство, инициировавшее соединение, становится ведущим, а остальные — подчиненными, но в пределах одной пикосети может быть только одно ведущее устройство.

7.5.3.1. Службы

Bluetooth предлагает два типа служб: синхронный канал связи, ориентированный на установление соединения, и асинхронный канал связи без установления соединения.

- **Синхронный канал связи, ориентированный на установление соединения (SCO)**. Стандартные телефонные соединения требуют симметричных двухточечных соединений с коммутацией каналов. Для соединений такого типа ведущее устройство резервирует два последовательных временных интервала, располагающихся через определенные фиксированные промежутки, — слот передачи и слот приема.
- **Асинхронный канал связи без установления соединения (ACL)**. Типичные приложения передачи данных требуют симметричных или асимметричных (например, информационный объем в WWW) многоточечных (в том числе широковещательных) сценариев передачи с коммутацией пакетов. В этом случае ведущее устройство использует схему опроса.

Одновременно система Bluetooth может поддерживать один канал ACL, три канала SCO или один канал ACL и одни канал SCO. Каналы SCO всегда поддерживают синхронные соединения со скоростью 64 Кбит/с. Каналы ACL могут быть симметричными или асимметричными, поддерживая различные скорости передачи в зависимости от типов пакетов, как объясняется в следующем разделе. Скорость передачи данных в симметричном канале связи с использованием пяти последовательных слотов и незащищенных данных достигает 432,6 Кбит/с. Асимметричные каналы могут переносить до 721 Кбит/с в одном направлении и 57,6 Кбит/с в обратном также при использовании пяти последовательных слотов и незащищенных данных.

7.5.3.2. Формат пакета

Структура пакета, передаваемого в одном слоте, представлена на рис. 7.31. Каждый пакет данных начинается с 72-битового поля — уникального кода доступа, выведенного из индивидуальных характеристик ведущего устройства, уникальных для этого канала.

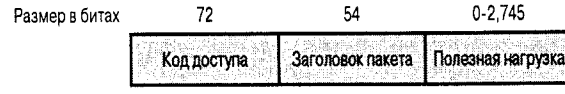

Рис. 7.31. Формат пакета Bluetooth, передаваемого в одном слоте

Устройства Bluetooth при приеме пакета в первую очередь проверяют код доступа. Если он не соответствует коду пикосети, пакет игнорируется. Более того, код доступа — это средство синхронизации. После этого поля следует поле заголовка пакета (см. рис. 7.32) и от 0 до 2 745 бит полезной нагрузки.

Рис. 7.32. Формат заголовка пакета Bluetooth

Заголовок пакета имеет следующую структуру. Первые три бита — MAC-адрес устройства в пределах пикосети, а значит, максимальное количество устройств, входящих в пикосеть, со-

ставляет восемь — одно ведущее и до семи подчиненных. С помощью следующего поля, занимающего 4 бита, может быть определено 16 типов пакетов. Далее идет поле управления потоком данных, ARQN — бит подтверждения приема, SEQN — порядковый номер для упорядочения пакетов. Последнее поле — 8-битовая контрольная сумма заголовка HEC. Для обеспечения надежности передачи заголовка дополнительно используется схема прямой коррекции ошибок со степенью кодирования 1/3. Таким образом, для передачи 18 бит заголовка требуется 54 бита.

Как правило, пакет данных занимает точно один слот. Для увеличения скорости пакет данных может занимать три или пять последовательных слотов (так называемые **многослотовые** пакеты), причем передача проходит без переключения, на одной несущей. При этом последовательность частотных скачков не прерывается, а остается такой же, как при передаче обычного пакета за один слот. Такой режим передачи пакета необходим для поддержания синхронизации в пикосети, так как все устройства пикосети работают на одной последовательности частотных скачков, устанавливаемой ведущим устройством.

При использовании канала **SCO** для передачи речи на скорости 64 Кбит/с применяется три типа однослотовых пакетов — без прямого исправления ошибок (FEC), с применением схемы 1/3 FEC и схемы 2/3 FEC. Выбор метода кодирования зависит от частоты появления ошибок в канале связи. Наиболее надежным, но и самым избыточным является метод 1/3 FEC — объем передаваемых данных увеличивается в три раза, но существенно уменьшается вероятность повторной передачи пакета. Для кодирования голоса используется **непрерывная дельта-модуляция с переменной огибающей (CVSD)**, высоконадежный метод кодирования, исключающий повторную передачу сообщений [Haartsen, 1998].

Асинхронный канал **ACL** предусматривает многослотовую передачу с выделением 3 или 5 слотов на пакет либо передачу однослотовых пакетов. Для повышения надежности передачи в каналах с большой частотой появления ошибок может применяться схема 2/3 FEC. Для каналов с низким уровнем ошибок в Bluetooth применяется схема автоматического запроса повторной передачи ARQ, предусматривающая передачу пакета подтверждения в следующем временном слоте после получения информации. При потере пакета отправитель повторяет передачу немедленно в следующем временном интервале после получения отрицательного подтверждения.

7.5.4. Организация сетей

Все устройства пикосети используют одну последовательность перестройки частот; таким образом, совместно используют один канал шириной 1 МГц. Увеличение числа активных устройств ведет к уменьшению удельной (по отношению к одному пользователю) пропускной способности сети; наличие одной сети, занимающей всю полосу 80 Мгц, не очень эффективно. Возникает идея создать группу пикосетей, использующих разные последовательности чередования несущих и имеющих возможность работать одновременно.

На рис. 7.33 приведен пример распределенной сети, состоящей из трех пикосетей, причем изображены устройства, принадлежащие одновременно двум разным пикосетям. Множество пикосетей использует один канал связи шириной 80 МГц, но применяют индивидуальные последовательности перескоков частот. При этом не исключена возможность конфликтов при совпадении несущих в разных пикосетях. Увеличение числа пикосетей приводит к росту вероятности ошибок в общей сети и увеличению конфликтов в отдельной пикосети.

Если устройство желает участвовать в нескольких пикосетях, оно должно синхронизировать свою последовательность скачков частот с последовательностью сети, в которую желает войти. Для этого устройство должно знать ведущее устройство и последовательность перескоков частот новой сети. Кроме того, перед уходом из старой сети устройство должно проинформировать ведущее устройство старой сети, что данное устройство будет недоступно определенное время. Остальные устройства старой и новой сети переход не затрагивает.

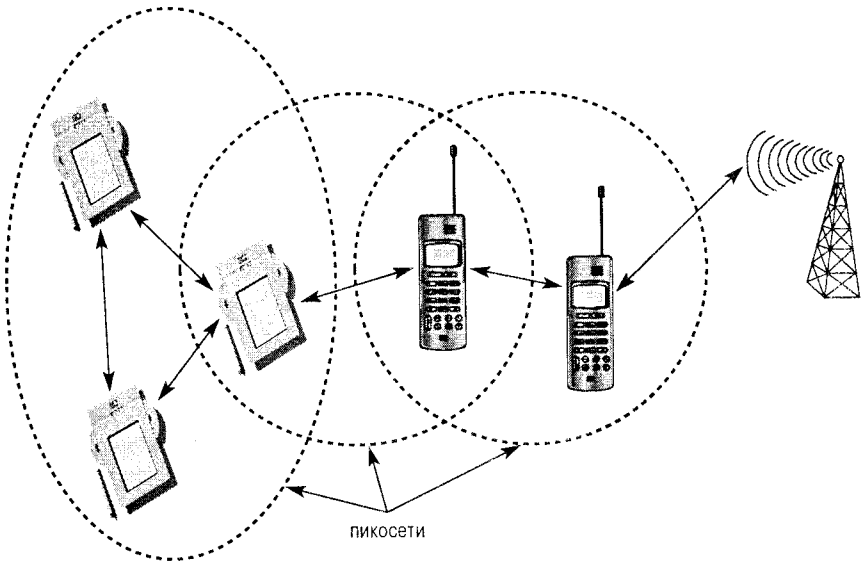

Рис. 7.33. Протокольная архитектура и мосты IEEE 802.11

Ведущее устройство также может покинуть текущую сеть и перейти в другую, где оно обязательно будет подчиненным (ведущее устройство одной пикосети не может стать ведущим другой, поскольку в этом случае поведение обоих сетей будет идентичным — последовательность перескоков частот определяется ведущим устройством). Как только ведущее устройство покидает сеть, информационный обмен приостанавливается до его возвращения.

Таким образом, связь пикосетей достигается посредством устройств, периодически переходящих из одной сети в другую и обратно, что обеспечивает изохронный поток данных.

7.5.5. Безопасность

Особенностью радиоинтерфейса является крайне простой доступ. Для пользователей Bluetooth, часто оперирующих с конфиденциальной информацией (например, регламент важного совещания, передаваемый с мобильного телефона на PDA, и т. д.), важно обеспечение секретности данных. Просто представим ситуацию, когда два PDA, находясь в кейсах на одном багажном конвейере аэропорта, обмениваются важной информацией! Для обеспечения информационной безопасности в Bluetooth на уровне MAC применяются механизмы аутентификации и шифрования данных.

Функции безопасности, предлагаемые Bluetooth, включают аутентификацию по схеме запрос/отклик, потоковое шифрование и генерацию ключа сеанса. Каждое конкретное соединение может требовать одно- или двухсторонней аутентификации или не требовать аутентификации вообще. Для шифрования могут использоваться 0-, 40- или 64-битовые ключи, что дает относительно слабую безопасность. В то же время все названные схемы должны реализовываться аппаратно, а более сильное шифрование (при необходимости) обеспечат вышестоящие уровни. Уровень защиты, предлагаемый Bluetooth, служит только для установления определенного доверия между устройствами локального домена.

Алгоритмы безопасности используют открытую идентификацию устройств, секретный частный ключ пользователя и генерируемый внешне случайный ключ. Для каждой транзакции генерируется новое случайное число, а управление ключами осуществляется на более высоких уровнях.

7.5.6. Управление каналом

Система Bluetooth управляется **администратором связей** (link manager — LM). Администратор связей занимается аутентификацией, согласовывает параметры аутентификации, осуществляет установление и завершение соединения, конфигурирует связи и т.п. Администраторы различных устройств сообщаются с использованием **протокола управления LMP**, кроме того, все администраторы используют службы, предоставляемые контроллером канала связи (LC).

Контроллер канала связи предлагает базовые службы передачи и приема данных и запроса имен других устройств. Более того, он настраивает соединения, инициирует режимы аутентификации и согласования параметров. Контроллер может также переводить устройство в режим прослушивания, когда устройство только прослушивает каждый n-й слот; значение величины n согласовывает администратор связей. Подобным образом контроллер может перевести устройство в режим приостановки или парковки, как рассказывается в разделе 7.5.2.

7.6. Резюме

В данной главе вводятся три различные технологии, разработанные для беспроводных ЛВС. Основной целью перехода к беспроводным технологиям является обеспечение большей гибкости узлов сети. Все описанные беспроводные ЛВС имеют недостатки (радиоинтерфейс и большая сложность по сравнению с проводными аналогами), но они дают новую степень свободы пользователям, находящимся в помещениях, зданиях или производственных цехах. В настоящее время беспроводные ЛВС уже используются на складах, в аудиториях, конференц-залах, больницах и многих других местах.

Представленные в главе технологии имеют различия. Самые первые разработки в области беспроводных ЛВС были частными решениями, теперь же большинство их поддерживает стандарт IEEE 802.11. Стоит отметить, что ко времени утверждения стандарта на рынке беспроводных ЛВС появились разработки, обогнавшие стандарт по многим параметрам.

Для HIPERLAN ситуация иная. Авторы стандарта (институт ETSI) разработали совершенно новый стандарт, для которого на настоящий момент еще нет доступных продуктов. Стандарт HIPERLAN объединяет много интересных свойств и характеристик, является более мощным, чем IEEE 802.11, поддерживает большую скорость передачи данных (23,5 Мбит/с против 2 Мбит/с), но еще неизвестно, насколько коммерчески успешным он окажется. Кроме того, выбранный для HIPERLAN частотный диапазон 5 ГГц не так доступен во всем мире, как полоса ISM 2,4 ГГц, используемая IEEE 802.11. Хороший обзор возможностей схем MAC стандартов IEEE 802.11 и HIPERLAN представлен в работе [Anastasi, 1998]. Исследованы также принципы обеспечения качества обслуживания (QoS), требуемого для беспроводной технологии ATM (см. главу 8).

Совершенно другая ситуация сложилась при разработке Bluetooth. Группа компаний, занимающаяся разработкой, не задавалась целью создания фактического промышленного стандарта. До настоящего момента ни одна организация по стандартизации не издала ни одной спецификации, относящейся к Bluetooth; в то же время первые продукты с этим решением должны были выйти к концу 1999. Основной целью Bluetooth является не комплексный стандарт, охватывающий многие аспекты организации беспроводных сетей, а быстрое и очень дешевое решение, позволяющее эпизодическую связь персональных компьютеров на близком расстоянии с использованием диапазона ISM 2,4 ГГц.

Типичные мобильные устройства будущего будут объединять несколько технологий с возможностью подключения к сетям различного типа, например, сопровождающим службам GSM (такие как UMTS) для связи на больших расстояниях, спутниковым антеннам (например, для связи с системой Iridium) и различным устройствам беспроводных ЛВС (например, IEEE 802.11 и Bluetooth). В зависимости от стоимости услуг, сферы применения и расположения устройство будет автоматически выбирать оптимальное устройство связи и сеть.

Обеспечение межсетевого роуминга (в частности, приемлемого качества обслуживания) остается пока открытым вопросом, который должен решаться на более высоких уровнях.

7.7. Упражнения

1. Как ограничена мобильность клиентов беспроводных локальных сетей? Какие дополнительные элементы необходимы для обеспечения межсетевого роуминга? Как обеспечивается роуминг в беспроводных локальных сетях? В ответах обратите внимание на функции уровня 2.
2. В чем различия и общие черты беспроводных глобальных сетей и беспроводных локальных сетей? Рассмотрите режимы работы, управление, используемые частотные диапазоны, возможности и услуги, национальное и межгосударственное регулирование.
3. С какими проблемами приходится сталкиваться при обеспечении безопасной работы сетей? Какой уровень защиты обеспечивается в беспроводных локальных сетях и что дополнительно можно предпринимать в этом отношении?
4. Сравните технологии IEEE 802.11, HIPERLAN и Bluetooth с точки зрения возможности организации эпизодической сети. В чем принципиальные отличия?
5. В чем причина сохранения локальных сетей в ИК-диапазоне в условиях коммерческого успеха Bluetooth на рынке беспроводных локальных сетей малого радиуса охвата?
6. Охарактеризуйте способы беспроводной передачи на физическом уровне IEEE 802.11 и HIPERLAN.
7. Сравните механизмы обеспечения энергосбережения сетевых устройств. Как влияет поддержка режима энергосбережения на качество обслуживания (Quality of Service — QoS) сети? Насколько совместимы режим энергосбережения и высокое QoS?
8. Сравните качество обслуживания (QoS) трех беспроводных эпизодических сетей. В чем преимущества сетей с инфраструктурой с точки зрения QoS? Опишите характеристики QoS в Bluetooth. Обеспечивает ли одна из трех технологий гарантированное качество QoS?
9. Как решается проблема скрытых узлов в IEEE 802.11 и Bluetooth?
10. Опишите механизмы доступа к каналу передачи для узлов в 802.11 и HIPERLAN. Как способ энергосбережения отражается на времени ожидания доступа?
11. Какие решения и методы применяются для повышения надежности передачи данных в сетях?
12. По каким причинам возникают конфликты при передаче данных в рассматриваемых сетях? В чем различие конфликтов на уровне MAC и физическом уровне? Какие способы применяются для предотвращения или уменьшения вероятности возникновения конфликтов в трех сетях?
13. Сравните избыточность схем доступа к среде передачи (объемы служебных издержек). Как связана вероятность возникновения конфликтов с загруженностью среды? Как решается проблема возрастания вероятности конфликтов при увеличении числа клиентов сети? Каковы масштабы расширения сетей по количеству клиентов?
14. Как обеспечивается роуминг, как режим роуминга отражается в изменении топологии сети? Чем отличается обеспечение роуминга в эпизодических сетях и сетях с инфраструктурой?
15. В чем преимущества и сложности реализации механизма ретрансляции в эпизодических сетях с учетом безопасности, энергосбережения и устойчивости сети? Сравните Bluetooth и HIPERLAN.

7.8. Рекомендуемая литература

Anastasu, G.; Lenzini, L.; Mingozzi, E.; Hettich, A.; Krämling, A. (1998) "MAC protocols for wide-band wireless local access: evolution toward wireless ATM", *IEEE Personal Communications*, 5, (5). October.

Bisdikian, C.; Bhagwat, P.; Gaucher, B. P.; Janniello, F. J.; Naghshinen, M.; Pandoh, P.; Korpeoglu, I. (1998): "WiSAP A wireless personal access network for handheld computing devices", *IEEE Personal Communications*, 5, (6).

The Bluetooth Consortium, (1998) `http://www.bluetooth.com`.

Chhaya, H. S.; Gupta, S. (1996) "Performance of asynchronous data transfer methods of IEEE 802.11 MAC protocol", *IEEE Personal Communications*, 3 (5).

Chhaya, H. S.; Gupta, S. (1997) "Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocol", *Wireless Networks* 3 (1997), Baltzer Science Publishers.

ETSI (1996) *Radio Equipment and Systems (RES); High Performance Radio Local Area Network (HIPERLAN) Type 1; Functional specification*, European Telecommunication Standard, ETS 300 652, European Telecommunications Standards Institute.

ETSI (1998) *Broadband Radio Access Networks (BRAN); High Performance Radio Local Area Network (HIPERLAN) Type 1; Functional specification*, EN 300 652 v1.2.1, European Telecommunications Standards Institute.

Haartsen, J. (1998 "Bluetooth — the universal radio interface for ad hoc, wireless connectivity", *Ericsson Review* No. 3, `http://www.ericsson.com`.

Halsall, F. (1996) *Data communications, computer networks and open systems*. Addison Wesley Longman.

IEEE (1990) *Local Area Network and Metropolitan Area Network — Overview and Architecture*, The Institute of Electrical and Electronics Engineers, IEEE 802.1a.

IEEE (1997) *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, The Institute of Electrical and Electronics Engineers, IEEE 802.11.

ISO (1995) *Information technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Common specifications — Part 1: Medium Access Control (MAC) service definition*. International Organization for Standardization, ISO 15802-1.

Perlman, R. (1992) *Interconnections: bridges and router*. Addison Wesley Longman.

Santamaria, A.; López-Hernández, F. (1994) *Wireless LAN systems*, Artech House.

Wesel, E. K. (1998) *Wireless multimedia communications*. Addison Wesley Longman.

Woesner, H.; Ebert, J.-P.; Schläger, M.; Wolisz, A. (1998) "Power-saving mechanisms in emerging standards for wireless LANs", *IEEE Personal Communications*, 5, (3).

Беспроводная ATM 8

В этой главе представлен обзор технологии **беспроводного асинхронного режима передачи (wireless ATM — WATM)**. WATM — это не только технология передачи, подобно ATM для стационарных сетей [Händel, 1994], но также попытка специфицировать систему связи полностью. Если многие аспекты систем, представленные в главах 7, 9, 10 и 11, берут свое начало в сфере систем связи, многие аспекты WATM пришли из телекоммуникационной индустрии. Эту специфическую ситуацию можно сравнить со случаем конкуренции и слияния идей TCP/IP и ATM.

В следующих разделах описаны некоторые аспекты WATM и требуемые, в связи с мобильностью, дополнения к стационарным сетям ATM. Прежде всего, в эталонной модели WATM показаны некоторые объекты, родственные ATM, также делается акцент на необходимых дополнениях. В мобильных терминалах должны применяться многие новые функции, и обсуждению этой темы посвящен отдельный раздел. Несколько новых механизмов обсуждает рабочая группа WATM форума ATM Forum. Для беспроводных сетей, поддерживающих мобильность пользователя, фундаментальными являются понятия переключения, маршрутизации, управление локализацией и т.д. Особенно интересным в данном контексте представляется фактор "мобильного качества обслуживания", поскольку именно здесь подчеркивается основное различие между WATM и другими технологиями, которые только поддерживают наилучшую работу.

Очень важно отметить, что до настоящего времени WATM не является стандартом, в котором определены все обсуждаемые здесь аспекты. Спецификация в области WATM состоит из множества предложений (например, служба ATM с поддержкой мобильности AMES [Bhat, 1998a]), обсуждений [Rauhala, 1998б], полевых испытаний [Ayanoglu, 1996]. Поскольку не существует стандарта, охватывающего все эти аспекты, в следующих разделах предпринята попытка отразить некоторые предложения, выдвинутые исследователями и компаниями, связанными с рабочей группой WATM форума ATM Forum, и существующее на данный момент положение дел относительно ведущихся дискуссий и существующих проектов. Эти предложения позволяют хорошо представить существующие на данный момент проблемы, кроме того, предлагаются какие-то решения, однако, эти решения не стоит рассматривать как единственно возможные, поскольку они представляют собой лишь отпечаток текущей работы. Сравнение некоторых документов из разных источников может вызвать определенное замешательство, поскольку названия объектов и сообщений протокола WATM в настоящее время еще не стандартизованы. Дабы не вносить еще большей путаницы, в главу включены не все подробности проводимых дискуссий.

8.1. Мотивы разработки WATM

Существует несколько причин, приведших к разработке WATM.

- Основной причиной разработки WATM является необходимость прямой интеграции беспроводных терминалов в сеть ATM. Эта интеграция является основным требованием для поддержания интегрированных служб и различных типов потоков информации, подобных поддерживаемым технологией ATM в стационарных сетях.
- Сети ATM хорошо масштабируются от локальных до глобальных, а мобильность нужна как в локальных, так и в глобальных сетях. Таким образом, требуется стратегия расширения ATM на беспроводный доступ в локальной и глобальной конфигурациях.
- Для успешного развития, технология ATM должна расширяться на беспроводную связь. В противном случае она не сможет применяться в быстро развивающейся области мобильной связи.
- Технология WATM может обеспечивать качество обслуживания (quality of service — QoS) для адекватной поддержки потоков мультимедийных данных. Многие другие беспроводные технологии (например, сети IEEE 802.11) обычно только предлагают службы без обязательств или несколько расширенные срочные службы. Однако, эти службы не обеспечивают такое количество параметров QoS, которое дают сети ATM.
- Для поставщиков телекоммуникационных услуг естественным кажется, что слияние технологий беспроводной мобильной связи и ATM даст беспроводную ATM. В таком контексте одной из целей является прямая интеграция мобильности в сеть B-ISDN, которая уже использует ATM в качестве технологии передачи.

При рассмотрении всех этих аспектов ATM становится ясно, что эта система будет более сложной, чем большинство других беспроводных систем. В то время как, например, IEEE 802.11 охватывает только методы доступа к локальным областям, Bluetooth наращивает пикосети, а Mobile IP работает только на сетевом уровне, WATM — это попытка построить комплексную систему, охватывающую физический уровень, доступ к среде, маршрутизацию, интеграцию в стационарную сеть ATM, интеграцию служб в B-ISDN и т.д.

8.2. Рабочая группа беспроводной ATM

Для разработки описанной довольно сложной системы форум ATM Forum в 1996 году сформировал **Рабочую группу беспроводной ATM** (Wireless ATM Working Group). Эта группа занимается разработкой набора спецификаций, которые распространят использование технологии ATM на беспроводные сети. Беспроводные сети могут охватывать много различных сценариев работы с сетью, таких как частный и общий, локальный и глобальный, мобильный и беспроводный доступ [Raychaudhuri, 1996].

Основная цель этой рабочей группы — обеспечить совместимость всех новых предложений с существующими стандартами форума ATM Forum. Таким образом, становится возможным усовершенствование существующих сетей ATM, то есть, коммутаторов ATM и оконечных систем ATM, с помощью определенных функций, позволяющих, если это требуется, поддерживать мобильность и радиосвязь с абонентами. В таком контексте было определено две основных группы открытых вопросов: с одной стороны, это расширение, которое требуется "стационарной" ATM для поддержки мобильности, с другой стороны, это все протоколы и механизмы, связанные с радиосвязью между абонентами.

Для **мобильной ATM** необходимо рассмотреть следующие дополнительные расширения системы ATM.

- **Управление локализацией**. Подобно другим сотовым сетям, сети WATM должны быть в состоянии определить местоположение беспроводного терминала или мобильного пользователя, то есть, найти текущую точку доступа терминала к сети.
- **Мобильная маршрутизация**. Даже если системе известно местоположение терминала, она все же еще должна маршрутизировать транспортный поток по сети к точке доступа, в данный момент отвечающей за беспроводный терминал. Всякий раз, когда пользователь перемещается в новую точку доступа, система должна изменить маршрут транспортного потока.
- **Передачи сигналов с переключением**. Сеть должна предоставлять механизмы поиска новых точек доступа, позволять установление новых соединений между промежуточными системами и сигнализировать о реальных изменениях точки доступа.
- **Управление QoS и информационным обменом**. В отличие от беспроводных сетей, предлагающих только информационный обмен без обязательств, и сотовых сетей, предлагающих лишь несколько различных типов транспортного потока, WATM должна предлагать большее количество параметров QoS. Для эксплуатации этих параметров все действия, такие как изменение маршрута, переключение и т.д., должны быть управляемыми. Кроме того, сеть должна обращать внимание на входящий транспортный поток (и проверять, согласуется ли он с каким-либо контрактом на транспортный поток) аналогично тому, как это происходит в существующих на данный момент ATM.

Для обеспечения беспроводного доступа рабочая группа в настоящее время рассматривает нижеследующие вопросы, связанные с **уровнем радиодоступа** (radio access layer — RAL).

- **Управление ресурсами радиосвязи**. Как и для других беспроводных сетей, в данном случае должны быть определены радиочастоты, схемы модуляции, антенны, канальное кодирование и т.д.
- **Беспроводный доступ к среде**. Возможны различные схемы доступа к среде, каждая из которых имеет свои специфические сильные и слабые стороны, например, для мультимедийных и голосовых приложений. Можно представить различные схемы централизованного или распределенного доступа, работающие в ячейках ATM.
- **Беспроводное управление каналом передачи данных**. Уровень управления каналом передачи данных может предложить сжатие заголовка для ячейки ATM, который представляет собой почти 10% (5 байт) служебных издержек для ячейки 53 байт. Кроме того, этот уровень может для улучшения надежности применить схемы ARQ или FEC.
- **Вопросы переключения**. Во время переключения ячейки могут не только теряться, но и размещаться в неверном порядке (в зависимости от механизмов переключения). Таким образом, ячейки должны быть собраны в новую последовательность и, если требуется, должны быть повторно переданы потерянные ячейки.

8.3. Службы WATM

В следующий раздел включено несколько примеров сценариев, в которых может использоваться технология WATM. Эти примеры показывают, что принцип работы WATM — это не просто предоставление беспроводного доступа или конструирование беспроводной ЛВС. Предлагаемые службы охватывают многие аспекты современной беспроводной и мобильной связи.

Системы WATM предназначены для передачи голоса, классических данных, видео- (от низкокачественного до профессионального), мультимедийных данных, коротких сообщений и т.д. Можно определить несколько **сценариев служб** [Rauhala, 1998a], [Barton, 1998], подобных перечисленным ниже.

- **Офисные среды**. Этот вариант включает в себя все типы расширений существующих стационарных сетей, предлагающий широкий диапазон доступа к Internet/Intranet, проведение мультимедийных конференций, доступ к мультимедийным базам данных в реальном времени и дистанционное присутствие. Используя технологию WATM, офис может фактически расширяться на реальное местоположение его сотрудников.
- **Университеты, школы, центры обучения**. В этом сценарии акцент делается на дистанционном обучении, беспроводном и мобильном доступе к базам данных, доступе к Internet или обучении в сфере мобильной обработки мультимедийных данных.
- **Промышленность**. WATM может предложить расширение Intranet, поддерживающее соединение с базами данных, информационный поиск, наблюдение, а также передачу данных в реальном времени и управление предприятием.
- **Больницы**. Благодаря качеству обслуживания при передаче данных технология WATM — это первейший кандидат для создания надежных мобильных и беспроводных сетей с высокой пропускной способностью. В число практических приложений может входить передача медицинских изображений, дистанционный доступ к регистрационным документам пациентов, дистанционное наблюдение за пациентами, дистанционная диагностика на дому или в машине скорой помощи, а также телемедицину. В последнем случае требуются высоконадежные сети с гарантированным качеством обслуживания, допускающим, например, дистанционную хирургию.
- **Домашние среды**. Многие бытовые электронные устройства (например, теле-, радиооборудование, проигрыватели компакт-дисков, ПК с доступом к Internet) можно связать, используя технологию WATM. Здесь WATM позволяет осуществлять различные беспроводные соединения, например, связать персональное информационное устройство с телевизором.
- **Объединенные в сеть транспортные средства**. Все транспортные средства, используемые для транспортировки людей или вещей, в будущем будут иметь доступ к сети и составлять локальные сети. Если в настоящее время транспортные средства, такие как грузовики, самолеты, автобусы или автомобили, имеют весьма ограниченные возможности связи (например, посредством GSM), технология WATM могла бы предоставить им высококачественный доступ к Internet, базам данных компаний, дать возможность проведения мультимедийных конференций и т.д. Существует и другое применение WATM — уровни локальных сетей транспортных средств определенной области с все возрастающей важностью, предназначенные, например, для предотвращения аварий или увеличения пропускной способности дорог путем формирования на дороге поезда из автомобилей или грузовиков с очень малым безопасным расстоянием между отдельными транспортными средствами.

Мобильность внутри сети ATM будет обеспечиваться **службой ATM с поддержкой мобильности (AMES)**. AMES будет способствовать использованию этих сетей ATM различным оборудованием и приложениями, требующими мобильности. Беспроводное оборудование должно получать от сети равноценное (по сравнению с проводными терминалами) обслуживание. AMES включает расширения, необходимые для поддержки мобильности терминалов домашнего и промышленного использования. Пользователи могут переупорядочивать устройства, не теряя доступ к сети ATM и сохраняя гарантированное качество обслуживания.

Кроме того, WATM будет предлагать **службу доступа** для персональных сотовых систем (PCS). Персональные сотовые системы, такие как GSM, IS-95, UMTS и т.п., могут использовать возможности поддержки мобильности стационарной сети ATM для маршрутизации потока информации к соответствующему контроллеру базовой станции. Общественными службами для пользователей могут быть служба мультимедийной телефонии, симметричная служба, предлагающая передачу речи и видеоизображения при низкой скорости передачи данных со

средней мобильностью, а также асимметричная служба передачи данных в режиме реального времени, например, просмотр Web-ресурсов, использование электронной почты и загрузка файлов. Частные службы могут включать мультимедийный беспроводной телефон с более высоким качеством, чем общественные версии. Кроме того, на территории учебных заведений могут развертываться специальные частные службы по передаче данных.

Другая сфера услуг предоставляется **спутниковыми службами ATM (SATM)**. В будущем спутники будут предлагать большое количество разнообразных телевизионных услуг, мультимедийных служб, служб интерактивного видео, Internet, телефонии и других (см. главу 5.). Основным преимуществом в данном контексте является повсеместное покрытие глобальной сетью в отдаленных, сельских и городских районах. Спутники могут использоваться непосредственно (прямой доступ пользователя к службе), например, посредством мобильного телефона или терминала с антенной, которые предоставляют пользователю возможность непосредственного доступа к сети ATM. Кроме того, вся сеть может быть связана со спутником через мобильный коммутатор (служба стационарного доступа). Например, все компьютеры в отдаленных школах могут быть связаны с коммутатором, который связывается со спутником. Даже морские судна могут нести сети ATM и использовать прямую интеграцию своей бортовой сети ATM с глобальной сетью ATM (служба на мобильной платформе).

8.4. Эталонная модель

По сравнению со стационарной ATM, беспроводная ATM предоставляет несколько дополнительных функций. В следующих разделах представлены примеры эталонных моделей, которые иллюстрируют интеграцию беспроводных компонентов, точек доступа и расширений для поддержания мобильности в архитектуру ATM.

8.4.1. Примеры конфигураций

WATM поддерживает различные типы беспроводного доступа, такие как беспроводная местная линия связи или беспроводные мобильные конечные пользователи. Поскольку в этой книге основное внимание акцентируется на мобильной связи, ниже рассматриваются только протоколы и архитектуры, поддерживающие мобильность и беспроводный доступ.

На рис. 8.1 показана упрощенная эталонная модель сети ATM, поддерживающей мобильность и беспроводный доступ [Bhat, 1998a]. **Беспроводные мобильные терминалы ATM (WMT)** обладают каналами радиосвязи с **радиотрансиверами** (RT). Эти радиотрансиверы связаны с **точкой доступа** (AP) проводным образом. Одна точка доступа может контролировать несколько радиотрансиверов, она реализовывает все функции радиодоступа и, таким образом, осуществляет управление радиоресурсами, доступом к радиосреде и выполняет другие специфические функции радиосвязи. Между AP и WMT происходит передача информации, связанной с радиодоступом. Несколько точек доступа связаны с **коммутатором ATM (объектом), поддерживающим мобильность конечного пользователя** (EMAS-E). Объекты EMAS-E на уровне ATM обмениваются сообщениями с терминалами WMT и предлагают дополнительные функции, связанные с мобильностью, как описано в следующих разделах. В архитектуре WATM разделены передача радиосигналов между точкой доступа и WTM и передача сигналов ATM между EMAS-E и WMT. Чтобы скоординировать эти два уровня передачи сигналов используется **протокол управления точкой доступа (access point control protocol — APCP)**. Протокол APCP интегрирован в дополнение к протоколу UNI с поддержкой мобильности (UNI+M). Объекты EMAS-E сообщаются с другим **коммутатором ATM (сетью), поддерживающим мобильность конечного пользователя** (EMAS-N) посредством разновидности протокола NNI с поддержкой мобильности, NNI+M.

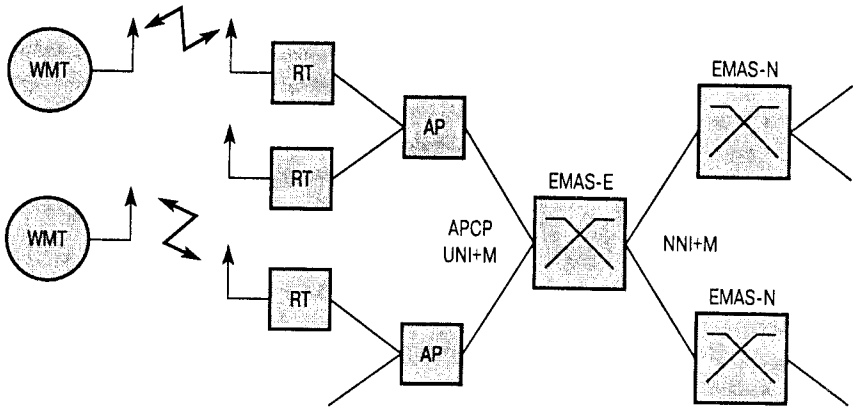

Рис. 8.1. Пример эталонной модели беспроводного доступа [Bhat, 1998a]

Эталонная модель, показанная на рис. 8.1, — это только одно из возможных решений. Функции точки доступа также могут быть интегрированы в EMAS-E. Интересно отметить, что в сценарии беспроводного доступа точка доступа выступает как виртуальный мультиплексор для потоков данных ATM. Точка доступа получает ячейки ATM (возможно модифицированные под радиопередачу) и уплотняет их в выходной канал связи (уплотнение ячеек в технологии ATM).

Приведенная здесь эталонная модель подобна используемой в сети GSM (см. раздел 4.1). Радиотрансиверы аналогичны базовой трансиверной станции (BTS), точки доступа реализуют функции, подобные функциям контроллеров базовой станции (BSC) и, наконец, EMAS-E действует подобно контроллеру мобильного коммутатора (MSC).

WATM не ограничивается базовым сценарием, показанным на рис. 8.1. Другие сценарии включают беспроводные эпизодические сети, мобильные коммутаторы (например, на борту самолета) или стационарный беспроводный доступ к местным линиям связи (см. рис. 8.20 и [Rauhala, 1998a]).

8.4.2. Общая эталонная модель

Ниже представлены предложения рабочей группы WATM [Raychaudhuri, 1996a], [Raychaudhuri, 1996b] и текущие проекты стеков протоколов беспроводной ATM [Bhat, 1998a]. Они иллюстрируют изменения, которые необходимо внести в стационарную сеть ATM, чтобы сделать ее пригодной для мобильного использования с поддержкой беспроводного доступа.

Для беспроводного доступа, мобильного доступа и, как показано на рис. 8.2, беспроводного мобильного доступа к сети ATM существует несколько эталонных моделей. Терминалы мобильной ATM (mobile ATM — MATM) для получения беспроводного доступа к точке доступа WATM используют терминальный адаптер WATM. Терминалы MATM могут представляться, например, портативными компьютерами, использующими адаптер ATM для беспроводного доступа плюс программное обеспечение для мобильности. Терминальный адаптер WATM может предоставлять беспроводный доступ, то есть, он включает в себя трансивер и прочее нужное оборудование, однако он не поддерживает мобильности. Точка доступа с радиотрансиверами связана с коммутатором ATM (объектом) с поддержкой мобильности (EMAS-E), который, в свою очередь, связывается с сетью ATM с коммутаторами (сетями) с поддержкой мобильности (EMAS-N), и другими стандартными коммутаторами ATM. Наконец, партнером по связи может стать проводная, не поддерживающая мобильность, конечная система ATM.

Радиосегмент простирается от терминала и терминального адаптера к точке доступа, тогда как стационарный сегмент сети простирается от точки доступа к стационарной конечной системе. Кроме

того, стационарная сеть с поддержкой мобильности, содержащая в себе все приспособленные под мобильность коммутаторы EMAS-E и EMAS-N, может отличаться от стандартной сети ATM с неприспособленными под мобильность коммутаторами и конечными системами.

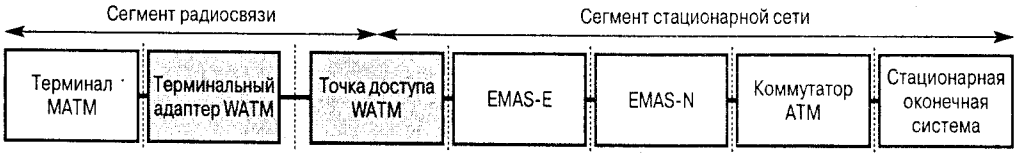

Рис. 8.2. Пример общей эталонной модели WATM

ATM отделяет пользователя от уровня управления, что отражено на протокольных уровнях эталонной модели (см. рис. 8.3 и рис. 8.4). Мобильный и беспроводный доступ незначительно влияет на пользовательский уровень. Поскольку пользовательские процессы не должны замечать беспроводный доступ, высшие протокольные уровни остаются одинаковыми как для беспроводной, так и для проводной передачи. Передача данных по радиоканалу, однако, может полностью отличаться от проводной передачи данных. **Уровень радиодоступа** (**RAL**) может выполнять сжатие заголовка для увеличения эффективности, использовать схему FEC для повышения надежности или вводить новую управляющую информацию в заголовок ячейки ATM. Уровень RAL включает уровни LLC, MAC и физический (см. раздел 8.6). В то же время, с пользовательской стороны терминального адаптера и с сетевой стороны точки доступа не видно никаких отличий от обычного потока ячеек ATM.

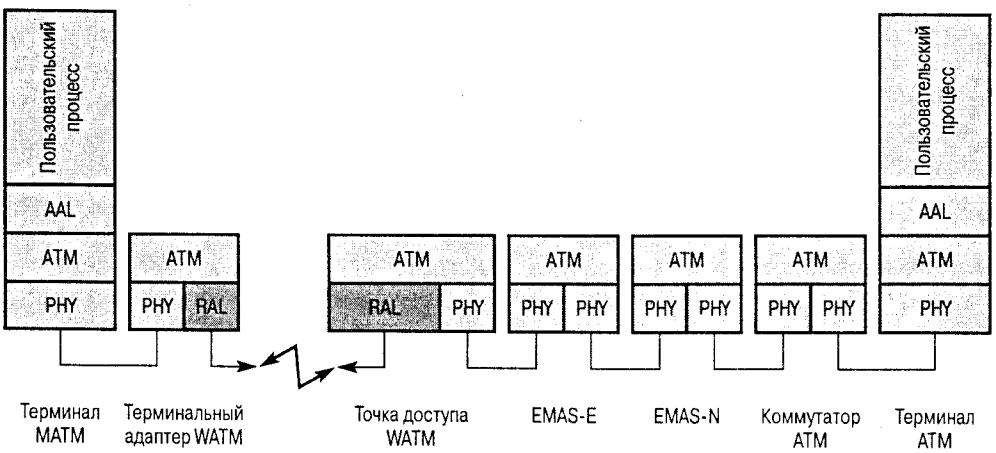

Рис. 8.3. Протокольные уровни пользовательской плоскости в эталонной модели WATM

Если на пользовательской плоскости требуется не слишком много новых свойств, плоскость управления, наоборот, интегрирует множество новых функций в различные объекты сети; как показано на рис. 8.4. MATM требуются дополнительные функции передачи сигналов для интерфейса "пользователь-сеть" (user-to-network interface — UNI) для поддержки мобильности. Результирующий интерфейс, предлагающий все традиционные функции UNI и поддерживающий новые, обозначается **UNI+M**. Для управления радиотрансивером в терминальный адаптер и точку доступа требуется внедрить дополнительные функции (**W-CTRL**). Эти функции могут контролировать интенсивность сигнала или другие специфические для радиопередачи параметры.

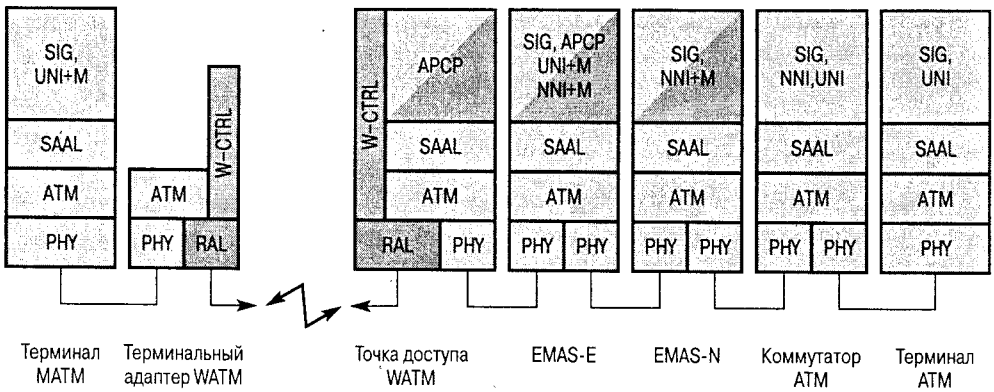

Рис. 8.4. Протокольные уровни плоскости управления в эталонной модели WATM

Для обмена управляющей информацией между точкой доступа и EMAS-E необходима реализация **APCP**. APCP используется только для информации, связанной с точками доступа, но не для каких-либо иных данных, связанных с ATM. Точка доступа не производит коммутации, так что плоскость управления EMAS-E реализовывает дополнение к интерфейсу UNI+M терминала и APCP точки доступа. Кроме того, она содержит межсетевой интерфейс NNI с функциями поддержки мобильности (**NNI+M**). Следующий коммутатор, EMAS-N, поддерживает мобильность с NNI+M, однако у него нет связи с точками доступа. Эти коммутаторы могут также связываться со стандартными коммутаторами ATM.

Стоит отметить, что еще не ясно, как много нужно знать стационарной конечной системе о мобильности. Однако в сети будут функции, позволяющие связь между существующими конечными системами ATM и мобильными терминалами.

8.5. Функции

Для поддержки мобильности, беспроводного доступа, безопасности и т.д., как с пользовательской, так и с сетевой стороны необходимо множество дополнительных функций. Примеры таких функций приводятся в следующих разделах. Следует, впрочем, заметить, что некоторые из представленных примеров еще только обсуждаются [Bhat, 1998a], и что этот список далеко не полон. На рис. 8.5 представлен обзор функций и их логических взаимосвязей. Модуль идентификации пользователя (UIM) является частью беспроводного мобильного терминала (WTM). Посредством радиоканала WTM связывается с точкой доступа (access point — AP), которая связана с коммутатором ATM, поддерживающим мобильность конечного пользователя на границе стационарной сети (EMAS-E).

8.5.1. Функции беспроводного мобильного терминала

Для поддержки администрирования, служб передачи данных, аутентификации и т.д. были определены следующие функции (индекс T указывает, что они относятся к терминалу).

- **Функция управления мобильностью (MMF_T).** Используется для анализа и наблюдения за сетью, инициации обновления информации о размещении и избирательного отклика.
- **Функция управления вызовами и соединениями (CCF_T).** Подобно стационарным сетям эта функция управляет настройкой вызова и разъединением. Таким образом, она также включает в себя управление доступом, управление соединением и запросы, связанные с QoS. Функция CCF_T также использует и модифицирует соединения ATM в $ATMC_T$.

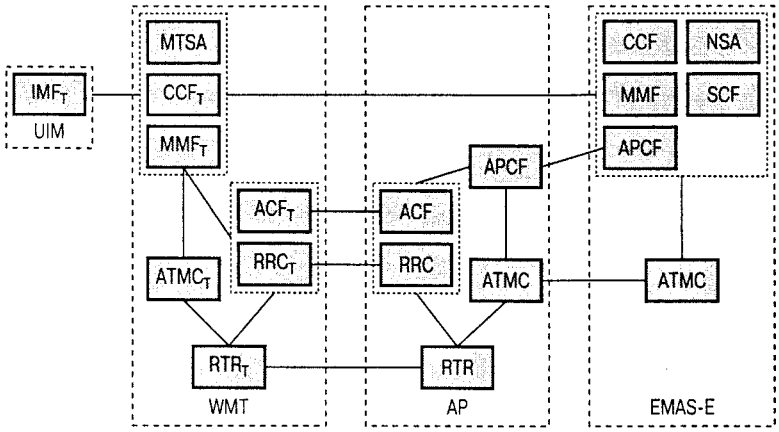

Рис. 8.5. Функции и взаимосвязи, необходимые для поддержки мобильности

- **Функция управления тождественностью (IMF$_T$).** Подобно функциям тождественности абонента и оборудования мобильных терминалов GSM, пользователю мобильных терминалов ATM требуются функции для хранения информации, связанной с идентификацией и безопасностью.
- **Агент безопасности мобильного терминала (MTSA).** Самому терминалу (в целях безопасности) могут требоваться дополнительные функции, независимые от пользователя и информации, сохраненной в UIM. Эта функция может помочь аутентифицировать терминал и аппаратное средство, которое не обязательно связано с определенным пользователем.

Для управления беспроводным и радиодоступом нужны дополнительные функции.

- **Радиопередача и прием (RTR$_T$).** Это основная функция, делающая возможными отправку и прием данных. RTR$_T$ включает в себя уровни управления логическим каналом (logical link control — LLC), управления доступом к среде (medium access control — MAC) и физический (PHY) уровень. В число включенных механизмов входят защита от ошибок с целью улучшения качества передачи (повторная передача и/или прямое исправление ошибок), уплотнение и разуплотнение, шифрование и дешифрование. Физический уровень управляет модуляцией, скремблированием, режимами работы и другими функциями, связанными с радиоканалом.
- **Контроль радиоресурсов (RRC$_T$).** Ресурсы радиосвязи должны контролироваться, т.е. выбираться, освобождаться, резервироваться, как определено плоскостью управления. Кроме того, RRC$_T$ может инициировать переключения, наблюдать за радиосвязью с абонентами и анализировать ее.
- **Функция управления ассоциациями (ACF$_T$).** Эта функция отвечает за сопоставление беспроводного терминала с его точкой доступа. Она включает в себя установку и освобождение ассоциации, поддержку переключения внутри точки доступа, переход в экономичный режим питания и переключение радиосигнала.
- **Функция соединения ATM (ATMC$_T$).** Эта функция отвечает за соединения ATM, то есть, настройку, разрыв и модификацию соединения, а также предоставляет службу-носитель для плоскости управления и пользовательской плоскости. Эта функция управляет элементами соединения ATM, предоставляя стандартные службы ATM, то есть, постоянную скорость передачи (constant bitrate — CBR), переменную скорость передачи в реальном времени (real-time variable bitrate — rt-VBR), переменную скорость передачи не в реальном времени (non real-time variable bitrate — nrt-

VBR), доступную скорость передачи (available bitrate — ABR) и не уточненную скорость передачи (unspecified bitrate — UBR) [Händel, 1994], [ATMForum, 1996].

8.5.2. Функции сети

Сетевая сторона реализует дополнения функций терминала и некоторые дополнительные механизмы, предназначенные для сетевой интеграции и администрирования. Управление службами и соединениями поддерживается перечисленными ниже функциями.

- **Функция управления точкой доступа (APCF)**. Реализуемая в точке доступа и в объекте EMAS-E, эта функция используется для избирательного доступа, переключения, управления точкой доступа и сообщения об условиях связи и загруженности линии связи.
- **Функция управления вызовами и соединениями (CCF)**. Для установления соединения с какой-либо другой системой CCF устанавливает, поддерживает и разрывает соединения. Кроме того, она осуществляет управление включением в соединение, что необходимо в сетях ATM, и запрашивает сетевые и радиоресурсы.
- **Агент безопасности сети (NSA)**. Он важен для сетей с беспроводным доступом, поскольку осуществляет управление беспроводными терминалами. Таким образом, эта функция управляет тождественностью, аутентифицирует абонентов, контролирует конфиденциальность и шифрует данные.
- **Функция управления службами (SCF)**. Эта функция высокого уровня поддерживает хранение профилей служб и доступ к ним, также она обеспечивает проверку согласованности этих профилей. Эта функция имеет наивысший уровень в среде всего управления службами.
- **Функция управления мобильностью (MMF)**. Наконец, MMF — это функция, которая поддерживает мобильность со стороны сети, подобно управлению размещением, переключению, запоминание данных о расположении и личности абонента. Кроме того, она контролирует избирательный доступ, маршрутизацию и обновление местоположения.

Подобно беспроводным терминалам, со стороны сети также требуются функции, обеспечивающие управление радиопередачей и доступ к среде. Эти функции перечисляются ниже.

- **Функция управления ассоциациями (ACF)**. Эта функция управляет установлением и разрывом связи между мобильным терминалом и точкой доступа. Она включает в себя избирательный доступ, принятие решения о переключении и его реализацию, наблюдение за каналом связи и его анализ. Кроме того, она ретранслирует системную информацию и поддерживает режим энергосбережения.
- **Функция управления радиоресурсами (RRC)**. RRC реализует основное управление радиоресурсами, совместно используемыми многими беспроводными терминалами. Она управляет каналами радиосвязи, анализирует конфигурацию среды мобильной радиосвязи и может инициировать переключение при изменении условий передачи.
- **Функция радиопередачи и приема (RTR)**. Функция RTR со стороны сети также (как и со стороны терминала) содержит уровни LLC, MAC и физический. Основным дополнительным механизмом для сетевой стороны является уплотнение ячеек ATM согласно контракту об информационном обмене и параметрам QoS. Этот момент также является одним из основных отличий сетей WATM от многих других беспроводных сетей. WATM пытается поддерживать столько же параметров информационного обмена, сколько поддерживает ATM.
- **Функция соединения ATM (ATMC)**. Эта функция управляет соединениями терминала. Дополнительно она включает уплотнение потоков информационного обмена ATM.

8.6. Уровень радиодоступа

На рис. 8.3 и 8.4 показано положение уровня радиодоступа (radio access layer — RAL) в эталонной модели WATM. В следующих разделах перечислены требования к уровню RAL и приводится пример — стандарт ETSI BRAN (широкополосная сеть радиодоступа).

8.6.1. Требования

Беспроводный доступ к WATM будет обеспечиваться уровнем радиосвязи RAL, расположенным под уровнями ATM. Уровень RAL должен работать в частотном диапазоне 5 ГГц и удовлетворять требованиям, которые перечислены ниже.

8.6.1.1. Физический уровень

Физический уровень должен задавать частоты, эффективное многократное использование частоты, антенны, мощность передатчика, максимальный диапазон, характеристики ячейки, несущие частоты, скорость передачи символа, схему модуляции, канальное кодирование, настроечные последовательности и т.д. Кроме того, должны быть определены интерфейсы для потока данных и потока управляющей информации от устройства радиосвязи и к этому устройству.

Ниже перечислены требования, определенные для физического уровня RAL.

- **Частота появления ошибочных битов (bit error rate — BER)**. Эта величина должна быть меньше 10^{-4} при доступности 99,5% в точке доступа к службе физического уровня.
- **Размер пакета**. Физический уровень должен эффективно поддерживать PDU единичной ячейки. Однако, более эффективным был бы общий случай нескольких ячеек на PDU, следовательно, он также должен быть возможен.
- **Скорость передачи данных**. Для уровня ATM над уровнем RAL должна быть доступна скорость 25 Мбит/с (то есть, более 50 000 ячеек/с).
- **Максимальный диапазон**. В зависимости от местной конфигурации и среды, должен быть возможен охват 30–50 м внутри помещений, а вне помещений — 200–300 м.
- **Исправление ошибок**. Вследствие большого риска появления битовых ошибок физический уровень должен включать некоторые механизмы исправления ошибок.
- **Мощность передачи**. Обычно, мощность передачи ограничивается местной конфигурацией. Требования физического уровня предполагают наличие, по крайней мере, 100 мВт EIRP.

8.6.1.2. Уровень управления доступом к среде

Уровень MAC (medium access control) должен управлять одновременным доступом к среде нескольких беспроводных мобильных терминалов. Это включает определение протокола MAC, формата PDU и алгоритма управления MAC. Кроме того, уровень MAC должен поддерживать мобильность пользователя и предоставлять интерфейсы к физическому уровню и уровню управления логическим каналом. Были определены следующие основные задачи.

- **Логические каналы**. Как показано на рис. 8.3, уровень MAC в уровне RAL связывает уровни ATM точек доступа и беспроводный мобильный терминал. Таким образом, уровень MAC должен также поддерживать логические каналы для виртуальных соединений ATM, каждый из которых связан с набором параметров информационного обмена и QoS ATM.
- **Параметры QoS/информационного обмена**. Уровень MAC должен поддерживать параметры QoS и информационного обмена согласно спецификации ATM TM 4.0 [ATMForum, 1996]. Сквозное QoS не должно зависеть от MAC, что является одним из преимуществ ATM перед

другими технологиями без обязательств, такими как Mobile IP (см. раздел 9.1), в комбинации с сетью WLAN (например, сетью IEEE 802.11, см. раздел 7.3).

- **Архитектура**. В типичной архитектуре WATM использована инфраструктурная сеть (см. рис. 8.1). Поддержка этой архитектуры является обязательной для уровня MAC. Необязательным требованием является поддержка эпизодических конфигураций.
- **Предоставление услуг**. Услуги, предлагаемые уровнем MAC, должны включать в себя следующие классы служб ATM: CBR, VBR-rt, VBR-nrt, ABR и UBR с контролем QoS. Эффективность уровня MAC должна превышать 60–75% (последние исследования показывают, что возможна эффективность свыше 90% [Anastasi, 1998], [Petras, 1997], [Sanchez, 1997]). Максимальная поддерживаемая скорость передачи данных равна, по меньшей мере, 25 Мбит/с при долговременной скорости передачи данных 6 Мбит/с. MAC должен также поддерживать низкие скорости передачи данных, такие как 32 Кбит/с CBR для телефонии.

8.6.1.3. Уровень управления логическим каналом

Уровень управления логическим каналом (logical link control — LLC), который размещен между уровнями ATM и MAC, должен решать специфические проблемы в области беспроводной передачи и сочетаться с уровнем MAC для образования уровня управления каналом передачи данных (data link control — DLC). Здесь снова должны быть определены PDU протокола LLC и собственно протокол LLC, что включает специальные заголовки беспроводной передачи для пакетов данных, управляющие сообщения и, возможно, специальные функции для классов служб ATM. Уровень LLC главным образом отвечает за защиту от ошибок, выявление и исправление ошибок, используя выборочную повторную передачу и прямое исправление ошибок. Ниже перечислены основные требование к уровню LLC.

- **Автоматический запрос повторенной передачи (automatic repeat request — ARQ)**. Для уровня LLC обязательной является поддержка схемы ARQ. Однако необязательно для различных классов использовать различные схемы ARQ. Кроме того, необязательно поддерживать малый разброс задержек в ячейке (CDV) с использованием специальной схемы ARQ.
- **Прочие необязательные свойства.** LLC может предоставлять специальные механизмы прямого исправления ошибок для поддержки классов служб реального времени. Кроме того, уровень избыточности и схема FEC могут адаптироваться к требованиям, определяемым QoS. LLC может также предоставлять средства поддержки переключения и сохранения последовательности ячеек при переключении. В дополнение к этому, в LLC могут реализовываться аутентификация, кодирование и дополнительные механизмы безопасности.

8.6.2. BRAN

Широкополосная сеть радиодоступа (broadband radio access network — BRAN), которая была стандартизована Европейским институтом стандартов по телекоммуникациям (ETSI), является возможным выбором RAL для WATM. Хотя сеть BRAN была стандартизована независимо от WATM, они объединяются для сосредоточения общих усилий на одной цели.

Поводом к разработке BRAN послужила отмена государственного регулирования и приватизация телекоммуникационного сектора в Европе. Многие новые поставщики услуг сталкиваются с проблемой получения доступа к потребителям, поскольку телефонная инфраструктура принадлежит нескольким большим компаниям. Одной из возможных технологий предоставления доступа пользователей к сети является радиосвязь. Преимущество такого подхода заключается в высокой гибкости и быстрой установке. Поддерживаются различные типы информационного обмена, транспортный поток можно уплотнять для повышения эффективности, а соединение может быть асимметричным (как, например, в типичном сценарии WWW, когда

многие потребители скачивают существенные объемы информации с серверов, однако на серверы отправляют лишь небольшое количество данных). Кроме того, радиодоступ допускает экономичный рост пропускной способности доступа. Если нужна большая пропускная способность, легко можно установить дополнительный трансивер, в то время как для проводных терминалов это включало бы прокладку дополнительной проводки. Основной рынок для BRAN включает в себя частных потребителей, а также мелкие и средние компании с приложениями Internet, проведение мультимедийных конференций и виртуальные частные сети.

Стандартизация BRAN охватывает довольно большие масштабы: мобильность внутри помещений и на территории учебных заведений, скорости передачи данных 25–155 Мбит/с, радиус передачи в пределах 50 м–5 км. Работы по стандартизации координируются с ATM Forum, IETF, другими группами из ETSI, IEEE и т.д. Под общим заголовком BRAN были специфицированы четыре различных типа сетей, которые перечислены ниже [ETSI, 1998a].

- **HIPERLAN 1**. Высокоскоростная линия WLAN поддерживает мобильность при скоростях передачи данных свыше 20 Мбит/с. Охватываемый диапазон равен 50 м, соединения многоточечные с использованием эпизодических или инфраструктурных сетей [ETSI, 1998б]. Подробно HIPERLAN 1 обсуждается в разделе 7.4.
- **HIPERLAN 2**. Эта технология может использоваться для беспроводного доступа к сетям ATM или IP и поддерживает скорость до 25 Мбит/с в многоточечной конфигурации. Охватываемый диапазон — 50 м с поддержкой медленной (менее 10 м/с) мобильности [ETSI, 1997].
- **HIPERACCESS**. Эта технология может использоваться для покрытия "последней мили" до потребителя посредством стационарного канала радиосвязи. Таким образом, это альтернатива кабельных модемов или технологий xDSL [ETSI, 1998в]. Охватываемый диапазон — до 5 км, а скорость передачи — до 25 Мбит/с.
- **HIPERLINK**. Для соединения различных точек доступа HIPERLAN или узлов HIPERACCESS высокоскоростной линией связи можно выбрать технологию HIPERLINK. HIPERLINK предоставляет стационарное двухточечные соединения со скоростями до 155 Мбит/с.

Общие характеристики HIPERLAN 2, HIPERACCESS и HIPERLINK включают поддержку следующих классов служб ATM: CBR, VBR-rt, VBR-nrt, UBR и ABR. Таким образом, очевидно, что только HIPERLAN 2 может использоваться как уровень RAL в WATM. Эта технология удовлетворяет требованиям поддержки качества обслуживания ATM, мобильности, беспроводного доступа и высокой пропускной способности.

Как сеть доступа, технология BRAN не зависит от протоколов стационарной сети. Таким образом, BRAN может использоваться для сетей ATM и TCP/IP, что проиллюстрировано на рис. 8.6. Базирующийся, возможно, на различных физических уровнях, уровень DLC сети BRAN предлагает общий интерфейс для более высоких уровней. Для охвата специальных характеристик беспроводных линий связи и непосредственной адаптации к различным сетевым технологиям более высоких уровней BRAN предоставляет подуровень сетевой конвергенции. Это уровень, который может использоваться беспроводными сетями ATM или сетями IP. Если в качестве уровня RAL для WATM используется BRAN, базовая сеть ATM будет использовать службы подуровня конвергенции сети BRAN.

Изучая эталонную модель HIPERLAN 2 (рис. 8.7) [ETSI, 1997], можно найти много общего с общей эталонной моделью WATM (рис. 8.2). Снова внешняя сеть связана с коммутатором, поддерживающим мобильность. Этот коммутатор имеет точку беспроводного доступа, которую можно разделить на две части: контроллер и трансивер. Следующим компонентом радиоинтерфейса H2.1.1 является беспроводный терминальный адаптер. Вместе точка доступа и беспроводный терминальный адаптер образуют беспроводную подсистему. И наконец, модуль

пользователя связан с беспроводным терминальным адаптером. Этот пользователь должен теперь управлять мобильностью, но не проблемами беспроводного доступа.

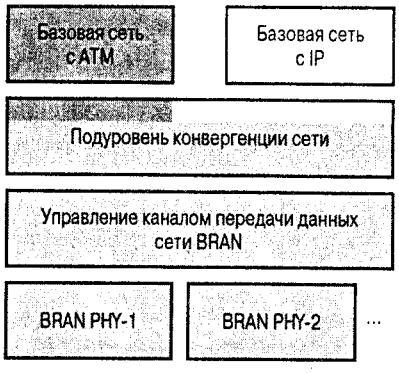

Рис. 8.6. Уровневая модель сетей доступа BRAN

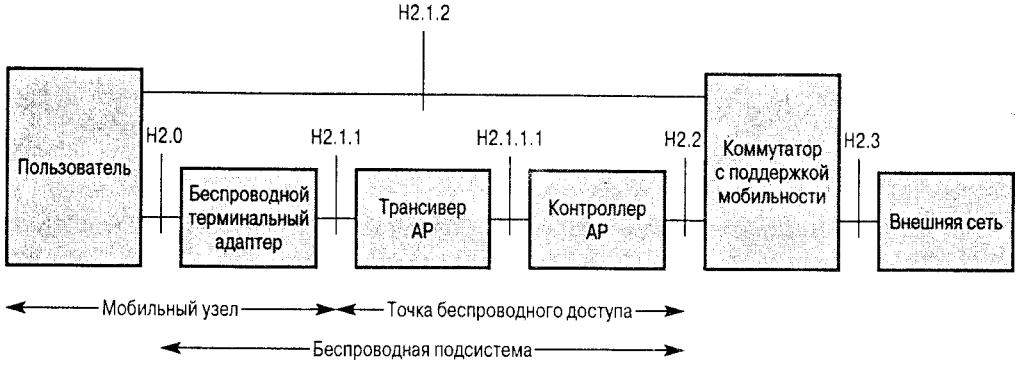

Рис. 8.7. Эталонная модель HIPERLAN 2

8.7. Переключение

Одной из наиболее важных тем в среде WATM является переключение. В протоколах без обязательств, без установления соединения, поддерживающих переключение (например, Mobile IP), не требуется особенно много внимания уделять заботе о качестве переключения. Эти протоколы не гарантируют определенных параметров информационного обмена, как это делает WATM. Основной проблемой WATM при переключении является повторная маршрутизация всех соединений с сохранением качества соединения. В то время, как переключение в среде без обязательств, без установления соединения главным образом включает изменение маршрута потока пакетов без заботы о надежной передаче, конечные системы в сетях WATM могут поддерживать большое число соединений, каждое из которых имеет различные требования к качеству обслуживания (например, ограниченная задержка, ограниченная дрожь, минимальная пропускная способность и т.д.). Таким образом, переключение включает не только изменения маршрута соединений, но также и резервирования ресурсов в коммутаторах, проверка доступной пропускной способности радиополосы, отслеживание терминалов с целью осуществления упреждающего резервирования и т.д.

8.7.1. Эталонная модель переключения

Простая эталонная модель переключения в системе WATM показана на рис. 8.8. Сквозное соединение ATM разделено на несколько сегментов. **Стационарный сегмент** — это часть соединения, на которую не влияет переключение, чего нельзя сказать о **сегменте переключения**, который полностью размещен внутри **домена переключения**. Домен переключения может объединять несколько коммутаторов, точек доступа и скорее всего размещен внутри одного административного домена. Границей между сегментом переключения и стационарным сегментом является **точка анкера**.

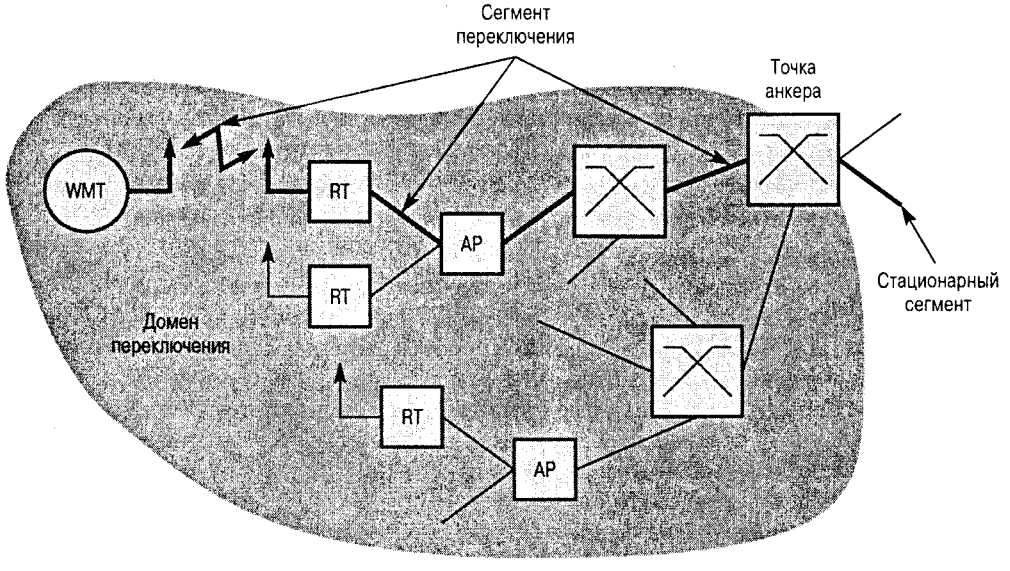

Рис. 8.8. Пример эталонной модели переключения

Если обе конечные системы являются беспроводными мобильными терминалами ATM, стационарного сегмента может не быть вообще. В этом случае точка анкера соединяет два сегмента переключения, которые могут находиться в различных доменах переключения.

8.7.2. Требования к переключению

Для переключения было установлено множество различных требований. В следующем списке представлены некоторые из них (согласно работам [Toh, 1997], [Bhat, 1998б]).

- **Переключение множественных соединений**. Поскольку ATM является технологией, ориентированной на установление соединений, когда конечные системы могут поддерживать множество соединений в одно и то же время, переключение в WATM должно поддерживать более одного соединения. Из этого требования вытекает, что после переключения изменяется маршрут каждого соединения. Однако возможна ситуация, когда доступные ресурсы не позволяют изменить маршруты всех соединений или приводят к ухудшению качества обслуживания. В таком случае терминал может решить принять более низкое качество или отбросить отдельные соединения.
- **Переключение одноточечного соединения в многоточечное.** Прямая поддержка связи одноточечных соединений с многоточечными является одним из главных преимуществ технологии ATM. Следовательно, переключение WATM также должно поддерживать эти типы соединений. Однако вследствие сложности схемы могут понадобиться некоторые ограничения. Пример: требова-

ние отсутствия точки расщепления[1] внутри сегмента переключения. Это ограничение позволяет избежать необходимости перемещать точку расщепления при переключении.

- **Поддержка QoS**. Переключение должно быть направлено на сохранение качества обслуживания всех соединений. Однако вследствие ограниченности ресурсов это не всегда возможно. Следовательно, могут потребоваться функции повторного согласования качества обслуживания и отбрасывания соединений на приоритетной основе. Потенциальные точки доступа должны сообщать о своих ресурсах терминалу, а эта информация затем может использоваться алгоритмом переключения для оптимизации переключения и для согласования различных точек доступа.
- **Целостность данных и безопасность**. Переключение WATM должно минимизировать потери ячеек и избежать дублирования ячеек или переупорядочивания. Кроме того, переключение не должно подвергать риску безопасность соединения между терминалом и сетью.
- **Поддержка передачи сигналов и маршрутизации**. WATM должна предоставлять средства идентификации коммутаторов с поддержкой мобильности, средства определения соседних коммутаторов и средства изменения маршрута отдельных соединений в домене переключения.
- **Производительность и сложность**. То, что системы WATM по своей природе сложны, в основном является следствием того, что они поддерживают соединения с QoS. Таким образом, основной целью при разработке переключения должна быть его простота. Модификации, вносимые в коммутаторы с поддержкой мобильности, должны быть крайне ограниченны, однако требуемые функции могут иметь довольно строгие требования ко времени обработки. Из-за необходимости поддержания заданной производительности коммутаторы ATM сильно зависят от аппаратной реализации и, следовательно, в них значительно сложнее произвести обновление и реализовать новые функции. Код переключения, необходимый терминалам, должен быть достаточно простым, поскольку больший размер кода требует также большей производительности, то есть, большей мощности батарей, которая, как правило, существенно ограничивает разработку мобильных терминалов.

8.7.3. Типы переключений

WATM будет поддерживать следующие типы переключений.

- **Жесткое переключение**. WATM не поддерживает мягкое переключение, то есть, она не допускает наличия в мобильных терминалах более одного радиосоединения одновременно. Обычно мягкое соединение поддерживается системами, основанными на технологии CDMA, где несколько точек доступа могут получать один и тот же сигнал одновременно. WATM в каждый момент времени поддерживает только одно соединение с одной точкой доступа. Следовательно, чрезвычайно важно избегать прерываний при переключении.
- **Активация терминалом**. Терминал WATM может решить инициировать переключение, основываясь, например, на текущем качестве сигнала.
- **Активация сетью**. Сеть, исходя из мощности сигнала или текущей загруженности, также может решить, что необходимо переключение.
- **Активация сетью при содействии терминала**. Беспроводный терминал может предоставлять сети информацию о текущем состоянии радиопередачи и других параметрах. Основываясь на этой информации, сеть может принять решение об инициации переключения.

---

[1] Точка расщепления – это коммутатор, в котором входная линия связи разделяется на два или более исходящих соединения.

- **Управление со стороны сети**. Сеть является последней инстанцией, принимающей решение относительно переключения, поскольку именно у нее постоянно есть информация о текущей загруженности, свободных ресурсах и т.д.
- **Обратное переключение**. Стандартный тип переключения именуется "обратным". Здесь беспроводный мобильный терминал WMT отмечает, например, затухание сигнала и инициализирует переключение на новую точку доступа. Во время переключения терминал хранит текущее радиосоединение и переключается на новую точку доступа AP только после того, как радиоресурс будет зарезервирован и все участвующие объекты подготовлены к переключению.
- **Прямое переключение.** Переключение называют "прямым", если WMT неожиданно попадает в новую точку доступа, которая теперь должна инициировать переключением на нее и далее управлять этим процессом. Это может произойти, если, например, беспроводный терминал неожиданно утратил связь со старой точкой доступа. В этом случае для системы WATM нет времени на установление обратного переключения. Причиной подобной потери соединения может быть интерференция или быстрое передвижение терминала.

8.7.4. Сценарии переключения

На рис. 8.9 показаны три основных сценария переключения для WATM.

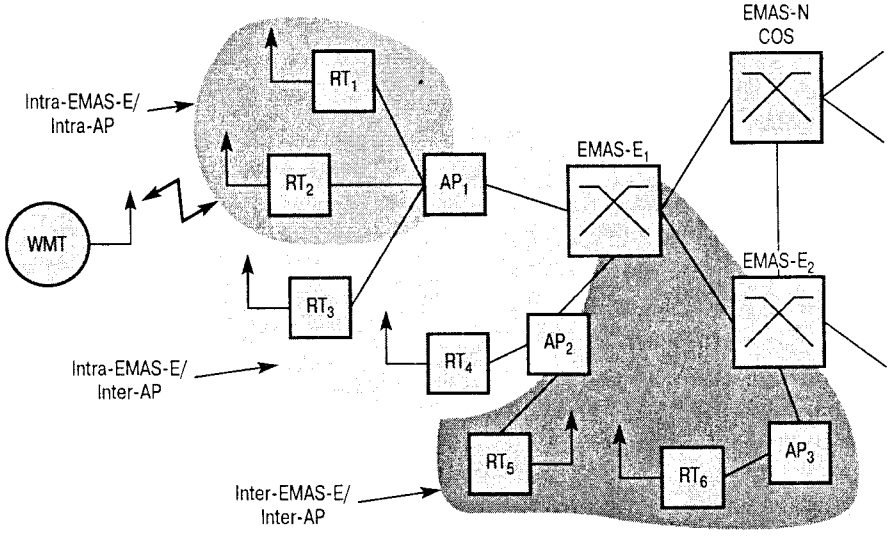

Рис. 8.9. Три основных сценария переключения для WATM

- **Intra-EMAS-E/Intra-AP.** Терминал WMT движется из радиотрансивера RT_1 в радиотрансивер RT_2. Переключение производится внутри домена текущей точки доступа AP_1 и текущего объекта $EMAS\text{-}E_1$. Основной задачей точки доступа является переключение между радиотрансиверами.
- **Intra-EMAS-E/Inter-AP**. WMT движется от RT_3 к RT_4. Переключение производится внутри домена $EMAS\text{-}E_1$, однако теперь коммутатор должен переключить соединение с AP_1 на AP_2. Таким образом, соответствующие точки доступа должны прервать или установить соединение.
- **Inter-EMAS-E/Inter-AP**. В переключении могут участвовать различные объекты EMAS-E. WMT теперь перемещается от RT_5 к RT_6. Другой промежуточный коммутатор, называемый

разделительным (cross over switch — COS), должен управлять переключением между коммутаторами EMAS-E_1 и EMAS-E_2. COS не обязательно должен быть следующим переключателем (см. рис. 8.9), он может находиться где угодно между точкой анкера (см. рис. 8.8) и двумя используемыми EMAS-E, в том числе, коммутатором COS может быть один из EMAS-E или сама точка анкера. В работе [Akyol, 1996] представлено несколько схем изменения маршрута, поддерживающих переключение inter-EMAS-E/inter-AP; там же эти схемы сравниваются с позиций полосы пропускания, необходимой для переключения, количества передаваемых сообщений, задействованных коммутаторов, времени выполнения переключения, надежности при неустойчивом переключении и т.д.

8.7.5. Обратное переключение

При обратном переключении система WATM может выбрать наилучшую (с точки зрения требований к терминалу и сети) точку доступа (см. рис. 8.10). Вначале WMT связан с RT_2 (см. рис. 8.9). Как только WMT замечает, что требуется переключение, он отправляет запрос BW_HO_REQEST на коммутатор EMAS-E_1. Это сообщение включает в себя список используемых целевых радиотрансиверов, например, RT_1, RT_4 и RT_6. EMAS-E_1 теперь ищет точки доступа, контролирующие трансиверы списка. (Это AP_1 для RT_1, AP_2 для RT_4 и AP_3 для RT_6.)

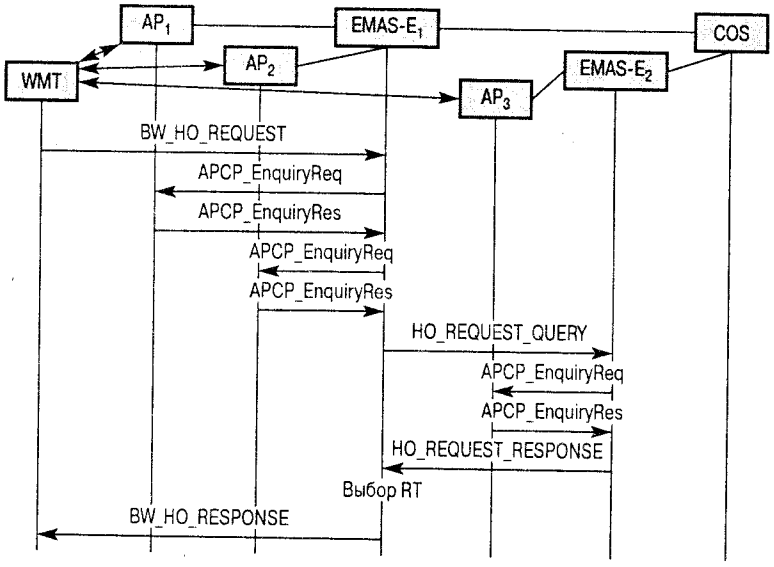

Рис. 8.10. Обратное переключение с множественными возможными точками доступа

Следующий шаг включает проверку доступности ресурсов радиосвязи. Следовательно, EMAS-E_1 посылает сообщение APCP_EnquiryReq на AP_1 и AP_2. EMAS-E_1 не может отправить сообщение непосредственно на AP_3, так что он отправляет сообщение HO_REQEST_QUERY на EMAS-E_2, который, в свою очередь, отправляет сообщение APCP_EnquiryReq на AP_3. Все три сообщения, отправляемые EMAS-E_1, ограничены таймером. После получения сообщений APCP_EnquiryRes или истечении времени ожидания запроса EMAS-E_1 выбирает один из трансиверов в качестве нового места назначения для WMT. Этот выбор может основываться на характере информационного обмена, текущих ресурсах и т.д. или может согласовываться с используемыми алгоритмами. После выбора трансивера EMAS-E_1 возвращает WMT сообщение

BW_HO_RESPONSE, указывающее, что трансивер выбран и что теперь соединение может переключаться, как объясняется в следующем разделе.

8.7.5.1. Обратное переключение Intra-EMAS-E/Intra-AP

При обратном переключении Intra-EMAS-E/Intra-AP WMT также указывает переключение, посылая сообщение BW_HO_REQUEST на свой объект EMAS-E (см. рис. 8.11). EMAS-E теперь проверяет доступность радиоресурсов в точке доступа, с которой в настоящее время связан WMT. Кроме списка радиотрансиверов, полученного от WMT, сообщение APCP_EnquiryRes содержит все идентификаторы соединений, которые требуется переключить. Кроме того, каждое соединение соотнесено с информационным потоком ATM и параметрами QoS.

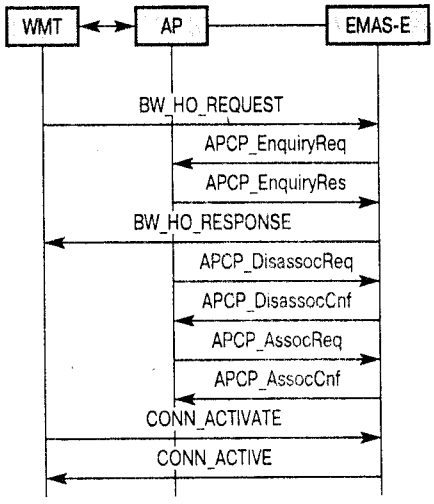

Рис. 8.11. Обратное переключение Intra-EMAS-E/Intra-AP

Как только точка доступа получает сообщение APCP_EnquiryReq, она проверяет доступные радиоресурсы и резервирует ресурсы, требуемые для переключения. Если точка доступа может выбирать между несколькими трансиверами, она выбирает поддерживающий максимальное количество соединений при заданных требованиях QoS. Точка доступа сообщает EMAS-E о выбранном трансивере в сообщении APCP_EnquiryRes. Это сообщение также включает в себя список соединений, которые могут переключаться.

Теперь EMAS-E может передать эту информацию с точки доступа на WMT в сообщении BW_HO_REQUEST. Затем WMT разрывает соединение со старым трансивером и начинает связываться с новым. Это включает определение расположения ресурсов, зарезервированных точкой доступа. После потери старого радиосоединения, точка доступа может отправить сообщение APCP_DisassocReq на объект EMAS-E, который, в свою очередь, отвечает подтверждением APCP_DisassocCnf. После соединения WMT могут последовать сообщения APCP_AssocReq и APCP_AssocCnf. Четыре последних сообщения указывают отключение старого и установление нового радиосоединения с EMAS-E.

Наконец, WMT сообщает EMAS-E о готовности возобновить отправку и получение пользовательской информации (сообщение CONN_ACTIVATE). EMAS-E отвечает сообщением CONN_ACTIVE, которое включает в себя ссылки на соединения, которые наконец-то были переключены.

8.7.5.2. Обратное переключение Intra-EMAS-E/Inter-AP

Если переключение осуществляется между различными точками доступа, связанными с одним и тем же EMAS-E (см. рис. 8.12), требуется еще несколько сообщений между WMT, точкой доступа и EMAS-E. Снова WMT отправляет сообщение BW_HO_REQUEST на коммутатор EMAS-E, в которое включен список соединений, которые должны быть переключены и список возможных новых радиотрансиверов. EMAS-E определяет подходящую точку доступа (одну или несколько, но в рассматриваемом примере — только AP_2) и отправляет на эту точку доступа сообщение APCP_EnquiryReq.

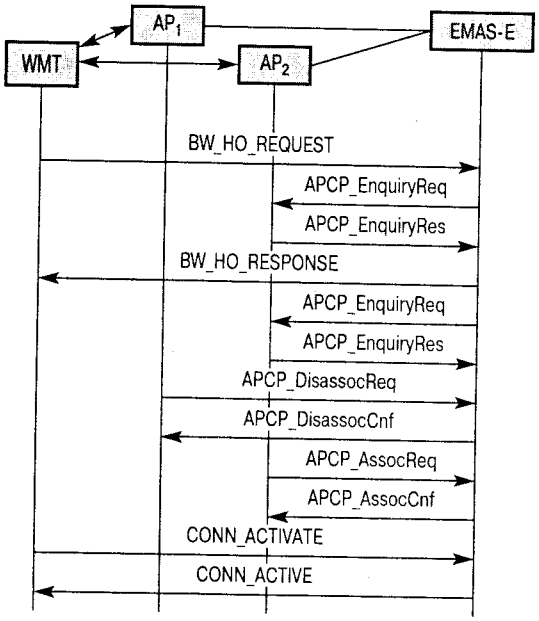

Рис. 8.12. Обратное переключение Intra-EMAS-E/Inter-AP

Получив это сообщение, точка доступа проверяет доступные ресурсы для поддержки соединений, указанных в списке переключения. Затем она выбирает трансивер и отправляет сообщение APCP_EnquiryRes, чтобы определить выбор и перечислить соединения, которые может поддерживать точка доступа. На этом этапе точка доступа не резервирует ресурсы для переключения. Теперь EMAS-E выбирает для переключения точку доступа, которая может поддерживать максимальное количество переключений или, например, гарантирует наивысшее качество. Сообщение BW_HO_RESPONSE доставляет эту информацию WMT.

Помимо этого, EMAS-E запрашивает резервирование ресурсов в выбранной точке доступа (сообщение APCP_Enquiry_Req), а точка доступа в ответ сообщает о реальном резервировании (APCP_EnquiryRes). После получения сообщения BW_HO_RESPONSE WMT может отключить соединение, о чем сообщается посредством APCP_DisassocReq со старой точки доступа (здесь — AP_1) и на что дается подтверждение APCP_DisassocCnf. Новая точка доступа, AP_2, информирует EMAS-E о прибытии WMT (сообщение APCP_AssocReq), на что также передается подтверждение (APCP_AssocCnf).

Теперь WMT снова готов отправлять данные, о чем информируется посредством сообщения CONN_ACTIVATE. EMAS-E отвечает сообщением CONN_ACTIVE, содержащим информацию о соединениях, которые были переключены.

8.7.5.3. Обратное переключение Inter-EMAS-E/Inter-AP

Наиболее сложным случаем является переключение из одной точки доступа в другую, если они принадлежат разным объектам EMAS-E (см. рис. 8.13). Теперь для переключения соединения с одного коммутатора на другой необходим разделительный коммутатор (cross over switch — COS). Снова WMT начинает с отправки сообщения BW_HO_REQUEST своему объекту EMAS-E_1, который обнаруживает, что он не отвечает за трансиверы, перечисленные в сообщении WMT. Соответственно, он посылает сообщение HO_REQUEST_QUERY на EMAS-E, который отвечает за соответствующую точку доступа, к которой прикреплен трансивер. Процедура нахождения подходящего EMAS-E зависит от схемы адресации трансиверов.

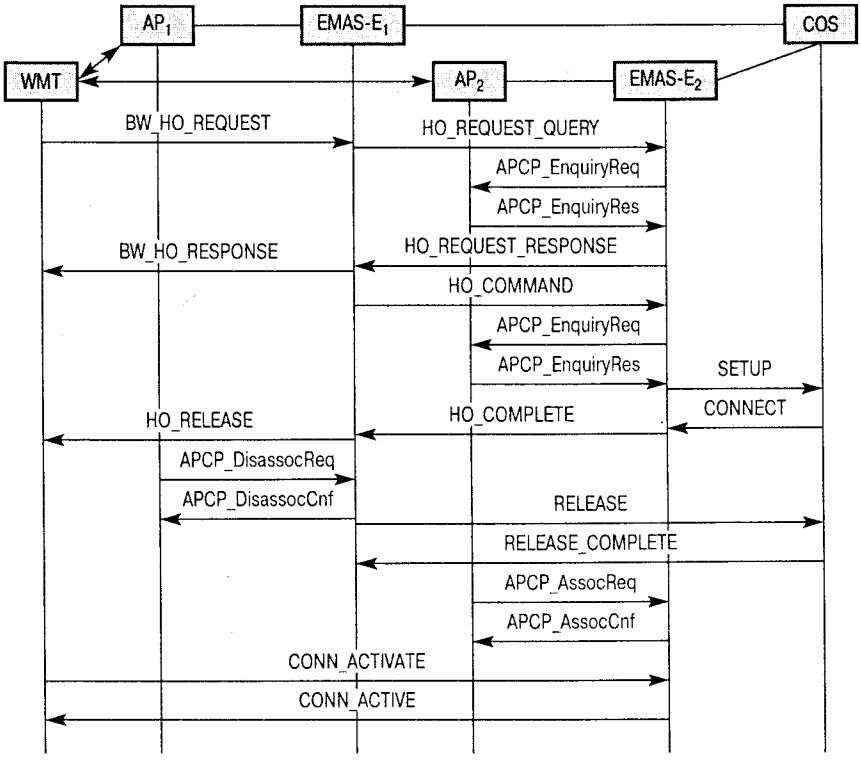

Рис. 8.13. Обратное переключение Inter-EMAS-E/Inter-AP

Как объяснялось выше, этот объект EMAS-E_2 проверяет доступные ресурсы в точке AP_2 с помощью запроса APCP_EnquiryRes и получает в ответ APCP_EnquiryRes. В зависимости от полученной информации EMAS-E выбирает наилучшую точку доступа и сообщает об этом старому объекту EMAS-E_1, используя сообщение HO_REQUEST_RESPONSE. EMAS-E_1 теперь может проинформировать WMT (сообщение BW_HO_RESPONSE) о трансивере, который следует выбирать. Кроме того, EMAS-E_1 выдает EMAS-E_2 сообщение HO_COMMAND, содержащее список соединений, задействованных в переключении, их требования QoS и (если известно) коммутатор COS, отвечающий за изменение маршрутов пользовательских соединений. Коммутатор COS может находиться в различных местах. Это может быть сам EMAS-E_1, и в этом случае переключение называется **расширением виртуального канала**, поскольку переключение расширяет текущее соединение со старого EMAS-E на новый. EMAS, действующий как коммутатор COS, может быть известен старому EMAS-E либо изначально (заданная конфигу-

рация), либо посредством динамического определения конфигурации при установлении соединения — этот случай называется **изменением маршрута на основе анкера**. Третьей альтернативой является **динамическое обнаружение коммутатора COS**, при котором EMAS-E_1 включает адреса потенциальных коммутаторов COS в сообщение HO_COMMAND.

EMAS-E_2 теперь резервирует ресурсы в точке AP_2, используя APCP_EnquiryReq и получает в ответ APCP_EnquiryRes. Если EMAS-E_2 получил адрес COS, он выдает на этот коммутатор сообщение SETUP. В противном случае, он должен использовать дополнительные механизмы для отправки сообщения SETUP на коммутатор COS [ATMForum, 1998a], [ATMForum, 1998б]. Если соединение можно установить между EMAS-E_2 и COS, EMAS-E_2 отправляет сообщение HO_COMPLETE на EMAS-E_1, это должно указать, что переключение соединения прошло успешно. Для каждого соединения, указанного в сообщении HO_COMMAND, требуется установка нового соединения между EMAS-E_2 и COS. У WMT может одновременно быть много активных соединений и, таким образом, каждое отдельное соединение требует переключения, для чего выдаются сообщения SETUP/CONNECT и уведомляется старый EMAS-E (сообщение HO_COMPLETE).

После того, как EMAS-E получил сообщение HO_COMPLETE для последнего соединения, он отправляет WTM сообщение HO_RELEASE. Это сообщение указывает WTM, что он может разрывать радиосоединение со старым трансивером, соотнесенным с AP_1. Точка AP_1 сообщает об этом событии, используя APCP_DisassocReq, EMAS-E_1 посылает подтверждение APCP_DisassocCnf. Это открепляет все ресурсы, связанные с AP_1. EMAS-E_1 может теперь отключить все соединения, которые он еще удерживал с COS и которые были частью соединений с WMT. Все это выполняется для каждого соединения, используя сообщение RELEASE, на которое COS отвечает сообщением RELEASE_COMPLETE. В момент разрыва коммутатором COS соединения со старым EMAS-E коммутатор переключается на новый EMAS-E, используя соединения, которые были установлены прежде.

Соотнесение WMT с новой точкой доступа инициирует сообщение APCP_AssocReq, на которое отвечает сообщение APCP_AssocCnf, отправленное с нового EMAS-E. Отправкой сообщения CONN_ACTIVATE WMT указывает, что он готов принимать данные. Новый EMAS-E отвечает на это сообщением CONN_ACTIVE, и переключение полностью завершается.

Следует отметить, что на рис. 8.13 предполагается отсутствие сбоев при переключении, хотя в действительности могут произойти события, препятствующие идеальному переключению всех компонентов. Если, например, новый EMAS-E не может установить соединение с коммутатором COS, он выдает старому EMAS-E сообщение HO_FAILURE, который затем информирует WMT о сбое. Затем WMT может решить удалить соединение или попытаться инициировать прямое соединение. Кроме того, все операции контролируются таймерами, а следовательно, если какая-то операция осуществляется слишком долго, коммутаторы могут сменить EMAS-E либо дать WMT сигнал о сбое.

Может также случиться, что в результате быстрого движения WMT обмен сообщениями не может быть полностью закончен одним из объектов. В таком случае коммутатор возвращается в изначальное состояние вследствие превышения блокировки по времени. В этом примере показано, что на один EMAS-E используется только одна точка доступа. Однако в процесс переключения может задействоваться множество точек доступа, что в целом делает переключение более сложным. Приведенный пример показывает, что переключать соединения с одновременным поддержанием качества обслуживания достаточно сложно, по сравнению с другими технологиями, не гарантирующими качества (таких, например, как протокол IP, см. раздел 9.1).

Снова напомним, что приведенные примеры переключения только представляют текущее состояние обсуждения и что еще не существует завершенного стандарта, задающего все идентификаторы и процедуры.

8.7.6. Прямое переключение

Преимущество обратного переключения заключается в возможности выбора для соединения с WMT наилучших точки доступа и трансивера. При прямом переключении WMT неожиданно обрывает старое соединение и пытается соединиться с новой точкой доступа. Эта новая точка не может резервировать ресурсы заблаговременно и может оказаться не самым лучшим выбором с точки зрения системы. В то же время, если эфирные помехи привели к разрыву текущего соединения, это является единственной возможностью достаточно быстро возобновить соединение. WMT должен оптимизировать свой выбор точки доступа локально. В следующих разделах для прямого переключения описывается та же ситуация, что и выше для обратного. Сообщения, показанные для прямого переключения, передают те же параметры, что и ранее, например, список соединений, ресурсы, такие как трансиверы и т.д.

8.7.6.1. Прямое переключение Intra-EMAS-E/Intra-AP

Если WMT разъединился с точкой доступа вследствие помех, точка доступа сообщает об этом событии, передавая APCP_DisassocReq, а EMAS-E отвечает подтверждением APCP_DisassocCnf. В сценарии, приведенном на рис. 8.14, WMT движется к другому трансиверу той же точки доступа, указанной сообщениями APCP_AssocReq, APCP_AssocCnf. Теперь WMT может отправить EMAS-E сообщение FW_HO_REQUEST, в которое включен список текущих активных соединений и их параметров QoS.

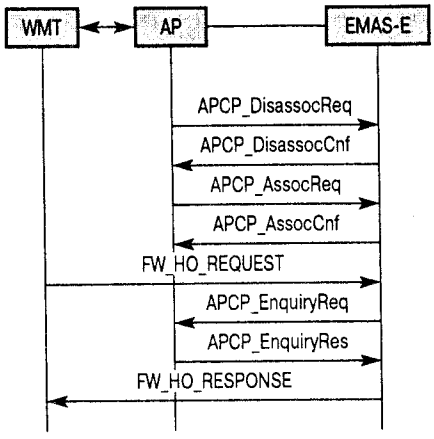

Рис. 8.14. Прямое переключение Intra-EMAS-E/Intra-AP

Затем EMAS-E может попытаться зарезервировать необходимые ресурсы в точке доступа с помощью запроса APCP_EnquiryReq, ответом на который со стороны точки доступа будет APCP_EnquiryRes, сообщение, содержащее список соединений, которые может переключить точка доступа. Количество соединений может уменьшаться вследствие ограниченности ресурсов. Коммутатор EMAS-E пересылает теперь этот список WMT в сообщении FW_HO_RESPONSE, так что WMT может немедленно возобновить связь со всеми соединениями, указанными в сообщении.

8.7.6.2. Прямое переключение Intra-EMAS-E/Inter-AP

При переключении между двумя точками доступа одного объекта EMAS-E (см. рис. 8.15) ситуацию можно сравнить с предыдущим сценарием. Единственным отличием является то, что EMAS-E обменивается парами сообщений (APCP_DisassocReq, APCP_DisassocCnf) со старой точкой доступа (то есть, с AP_1) и парами сообщений (APCP_AssocReq, APCP_AssocCnf) и

(APCP_EnquiryReq, APCP_EnquiryRes) с новой точкой доступа (то есть, с AP_2). Снова, после получения сообщения FW_HO_RESPONSE, WMT может возобновить связь со всеми соединениями, указанными в сообщении.

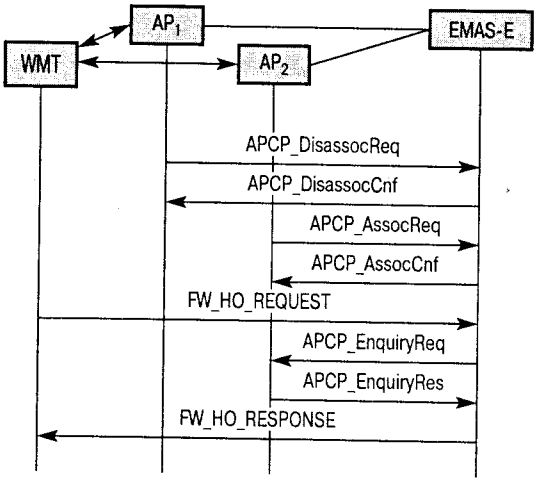

Рис. 8.15. Прямое переключение Intra-EMAS-E/Inter-AP

8.7.6.3. Прямое переключение Inter-EMAS-E/Inter-AP

Третий вариант переключения, Inter-EMAS-E/Inter-AP, снова является наиболее сложным (см. рис. 8.16). Вплоть до сообщения FW_HO_REQUEST, передаваемого с WMT на новый EMAS-E, ситуация аналогична переключению Intra-EMAS-E/Inter-AP. $EMAS\text{-}E_2$ ничего не знает о том, за что отвечает старый EMAS-E. Следовательно, сообщение FW_HO_REQEST содержит идентификатор старого трансивера, список текущих активных соединений, их параметры QoS и информацию о возможных коммутаторах COS. Опираясь на информацию об трансивере, $EMAS\text{-}E_2$ может отыскать старый EMAS-E этого WMT и затем отправить сообщение HO_NOTIFY, чтобы проинформировать старый EMAS-E об инициации переключения.

После проверки ресурсов путем обмена сообщениями APCP_EnquiryReq и APCP_EnquiryRes, $EMAS\text{-}E_2$ посылает WMT сообщение FW_HO_RESPONSE, указывая, что соединение может быть переключено. Для каждого из соединений, перечисленных WMT, нужно проделать следующие шаги: $EMAS\text{-}E_2$ устанавливает соединение с COS (SETUP, CONNECT) и отправляет WMT сообщение CONN_ACTIVE; COS информирует старый EMAS-E о завершении переключения, используя HO_COMPLETE; теперь $EMAS\text{-}E_1$ разрывает соединение с COS (RELEASE), который подтверждает разъединение (RELEASE_COMPLETE).

Как и ранее, каждую операцию контролирует таймер, а сообщение HO_FAILURE указывает на сбой при переключении. В то же время, при прямом соединении отсутствует возможность соединения с несколькими точкам доступа или EMAS-E. WMT сам решает, какой трансивер использовать и, следовательно, определяет новую точку доступа и EMAS-E.

8.8. Управление местоположением

Во всех сетях с поддержкой мобильности для поиска текущего положения мобильного терминала, для предоставления движущемуся терминалу постоянного адреса и для гарантий безопасности (конфиденциальность, аутентификация или авторизация) требуется наличие специальных функций. Эти и другие функции объединяются термином **управление местоположением**.

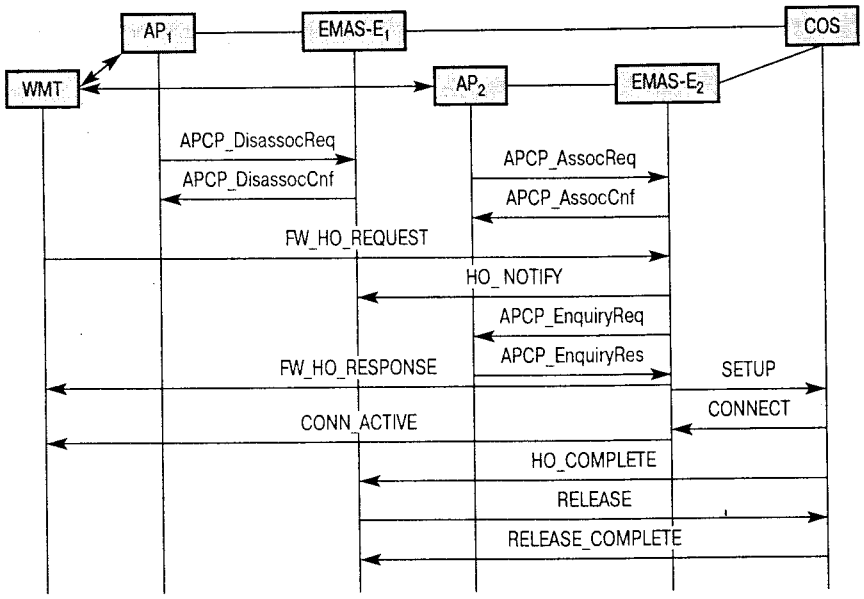

Рис. 8.16. Прямое переключение Inter-EMAS-E/Inter-AP

8.8.1. Требования к управлению местоположением

Было установлено несколько требований к управлению местоположением [Bhat, 1998б].

- **Незаметность мобильности**. При нормальной работе пользователь не должен замечать функцию управления местоположением. Все изменения местоположения должны происходить без вмешательства пользователя. Это накладывает некоторые ограничения на допустимую задержку функций, связанных с управлением местоположением. Более того, должен быть возможен незаметный роуминг между различными доменами (частный/частный, частный/общественный, общественный/общественный). Это может включать и роуминг между сетями, созданными на основе различных технологий, используя, например, терминал с двумя режимами работы.
- **Безопасность** Для обеспечения достаточно высокого уровня безопасности, приемлемого в ответственных системах (коммерция, аварийные ситуации и т.д.), системы WATM требуют наличия специальных функций. Вся информация о местоположении и пользовательские данные, собранные для управления местоположением и наблюдения за ней, должны быть защищены от несанкционированного разглашения. Эта защита чрезвычайно важна для профилей роуминга, которые позволяют точно отслеживать отдельные терминалы. Поскольку доступ к радиоинтерфейсу получить очень просто, нужно реализовать специальные ограничения на доступ, чтобы, например, не допустить пользователей сетей общего пользования к частным сетям WATM. Однако пользователь должен быть также в состоянии определить, к каким сетям имеет доступ его терминал. В число необходимых функций безопасности входят аутентификация пользователей, терминалов, а также точек доступа. Кроме того, необходимым является шифрование, по крайней мере, между терминалом и точкой доступа, хотя более желательным было бы сквозное шифрование.
- **Эффективность и масштабируемость.** Представим сети WATM с миллионами пользователей, подобные современным сетям мобильных телефонов. Каждая функция и система, задействованные в управлении местоположением, должна быть масштабируемой и эффективной. Это касается распределенных служб для хранения информации о местоположении,

учета и аутентификации. Кроме того, характеристики всех операций практически не должны зависеть от размеров сети, количества текущих соединений и загруженности сети. Таким образом, для увеличения общей производительности путем разделения нагрузки должно быть возможно разбитие коммутаторов по кластерам и введение иерархии доменов. В отличие от многих существующих сотовых сетей, WATM должна работать с более эффективными интегральными схемами передачи данных. Любая передача данных, требующая управления местоположением, должна, следовательно, внедряться в существующие механизмы передачи данных путем добавления к существующим сообщениям новых информационных элементов. Это позволяет использовать весьма эффективные существующие механизмы передачи данных в стационарных сетях ATM.

- **Идентификация**. Управление местоположением должно предусматривать методы идентификации всех объектов сети. Радиоячейки, сети WATM, терминалы и коммутаторы должны обладать уникальными идентификаторами и механизмами обмена информацией опознавания. Это требование включает также предоставления терминалу информации касательно его текущего положения (базовая сеть или чужая) и текущей точки подключения. Кроме того, помимо постоянного **адреса конечной системы ATM (ATM end system address — AESA)** терминалу, как только он выходит за пределы базовой сети, требуется маршрутизируемый временный адрес AESA. Этот временный адрес должен отправляться по базовому адресу терминала.
- **Обеспечение межсетевого обмена и стандарты**. Все функции управления местоположением должны взаимодействовать с существующими функциями ATM стационарной сети, в особенности касающимися маршрутизации. Кроме того, управление местоположением в WATM должно согласовываться с другими схемами, такими как управление местоположением в ныне существующих сетях GSM, будущих сетях UMTS, Internet с использованием Mobile IP или Intranet со специальными функциями. Эта согласованность может, например, привести к двухуровневому управлению местоположением, если Mobile IP используется над уровнем WATM. Все протоколы, использующиеся в WATM для обновления баз данных, регистрации и т.д., должны быть стандартизованы, чтобы допускать мобильность в пределах областей обслуживания различных поставщиков сетевых услуг. Однако в пределах административного домена могут применяться и личные улучшения и схемы оптимизации.

8.8.2. Процедуры и объекты

Было установлено несколько объектов, которые поддерживают управление местоположением. Мобильный терминал, который для рассматриваемых здесь приложений является беспроводным (WMT), реализовывает функции по инициации обновления местоположения и участвует в аутентификации и протоколах обеспечения конфиденциальности. Для идентификации сообщений установки соединения, предназначенных WMT, и для включения функций разрешения местоположений, требуется много коммутаторов, поддерживающих мобильность конечного пользователя (end-user mobility-supporting switch — EMAS). Помимо этого, EMAS перенаправляют соединения, а некоторые из них поддерживают серверы местоположения и аутентификации.

Серверы местоположения (location server — LS) поддерживают базы данных, содержащие текущий адрес AESA и постоянный адрес AESA терминала WMT. Они также отслеживают параметры терминала и пользователя, относящиеся к данной службе, такие как качество обслуживания соединений. Серверы местоположения могут быть частью EMAS. Для каждого WMT **сервер аутентификации (authentificaion server — AUS)** поддерживает безопасность баз данных с информацией, касающейся аутентификации и конфиденциальности. Это включает также права доступа и идентификацию оборудования.

На рис. 8.17 показан поток сообщений обновления регистрации и местоположения после перемещения WMT в новую точку доступа. Для идентификации сети, текущего местоположе-

ния и радиотрансивера точка доступа транслирует сообщение Broadcast_ID. WMT может теперь решить связаться с точкой доступа, если это согласуется с его правами доступа. Следующий этап включает отправку регистрационного сообщения на посещаемый EMAS-E, которое содержит адрес EMAS базовой сети, и далее запрашивается временный адрес AESA. Посещаемый EMAS-E выдает теперь EMAS базовой сети терминала WMT запрос Loc_Update_Home. Затем EMAS базовой сети проверяет права доступа и идентифицирует WMT посредством сообщения Auth_Req. Если отклик Auth_Req_Reply указывает, что все в порядке, EMAS базовой сети вводит в LS базовой сети новое местоположение. LS подтверждает прием вместе с EMAS базовой сети. Посещаемый EMAS-E возвращает временный адрес AESA терминалу WMT в сообщении Loc_Update_Reply.

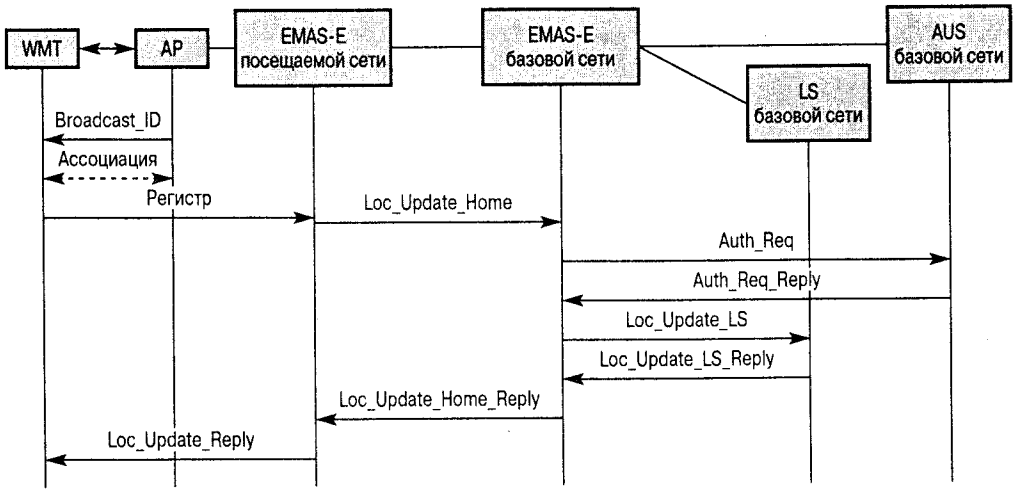

Рис. 8.17. Обновление регистрации и местоположения

Основные этапы установления соединения, инициированного произвольным терминалом ATM, отображены на рис. 8.18. В этом сценарии участвуют произвольная сеть (с поддержкой или без поддержки мобильности), **базовая сеть** WMT и сеть, посещаемая в данный момент. Базовая сеть — это сеть, которой принадлежит WMT согласно постоянному адресу. Базовая сеть содержит объект EMAS-E, на который направляются все соединения для WMT (в этом примере — EMAS-E_2). Все другие сети являются **чужими сетями** для WMT, чужая сеть, посещаемая в данный момент, является **посещаемой сетью**. В этом примере для определения местоположения WMT и установления соединения требуются следующие шаги.

1. Терминал T не обладает какой-либо информацией относительно мобильности или текущего местоположения WMT. T знает только постоянный адрес WMT и, следовательно, настраивает соединение с WMT. Коммутатор, отвечающий за WMT (EMAS-E_2), обнаруживает входящее соединение установления.
2. Теперь EMAS-E_2 должен найти текущее местоположение WMT, посылая запрос местоположения на сервер LS. LS содержит отображение постоянных адресов всех WMT, за которые он отвечает, в текущие адреса чужой сети. Если нужно провести несколько этапов иерархического поиска, LS передает постоянный адрес коммутатора EMAS посещаемой сети.
3. LS возвращает либо текущий адрес WMT, либо текущий адрес EMAS, ответственного за дальнейший поиск. Этот иерархический подход оправдан, если требуется избежать слишком частого обновления LS. Обычно WMT остается в определенной области, и, следова-

тельно, имеется высокая вероятность того, что WMT остается в пределах одной и той же сети. Если WMT перемещается в пределах этой сети, нет необходимости информировать LS базовой сети. WMT можно достичь, соединяясь с EMAS, отвечающим за дальнейший поиск в посещаемой сети.

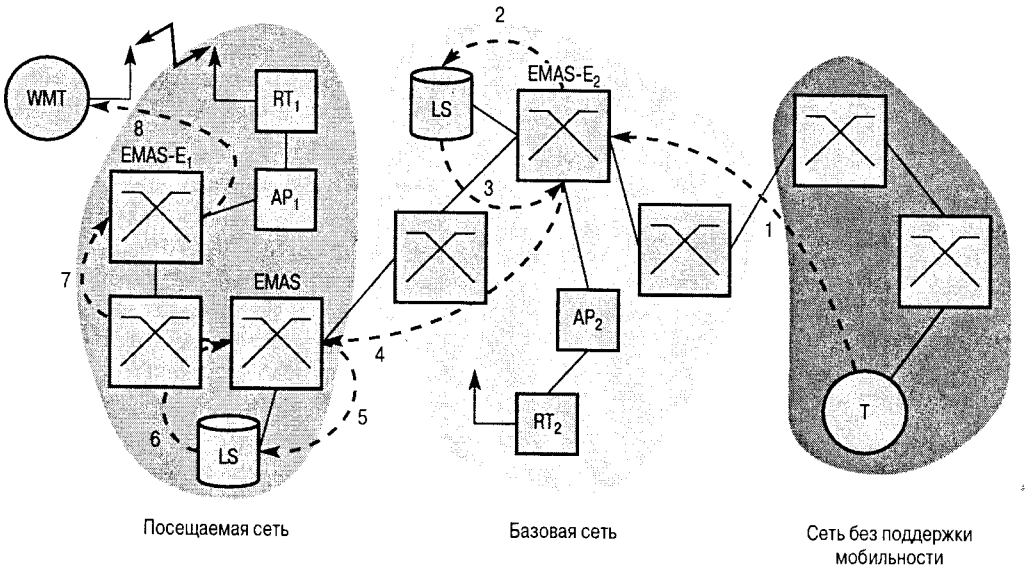

Рис. 8.18. Установление соединения с WMT в чужой сети

4. EMAS-E_2 теперь перенаправляет входящее соединение на EMAS посещаемой сети. Перенаправление может выполняться без разрыва соединения, то есть, запрос установления соединения передается с EMAS-E_2 вдоль нового пути. Для оптимизации маршрута соединения может использоваться частичный разрыв. Это означает, что некоторая часть соединения будет отключена и можно будет выбрать более короткое соединение. Это уменьшит неэффективность треугольной маршрутизации (т.е. наличия маршрута соединения, начинающегося на терминале, проходящего через EMAS-E_2 и заканчивающегося на WMT).
5. EMAS, отвечающий за сеть, посещаемую в данный момент, может очень точно определить местоположение WMT. Следовательно, он отправляет запрос на свой LS. Данный EMAS может быть либо EMAS-E, либо EMAS-N.
6. LS выдает в ответ точное текущее местоположение, которое представлено в его базе данных.
7. Теперь запрос установления соединения можно направить на EMAS-E_1, который непосредственно связан с точкой доступа, с которой также связан и WMT.
8. Последним шагом является установление связи с самим WMT посредством точки доступа и трансивера.

Для обновления информации о местоположении, регистрации и установления соединения объектам WATM необходима, по крайней мере, нижеследующая информация.

- **WMT**. Постоянные данные, сохраненные в WMT, содержат уникальный адрес базовой сети, соотнесенный с устройством, ключ для аутентификации терминала и перечень прав доступа. Временными данными является текущий адрес устройства (присваиваемый при регистрации) и трансиверов, с которыми в настоящий момент связан WMT. Еще не определено,

сколько пользовательской информации должно храниться на WMT. Информация может размещаться в микросхеме-ключе (как в GSM), хранящей ключ пользователя и другие пользовательские параметры (например, часто используемые адреса).

- **LS**. Чтобы отобразить адрес WMT в базовой сети в его текущий адрес LS базовой сети требуется таблица, содержащая статус текущего местоположения. Кроме того, для каждого пользователя может храниться профиль, включающий права пользователя. Учетная информация и информация о счете для каждого пользователя может храниться либо LS, либо отдельным сервером.
- **AUS**. Сервер AUS с целью идентификации должен хранить базовые адреса для каждого WMT (если для этого используется адрес в базовой сети) и ключ, который должен совпадать с ключом, хранящимся в WMT (требуется для аутентификации).
- **EMAS**. Каждый коммутатор EMAS хранит адрес своего сервера местоположения, поскольку этот адрес необходим для управления запросами на установку входящего соединения. Кроме того, требуется адрес сервера аутентификации. Если EMAS находится на краю сети (EMAS-E), то тогда он может дополнительно хранить список принадлежащих ему терминалов (то есть, зарегистрированных в данный момент и соединенных посредством точки доступа).

8.9. Адресация

Основные требования к адресам WATM определены в работе [Bhat, 1998б].

- WATM должен поддерживать все существующие на данный момент форматы адресов конечных систем ATM (AESA), однако можно применять и адреса со специальными структурами для мобильных терминалов.
- WMT должны иметь постоянные, т.е. независящие от местоположения, адреса. Этот адрес известен WMT и может использоваться другими сообщающимися сторонами для установления контакта с WMT. Кроме того, этот адрес должен соответствовать адресу, маршрутизируемому из "базовой сети", то есть, он должен быть соотнесен с EMAS в базовой сети.
- WMT должен поддерживать присваивание при регистрации в чужой сети временных, маршрутизируемых адресов. Формат временных и постоянных адресов может отличаться.

Основываясь на этих требованиях, спецификация WATM предлагает присваивать WMT два адреса. Один адрес является постоянным, уникальным, не зависящим от местоположения, и его основу составляет маршрутизируемый адрес ATM, соотнесенный с EMAS в базовой сети терминала WMT. Другой адрес является маршрутизируемым адресом в чужой сети, соотнесенным с EMAS-E, с которым WMT соединен в данный момент. Должен существовать автоматический механизм, при регистрации предоставляющий WMT временный адрес.

8.10. Мобильное качество обслуживания

Гарантии качества обслуживания (quality of service — QoS) являются одним из основных преимуществ сетей WATM по сравнению, например, с сетями пакетной радиосвязи Mobile IP. В то время, как протокол IP не гарантирует QoS, это делают сети ATM (ценой увеличения сложности). Следовательно, сети WATM должны обеспечивать мобильное качество обслуживания (mobile QoS — M-QoS). M-QoS состоит из трех различных частей.

- **Проводное QoS**. Инфраструктурная сеть, требуемая WATM, имеет те же свойства QoS, что и любая проводная сеть ATM. Типичными традиционными параметрами QoS являются задержка канала, дисперсия задержек по ячейкам, пропускная способность, частота появления ошибок в ячейке и т.д.

- **Беспроводное QoS**. Свойства QoS беспроводной части сети WATM отличаются от свойств проводной части. Как и ранее, можно задать задержку канала и частоту появления ошибок, однако теперь, как правило, частота появления ошибок на несколько порядков выше, чем, например, в волоконно-оптических каналах связи. Кроме того, механизмы резервирования каналов и уплотнения в радиодиапазоне сильно влияют на дисперсию задержек по ячейкам.
- **QoS переключения**. Новый набор параметров QoS вводится переключением. Здесь критическими факторами для QoS являются блокирование переключения вследствие ограниченности ресурсов в целевых точках доступа, потеря ячеек при переключении или скорость процедуры переключения в целом.

Система WATM должна при установлении соединения отображать QoS, заданное приложением, в эти наборы параметров QoS и должна проверять, можно ли удовлетворить затребованное QoS. Однако приложения будут подробно определять не одиночные параметры, а суммарные требования, такие как задержка или пропускная способность. Системы WATM должны теперь отображать, например, суммарную задержку в задержку ячейки в каждом сегменте, проводном и беспроводном. В связи со сложностью таких систем, сети WATM будут поначалу предлагать приложениям только набор различных классов служб.

Помимо этого, приложения должны быть достаточно адаптируемыми, чтобы выдержать следствия мобильности, в частности, значительные потери ячеек, значительную дисперсию задержек и т.д. Приложения могут, например, согласовывать диапазоны параметров, в пределах которых они смогут адаптироваться без разрыва соединения.

Решающим моментом в постоянном поддержании качества обслуживания является поддержка QoS в протоколах переключения. При переключении эти протоколы могут поддерживать два нижеследующих типа QoS.

- **QoS жесткого переключения**. Хотя благодаря ныне доступным ресурсам можно гарантировать QoS с текущей точкой доступа, но гарантировать QoS после переключения нельзя. Это можно, например, сравнить с традиционным случаем сетей GSM с голосовыми соединениями. Если терминал может установить соединение, то качество этого соединения гарантируется. Если после переключения ресурсов становиться недостаточно (в целевой ячейке находится уже слишком много пользователей), система разрывает соединение. Это является единственно возможным решением, если приложения и терминалы не могут адаптироваться к новой ситуации.
- **QoS мягкого переключения**. Даже для текущего беспроводного сегмента можно только статистически гарантировать QoS и помимо этого приложения должны еще и адаптироваться после переключения. Это предполагает наличие адаптирующихся приложений и, по крайней мере, позволяет сохранить некоторые гарантии QoS в течение периода перегрузки или, например, сильных помех.

Если терминал переключается в точку доступа со слабыми ресурсами, возможны следующие шаги: терминал может закрыть одно или несколько соединений, чтобы освободить ресурсы или он может попытаться адаптировать QoS одного или нескольких соединений. Помимо этого, точка доступа может отказаться от соединения в процессе переключения, поскольку WMT предоставляет список переключаемых соединений (см. рис. 8.13).

8.11. Протокол управления точками доступа

Как показано на рис. 8.10–8.16, каждая точка доступа обменивается множеством сообщений с объектом EMAS-E, соединенным с этой точкой. Протокол, используемый для связи этих двух объектов, называется **протоколом управления точками доступа (access point control protocol — APCP)**. Основная задача APCP заключается в резервировании и освобождении ресурсов в точке

доступа для соединений (память, радиочастоты и т.д.). Кроме того, посредством APCP коммутатор может подготовить точку доступа к новым соединениям и содействовать переключению. Точка доступа использует протокол для информирования EMAS-E о новоассоциированных WMT.

Существует следующие две причины разнесения точки доступа и EMAS-E и стандартизации интерфейса между ними (ср. также рис. 8.2 и рис. 8.7).

- Все радиопередачи осуществляются отдельно от коммутатора. Коммутатор нужен только для поддержки мобильности и, следовательно, его можно использовать только в случае мобильности без беспроводного доступа.
- Передача данных на более высоких уровнях (например, передача ATM) не зависит от передачи, используемой для точки доступа, то есть, APCP. Следовательно, точка доступа может также использоваться в архитектурах, не использующих передачу ATM.
- Разделение и стандартизованный интерфейс допускают разнообразные решения, то есть, коммутатор и точка доступа могут происходить от различных поставщиков и, следовательно, конечные пользователи могут использовать все выгоды большей гибкости и конкуренции.

На рис. 8.19 показана функциональная архитектура процессов, связанных с APCP. Точка доступа содержит процессы, непосредственно связанные с радиопередачей, такие как **контроль доступа беспроводного соединения (WCAC)** и **управление радиоресурсами (RRM)**. Кроме того, процесс **управления соединением точки доступа (APCM)** контролирует текущие соединения. Коммутатор не содержит каких-либо специфических для беспроводной связи процессов, а управляет мобильностью. Это достигается путем использования четырех процессов: **управления ресурсами (RM)**, **управления вызовами (CC)**, **управления доступом соединения (CAC)** и **управления мобильностью (MM)**.

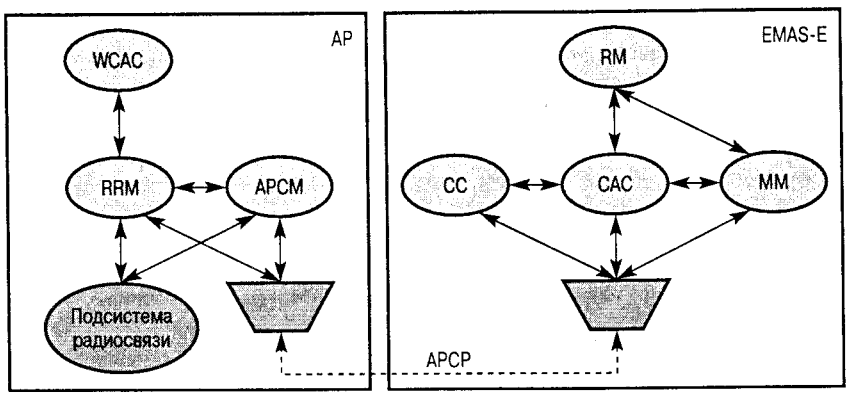

Рис. 8.19. Функции APCP в точке доступа (AP) и EMAS-E

Для переключения используются нижеследующие сообщения (см. рис. 8.10–8.16). Дальнейшие сообщения специфицированы аналогично изложенному в работе [Bhat, 1998a].

- **APCP_AssocReq**. Точка доступа информирует коммутатор о новой связи с WMT. В этом контексте связью называется только установление радиосоединения между WMT и точкой доступа. Это, например, случай, когда EMAS-E требует, чтобы точка доступа отсоединила WMT, например, если WMT не был авторизирован для доступа к этой точке доступа или регистрация не завершилась успешно по каким-либо другим причинам. Это сообщение нужно, по крайней мере, AESA терминала WMT.
- **APCP_AssocCnf**. Коммутатор подтверждает связь с точкой доступа. Это сообщение содержит WMT-AESA для идентификации и идентификатор соединения ATM для соединения.

- **ATCP_DisassocReq**. Этим сообщением точка доступа указывает, что WMT отсоединился. Это сообщение, по меньшей мере, содержит WMT-AESA.
- **APCP_DisassocCnf**. Это сообщение подтверждает APCP_DisassocReq и содержит WMT-AESA.
- **APCP_EnquiryReq**. EMAS-E посылает это сообщение для резервирования ресурсов для нового соединения в точке доступа. Это сообщение содержит несколько параметров, таких как WMT-AESA, идентификатор соединения ATM и дескрипторы информационного потока ATM, а также параметры QoS для входящей и исходящей линий связи (к WMT и от WMT).
- **APCP_EnquiryRes**. Здесь точка доступа сообщает, доступны ли требуемые ресурсы. Параметры сообщения включают в себя WMT-AESA, идентификатор соединения и перечень доступных в настоящее время ресурсов.

8.12. Резюме

В этой главе дан обзор работ по стандартизации, проведенных на данный момент рабочей группой WATM форума ATM Forum. WATM — это подход, направленный на расширение схемы ATM для связи, ориентированной на установление соединения, с предоставлением QoS в беспроводной и мобильной области. Основным отличием между WATM и другими подходами является интеграция в спецификацию всей системы. WATM определяет радиодоступ, управление мобильностью, схемы переключения, QoS мобильной связи, безопасность и т.д.

Кроме того, в главе обсуждались мотивы разработки WATM и перечислены службы WATM, которые должны будут предоставляться. Эти службы строятся на множестве новых функций, требуемых для поддержки мобильности в стационарной сети и на мобильных и беспроводных терминалах. Поскольку многие проблемы в WATM аналогичны встречающимся в других беспроводных сетях, ATM Forum пытается согласовывать свою работу с другими стандартами. Например, WATM может использовать технологию BRAN, которая стандартизована ETSI как уровень радиодоступа.

Однако, основная сложность WATM определяется функциями и протоколами, необходимыми для переключения. В этом контексте было представлено несколько сценариев, демонстрирующих сообщения, параметры и объекты, задействованные в переключении. Сложность WATM вызвана его возможностью сохранения параметров QoS для соединений при переключении и принципом ориентации на соединение (ATM). Вследствие таких особенностей возникает необходимость резервирования ресурсов, проверки доступности ресурсов в точке доступа и изменения маршрута соединений.

Поскольку WATM представляет собой интегральный подход, то в нем должны рассматриваться такие задачи, как управление местоположением, безопасность и эффективность системы в целом. Для минимизации количества служебных сигналов WATM пытается согласовать требуемые функции с теми, которые доступны в стационарной сети ATM. На рис. 8.20 показаны другие возможные сценарии доступа WATM. В то время, как в этой главе основное внимание было уделено беспроводному доступу мобильных терминалов ATM, возможно и несколько других конфигураций, которые обсуждаются в нестоящее время [Bhat 1998б]. В качестве дополнительных объектов на рис. 8.20 показаны следующие компоненты.

- **T (терминал)**. Стандартный терминал ATM, предлагающий услуги ATM для стационарных сетей ATM.
- **MT (мобильный терминал)**. Стандартный терминал ATM с дополнительной возможностью повторного соединения после изменения точки доступа, т.е. терминал может перемещаться между различными точками доступа в пределах определенного домена.
- **WT (беспроводный терминал)**. Доступ к этому терминалу получается посредством беспроводной линии связи, однако терминал сам по себе стационарен, т.е. терминал сохраняет свою точку доступа к сети.

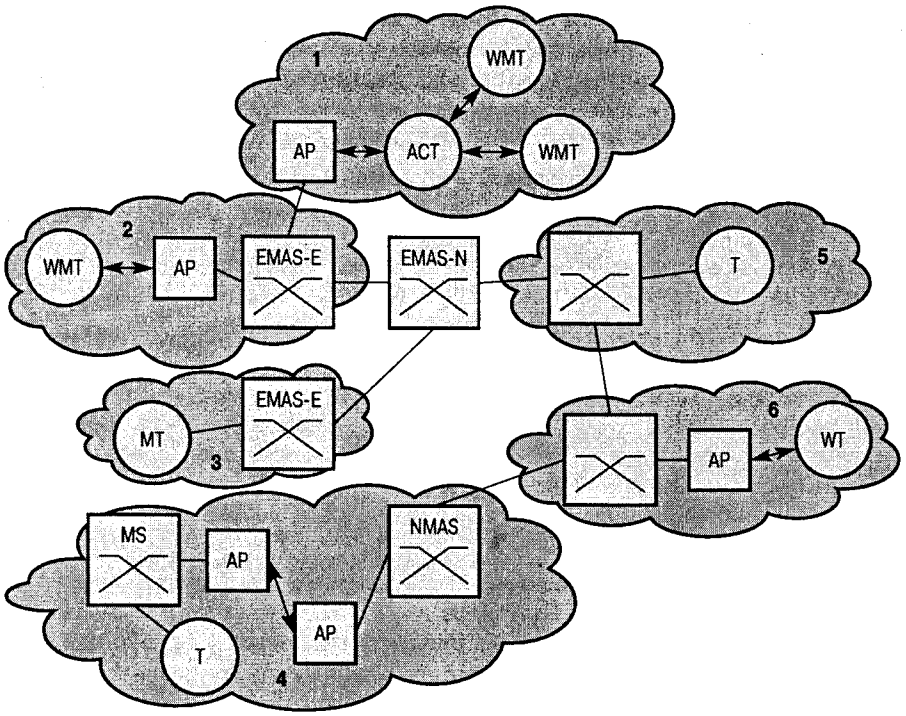

Рис. 8.20. Эталонная модель WATM, содержащая дополнительные сценарии доступа

- **WMT (беспроводный мобильный терминал)**. Комбинация беспроводного и мобильного терминалов дает в результате WMT. Это как раз тот тип терминала, который рассматривался в этой главе, поскольку он имеет возможность изменять свою точку доступа и использует радиодоступ.
- **AP (точка доступа)**. Как объяснялось в этой главе, это — точка доступа к сети через радиоканал.
- **EMAS (коммутатор ATM, поддерживающий мобильность конечного пользователя, -E: объект, -N: сеть)**. Коммутатор, который поддерживает мобильность конечного пользователя, см. 8.4.1.
- **NMAS (коммутатор ATM, поддерживающий мобильность сети)**. Мобильными могут быть не только терминалы, но также и сеть в целом. Следовательно, для поддержки такой мобильности из стационарной сети необходимы определенные дополнительные функции.
- **MS (мобильный коммутатор ATM)**. Коммутатор ATM также может быть мобильным и использовать беспроводный доступ к другим частям сети ATM.
- **ACT (терминал эпизодического контроллера)**. В конфигурации эпизодической сети в рамках беспроводной сети могут потребоваться специальные типы терминалов. Эти терминалы могут, например, контролировать беспроводный доступ без точки доступа.

На основе этих объектов можно определить несколько сценариев, которые должны поддерживаться WATM, при условии полной спецификации.

- **Беспроводная эпизодическая сеть ATM (сценарий 1)**. WMT могут связываться друг с другом, не требуя при этом наличия стационарной сети. Связь может устанавливаться без какой-либо инфраструктуры. Управление доступом может осуществляться посредством ACT. Если эпизодическая сеть нуждается в соединении со стационарной сетью, то это можно осуществлять посредством точки доступа. Дальнейшая спецификация этой конфигурации еще не проводилась.

- **Беспроводные мобильные терминалы ATM (сценарий 2).** В этой главе обсуждалась конфигурация с беспроводными мобильными терминалами, получающими доступ к стационарной сети посредством точки доступа. В этой конфигурации WMT не могут устанавливать связь без поддержки, предоставляемой объектами стационарной сети, такими как EMAS-E. В настоящее время этот сценарий специфицируется подробнее.
- **Мобильные терминалы ATM (сценарий 3).** Этот сценарий поддерживает портативность устройств и допускает простую перестройку структуры сети. Пользователи могут со временем изменять точки доступа своего оборудования ATM без необходимости перестраивать структуру вручную. Снова этот сценарий нуждается в поддержке объектов стационарной сети (например, EMAS-E).
- **Мобильные коммутаторы ATM (сценарий 4).** Еще более сложная конфигурация содержит мобильные коммутаторы, использующие беспроводный доступ к другим стационарным сетям ATM. Теперь в пределах стационарной сети нужны объекты, поддерживающие мобильность коммутаторов (NMAS). Для этого сценария существует много приложений, например, сети в самолетах, поездах или на кораблях. В пределах мобильной сети можно использовать как стационарные терминалы, так и мобильные, беспроводные или мобильные и беспроводные одновременно. Это наиболее сложная конфигурация, которая в настоящее время обсуждается в среде ATM.
- **Стационарные терминалы ATM (сценарий 5).** Эта конфигурация является стандартным случаем, который уже используется. Терминалы и коммутаторы не имеют возможности получать мобильный и беспроводный доступ. Эта конфигурация является также эталонной для приложений, работающих над уровнем сети ATM. Следовательно, уровень конвергенции должен скрывать специфические характеристики беспроводного и мобильного доступа, поскольку никакие специальные приложения не требуют представленного здесь сценария.
- **Стационарные беспроводные терминалы ATM (сценарий 6).** Для предоставления простого доступа к сетям ATM без прокладывания проводов идеальным решением являются стационарные беспроводные каналы связи. Для использования такого способа доступа к потребителям, не имеющим проводной инфраструктуры, используется и планируется множество альтернативных носителей. Этот сценарий не требует внесения каких-либо изменений или усовершенствований в стационарную сеть.

В настоящее время еще не ясно, какая технология или схема доступа будет наиболее успешной. Существует множество приложений для каждого из показанных сценариев; однако та или иная схема может оказаться удачной в зависимости от доступности служб и рыночных механизмов. С технической точки зрения уже совершенно ясно, что комбинация всех сценариев, как показано на рис. 8.20, является наиболее сложной системой связи, которая, возможно, слишком претенциозна для того, чтобы реализоваться в ближайшем будущем. Все конфигурации должны быть в состоянии взаимодействовать с существующими сотовыми системами и Internet-технологиями. В работе [Chakraborty, 1998] рассматриваются многие проблемы, возникающие при обеспечении межсетевого обмена с узкополосными сетями, такими как GSM, DECT, UMTS (см. главу 4) и стандартными сетями TCP/IP.

8.13. Упражнения

1. Назовите причины развития беспроводной ATM и институты, компании и организации, работающие над стандартизацией. Какие с этой точки зрения основные отличия между ATM и Internet-технологиями?
2. Сравните ATM и TCP/IP и укажите основные отличия. Приведите аргументы за и против обоих подходов и рассмотрите их влияние на мобильность и беспроводный доступ.

3. Какие дополнения к конфигурации ATM нужны для поддержки мобильности и беспроводного доступа?
4. Какие типы услуг для конечного пользователя предусматриваются в WATM? Сравните эти услуги с идеей беспроводных ЛВС, представленных в главе 7 и беспроводных телекоммуникационных систем, рассматриваемых в главе 4.
5. Вспомните основные эталонные модели WATM. Какие были определены объекты? Почему в спецификации разделены терминал и беспроводный адаптер?
6. Какие типы поддержки мобильности и беспроводного доступа необходимы в стационарных сетях? Покажите влияние поддержки мобильности на плоскость пользователя и плоскость управления.
7. Зачем WATM специфицирует отдельный протокол APCP? Назовите причину обособления точки доступа?
8. Каковы основные требования к RAL для WATM? Какие были определены уровни в рамках RAL? На что влияет предоставляемое качество обслуживания?
9. Какая часть BRAN может использоваться в WATM? Какие еще типы сетей определяет BRAN и какие более высокие уровни она может поддерживать?
10. Почему переключение так важно для WATM? Рассмотрите основные требования переключения и сравните их с другими беспроводными ЛВС/персональными сотовыми сетями и их сценариями переключения.
11. Каковы отличия между медленным затуханием радиосигналов и неожиданной потерей радиосоединения с точки зрения переключения? Какие проблемы могут быть связаны с неожиданной потерей канала радиосвязи?
12. Почему WATM исходно рассматривает только жесткое переключение? Каковы последствия такого выбора?
13. Сравните сложность, задействованные объекты и передачу служебных сигналов в трех представленных сценариях переключения.
14. Назовите компоненты управления местоположением и его функции, а также укажите, зачем они нужны.
15. Каковы функции процедуры регистрации и зачем нужна регистрация в дополнение к переключению?
16. Опишите процедуру исходного установления соединения с WMT в чужой сети. Назовите причину определения иерархии серверов местоположения? Как проходит установка соединения в базовой сети?
17. Почему качество обслуживания так важно в сетях WATM? Какие различные типы QoS должны поддерживаться в сетях WATM?
18. В этой главе основное внимание уделялось сценарию с терминалами WMT и точками доступа, но, как сказано в резюме, возможны и иные сценарии. Найдите дополнительные сферы применения каждого из сценариев и сравните их сложности. При рассмотрении основное внимание обратите на переключении в мобильной сети ATM, например, в самолетах с соединением через спутник.
19. Укажите причины успеха технологии WATM и попытайтесь определить причины, которые могли бы препятствовать успеху. Проверьте текущую ситуацию в этой области с помощью WWW-сервера форума ATM Forum.

8.14. Рекомендуемая литература

Acampora, A. (1996) "Wireless ATM: a perspective on issues and prospects", *IEEE Personal Communications*, 3, (4), August.

Akyol, B. A.; Cox, D. C. (1996) "Rerouting for handoff in a wireless ATM network", *IEEE Personal Communication*, 3, (5), October.

Anastasi, G.; Lenzini, L.; Mingozzi, E; Hettich, A.; Krämling, A. (1998) "MAC protocols for wideband wireless local access: evolution toward wireless ATM", *IEEE Personal Communications*, 5, (5).

ATM Forum (1996) *ATM traffic management specification version 4.0*, ATM Forum.

ATM Forum (1998) *Framework for dynamic COS discovery in Wireless ATM*, ATM98-0005, ATM Forum, February.

ATM Forum (1998) *Extension to backward COS discovery (BCD) approach* ATM98-0231, ATM Forum, April.

Ayanoglu, E.; Eng, K. Y.; Karol, M. J. (1996) "Wireless ATM: limits, challenges, and proposals", *IEEE Personal Communication*, 3, (4), August.

Barton, M.; Paine, R.; Chow, A. (1998) *Description of Wireless ATM Capability Set 1 Specification*, ATM Forum, BTD-WATM-01.09.

Bhat, R. R. (1998) *Draft baseline text for Wireless ATM Capability Set 1 Specification*, ATM Forum, BTD-WATM-01.09.

Bhat, R. R. (1998) *Wireless ATM Requirements Specification*, ATM Forum, RTD-WATM-01.02.

Chakraborty, S. S. (1998) "The interworking approach for narrowband access to ATM transport-based multiservice mobile networks", *IEEE Personal Communications*, 5, (4).

ETSI (1997) *Radio Equipment and Systems (RES), High Performance Radio Local Area Networks (HIPERLAN), Requirements and architectures for wireless ATM access and interconnection*, European Telecommunication Standards Institute, TR 101 031 V1.1.1 (1997-07).

ETSI (1998a) *Broadband Radio Access Networks (BRAN): Inventory of broadband radio technologies and techniques*, European Telecommunication Standards Institute, TR 101 173 V1.1.1 (1998-05).

ETSI (1998б) *Broadband Radio Access Networks (BRAN): High Performance Radio Local Area Network (HIPERLAN) Type 1, Functional specification*, European Telecommunication Standards Institute, EN 300 652 V1.2.1 (1998-07).

ETSI (1998в) *Broadband Radio Access Networks (BRAN): Requirements and architectures for broadband fixed radio access networks* (HIPERACCESS), European Telecommunication Standards Institute, TR 101 177 V1.1.1 (1998-05).

Händel, R.; Huber, N.; Schröder, S. (1994) *ATM Networks: concepts, protocols, applications*, Addison-Wesley.

Petras, D.; Krämling, A (1997) "Wireless ATM: Performance evaluation of a DAS++ MAC protocol with fast collision resolution by a probing algorithm", *International Journal on Wireless Information Networks*, 4, (4).

Rauhala, K. (1998) *Baseline Text for Wireless ATM specifications*, ATM Forum, BTD-WATM-01.07.

Rauhala, K. (1998) *Living list document of Wireless ATM working group*, ATM Forum, LTD-ATM-01.07.

Raychaudhuri, D.; Dellaverson, L. (1996) *Charter, scope and work plan for proposed wireless ATM working group*, ATM Forum Document, 96-0530.

Raychaudhuri, D. (1996) "Wireless ATM networks: architecture, system design and prototyping", *IEEE Personal Communications*, vol. 3, no. 4, August.

Sanchez, J; Martinez, R.; Marcellin, M. W. (1997) "A survey of MAC protocols proposed for wireless ATM", *IEEE Network Magazine*, November/December.

Toh, C.-K. (1997) *Wireless ATM and ad hoc networks*, Kluwer Academic Publishers.

Umehira, M.; Nakura, M.; Sato, H.; Hashimoto, A. (1996) "ATM wireless access for mobile multimedia: concept and architecture", *IEEE Personal Communications*, 3, (5), October.

Walke, B.; Petras, D.; Plassmann, D. (1996) "Wireless ATM: air interface and network protocols of the mobile broadband system", *IEEE Personal Communications*, 3, (4), August.

Мобильный сетевой уровень 9

В этой главе вводятся протоколы и механизмы, разработанные для поддержки мобильной связи на сетевом уровне. Самый известный пример подобных протоколов — Mobile IP, который рассмотрен в первом разделе и который вводит поддержку мобильности в протокол сетевого уровня IP. Тогда как системы, подобные GSM, изначально разрабатывались как протоколы с поддержкой мобильности, Internet возникла тогда, когда еще никто не предполагал широкого распространения мобильных компьютеров. Таким образом, в современной Internet отсутствуют механизмы поддержки пользователей, путешествующих по всему миру. Протокол IP — это общая основа для работы тысяч приложений в десятках различных сетей. Это и является причиной поддержки мобильности на уровне IP; мобильные телефонные системы, например, не могут предлагать такой тип мобильности для неоднородных сетей.

Другой тип мобильности, точнее портативность оборудования, поддерживается протоколом DHCP, который представлен во втором разделе. В начале эры своего развития компьютеры не так часто меняли свое местоположение. Сейчас же, при наличии портативных компьютеров или ноутбуков, например, студенты приходят в университет со своими компьютерами, предполагая либо подключиться к сети напрямую, либо воспользоваться беспроводным доступом. Разумеется, сетевые администраторы не желают каждый день заново конфигурировать десятки компьютеров или выдавать списки действующих IP-адресов, серверов DNS, префиксов подсети, стандартных маршрутизаторов и т.д. В такой ситуации для автоматического конфигурирования компьютеров используется протокол динамической конфигурации хоста (Dynamic Host Configuration Protocol — DHCP).

В предлагаемой главе эпизодические сети рассматриваются наряду с сетевым уровнем. Как можно осуществить маршрутизацию в динамической сети с постоянным изменением связности? Что делать, если нет маршрутизатора или базы данных, которые знают, где сейчас находится требуемый узел? В последнем разделе рассматриваются некоторые подходы, позволяющие осуществить маршрутизацию путем расширения стандартных алгоритмов, применяемых в Internet, или за счет использования сведений о текущем состоянии физической среды.

9.1. Mobile IP

Ниже дается обзор протокола Mobile IP, расширения, требуемого Internet для поддержки мобильных хостов. Отметим, что в этой главе мы не вдавались в мелкие подробности, а представили общую картину. Отличным справочным пособием по рассматриваемой теме является работа [Perkins, 1997], в которой подробно описывается разработка протокола Mobile IP, все форматы пакетов, механизмы, подробно обсуждаются протоколы и альтернативы. Здесь потребуется некоторое знание протоколов Internet, особенно IP. Очень хороший обзор и подробное описание предлагается в работе [Stevens, 1994].

9.1.1. Цели, предположения и требования

Как было показано в главе 1, мобильная вычислительная структура очевидно является парадигмой будущего. Internet — это сеть, предназначенная для глобального обмена информацией с сотнями миллионов пользователей. Так почему бы не использовать в Internet мобильный компьютер?

Причина достаточно проста: вы не получите ни единого пакета после того, как покинете свою базовую сети (т.е. сеть, сконфигурированную под этот компьютер) и подключитесь (с использованием проводной или беспроводной связи) к другой точке. Смысл этого более чем очевиден, если рассмотреть механизм маршрутизации в Internet. Хост отсылает IP-пакет с заголовком, в котором имеется поле адреса пункта назначения и другие поля. Адрес пункта назначения определяет не только получателя пакета, но также и физическую подсеть получателя. Например, целевой адрес 129.13.42.99 показывает, что получатель должен быть подключен к физической подсети с сетевым префиксом 129.13.42 (если не используется схема CIDR). Теперь маршрутизаторы в Internet изучают целевой адрес поступившего пакета и пересылают его в соответствии с внутренними таблицами соответствия. Чтобы избежать переполнения таблиц маршрутизации, в них хранятся только префиксы. В противном случае компьютеру пришлось бы хранить адреса всех компьютеров в Internet, что, очевидно, нереально. Пока получателя можно обнаружить в пределах физической подсети, он получает пакеты; как только получатель покидает подсеть, он больше не может принимать никаких пакетов. Таким образом, хосту необходим так называемый **топологически правильный адрес**.

9.1.1.1. Быстрые "решения"

Может возникнуть мысль "решить" эту проблему путем присвоения компьютеру нового, топологически правильного IP-адреса. Теперь изменение местоположения также будет означать присвоение нового адреса; правда, возникает другая проблема — этого нового адреса никто не знает. Практически невозможно найти в Internet (мобильный) хост, который только что изменил свой адрес. В частности, службе имен доменов DNS требуется некоторое время для обновления внутренних таблицы отображения логических имен в IP-адреса. Этот подход не работает, если мобильные узлы перемещаются довольно часто. Более того, Internet и DNS разрабатывались без расчета на частые обновления. Просто представьте миллионы узлов, перемещающихся одновременно. DNS никогда не сможет непротиворечиво охватить имена и адреса, для чего он использует кэширование. Быстрое обновление было бы слишком дорогим удовольствием.

Кроме того, существует серьезная проблема, связанная с протоколами высших уровней, подобными TCP, который зависит от IP-адресации. Изменение IP-адреса при установленном соединении TCP означает разрыв этого соединения. Соединение TCP определяется кортежем (IP-адрес источника, порт источника, IP-адрес назначения, порт назначения), известным как **сокет**. Следовательно, соединение TCP не может сохраниться при изменении какого-либо адреса. Разрыв соединения TCP недопустим, использование программ, подобных telnet, станет просто невозможным. К тому же мобильный узел должен будет оповещать всех партнеров по связи о своем новом адресе.

Другой подход — это создание специфических маршрутов к мобильному узлу. Маршрутизаторы всегда выбирают наиболее удобный префикс при решении о выборе маршрута. Если маршрутизатор в данный момент имеет запись для префикса 129.13.42 и адреса 129.13.42.99, то при поступлении пакета с целевым адресом 129.13.42.99 для перенаправления он выберет порт, связанный с последним адресом. Теоретически возможно изменить все таблицы маршрутизации во всем мире для создания специфических маршрутов для мобильного узла, но этот подход невозможно распространить на все узлы Internet. Маршрутизаторы настраиваются на максимально быструю передачу, а не на быстрое обновление таблиц маршрутизации. Если первое выполняется за счет аппаратной реализации, то последнее зависит от программного обеспечения, которое не может справиться с задачей интенсивных

обновлений. К тому же маршрутизаторы — это "мозг" Internet, который удерживает вместе всю сеть. Нет такого поставщика услуг или системного администратора, который лишь для удовлетворения нужд мобильных пользователей смог бы обеспечить изменение таблиц маршрутизации, не пожертвовав взамен устойчивостью.

9.1.1.2. Требования

Поскольку очевидно, что быстрые "решения" проблемы — это не выход, нужно разрабатывать более общую архитектуру. Большое число полевых испытаний и опыт частных систем в конце концов привели к разработке протокола Mobile IP — стандарта, допускающего мобильность в Internet. Разработка стандарта сопровождается несколькими требованиями.

- **Совместимость**. База установленного оборудования Internet, т.е. компьютеров, которые используют TCP/IP и подключены к Internet, огромна. Новый стандарт не должен требовать изменения уже используемых приложений или сетевых протоколов. Пользователи попрежнему предпочитают использовать знакомые Web-броузеры и не захотят изменять что-либо лишь ради поддержки мобильности. То же самое относится и к операционной системе. Никто не захочет устанавливать другую операционную систему лишь ради самой мобильности, поэтому протокол Mobile IP должен интегрироваться в уже существующие операционные системы или, по крайней мере, быть совместимым с ними. На маршрутизаторы Internet не должно устанавливаться новое программное обеспечение. Хотя можно расширить возможности некоторых маршрутизаторов для поддержки мобильности, почти невозможно изменить все маршрутизаторы. Более того, протокол Mobile IP должен быть совместимым со всеми нижними уровнями, используемыми в стандартном, немобильном протоколе IP. Это означает, что протокол Mobile IP не должен требовать специальных сред или протоколов MAC/LLC. Итак, протокол Mobile IP должен использовать те же интерфейсы и механизмы доступа к нижним уровням, что и IP. И наконец, конечные системы с Mobile IP должны быть способны сообщаться со стационарными системами без этого протокола. Mobile IP гарантирует всем пользователям возможность доступа ко всем другим серверам и системам Internet. Следует отметить, что это также подразумевает использование одинаковых форматов адресов и механизмов маршрутизации.
- **Прозрачность**. Мобильность должна быть "незаметной" для более высоких уровней протоколов и приложений за исключением, возможно, случаев понижения пропускной способности и перебоев в обслуживании. Более высокие уровни должны продолжать работу при изменении мобильным компьютером точки своего подключения к сети. Например, для TCP это означает, что компьютер должен сохранять свой IP-адрес, как это было объяснено выше. Если повреждение связности не слишком длительное, соединение TCP "выживает" при изменении точки подключения. Проблемы, связанные с производительностью TCP, рассматриваются в главе 10. Ясно, что многие современные приложения не разрабатывались для использования в мобильной среде. Таким образом, единственными следствиями мобильности должно быть увеличение задержек передачи и понижение пропускной способности. Однако, существуют некоторые приложения, для которых было бы неплохо изначально предполагать мобильность. Примеры: маршрутизация, зависимая от стоимости, или сжатие видео. Зная, что в настоящее время можно использовать различные сети, программное обеспечение должно выбирать самую дешевую из них. Или если видеоприложению известно, что сейчас доступно лишь соединение с низкой пропускной способностью, оно должно использовать иную схему сжатия. Таким образом, необходимы дополнительные механизмы, которые будут информировать эти приложения о мобильности [Brewer, 1998].
- **Масштабируемость и эффективность**. Введение новых механизмов в Internet не должно сказываться на производительности самой сети. Расширение IP для поддержки мобильно-

сти не должно порождать большого количества новых сообщений, наводняющих всю сеть. Кроме того, следует учитывать особое обстоятельство — более низкую пропускную способность беспроводных линий связи. Во многих мобильных системах с точкой подключения будет соединена беспроводная линия связи. Следовательно, должны требоваться лишь несколько дополнительных пакетов между мобильной системой и узлом сети. Если принять во внимание количество компьютеров в Internet и темпы развития мобильных систем связи, то очевидно, что бесчисленное множество устройств в Internet окажутся мобильными компонентами. Только подумайте о легковых автомобилях, грузовиках, мобильных телефонах, каждом посадочном месте в каждом самолете мира и т.д. — во многих из них будет реализована некоторая схема IP, и многие будут перемещаться между различными сетями, а значит, им понадобится протокол Mobile IP. Таким образом, данный протокол обязательно должен быть масштабируемым на огромное количество объектов Internet по всему миру.

- **Безопасность**. Мобильность создает много проблем безопасности. Минимальным требованием является аутентификация всех сообщений, связанных с управлением протоколом Mobile IP. При пересылке пакета на мобильный хост уровень IP должен быть уверенным, что именно этот хост является действительным получателем пакета. Уровень IP может лишь гарантировать, что IP-адрес получателя правильный. Предотвратить подделку IP-адресов или защититься от других атак нельзя. В соответствии с философией Internet за это отвечают более высокие уровни.

Таким образом, задачи протокола Mobile IP можно определить как "поддержка мобильности конечных систем при поддержании масштабируемости, эффективности и совместимости со всеми существующими приложениями и Internet-протоколами".

9.1.2. Объекты и терминология

Ниже приводятся определения некоторых объектов и терминов, необходимых для понимания протокола Mobile IP, в соответствии с RFC 2002 [Perkins, 1996a]. На рис. 9.1 показан пример сценария.

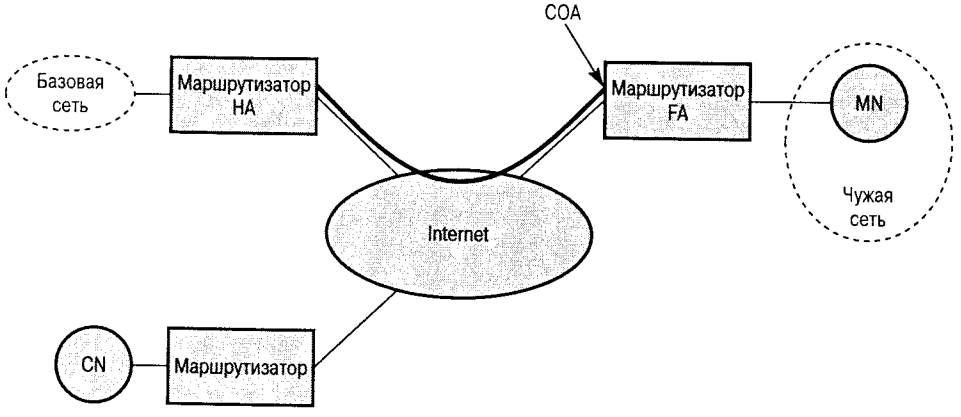

Рис. 9.1. Пример сети Mobile IP

- **Мобильный узел (mobile node — MN)**. Мобильным узлом является конечная система или маршрутизатор, который может изменять свою точку подключения к Internet с помощью протокола Mobile IP. MN сохраняет свой IP-адрес и может постоянно поддерживать связь с любой другой системой в Internet, пока предоставляется связь на канальном уровне. Мо-

бильные узлы не обязательно являются портативными устройствами, подобными лэптопам с антеннами или мобильным телефонам; мощным мобильным узлом может, например, быть маршрутизатор на борту самолета.

- **Корреспондентский узел (correspondent node — CN)**. Для связи необходим, по крайней мере, один партнер. При дальнейшем обсуждении в качестве партнера MN используется CN, который может быть стационарным или мобильным узлом.
- **Базовая сеть**. Базовая сеть — это подсеть, к которой (согласно IP-адресу) принадлежит MN. В базовой сети поддержка протокола Mobile IP не требуется.
- **Чужая сеть**. Чужой сетью для MN является подсеть, в которой он на данный момент находится и которая не является его базовой.
- **Чужой агент (foreign agent — FA)**. FA может предоставлять некоторые услуги MN при посещении чужой сети. FA может иметь адрес COA (см. ниже), действуя как конечная точка туннеля, и пересылать пакеты к MN. Кроме того, FA может быть стандартным маршрутизатором для MN. FA также могут предоставлять услуги безопасности для своих сетей. Для функционирования протокола Mobile IP наличие FA не обязательно. Как правило, FA реализуется на маршрутизаторе подсети, к которой подключен MN.
- **Адрес передачи (care-of address — COA)**. Адрес COA определяет текущее расположение MN с точки зрения IP. Все пакеты IP, адресованные MN, направляются по адресу COA, а не непосредственно по IP-адресу MN. Передача пакета MN осуществляется с использованием туннеля, как объясняется ниже. Следовательно, если быть точнее, COA отмечает конечную точку туннеля, т.е. адрес, где пакет покидает туннель. Существует два возможных расположения COA.
  - **COA чужого агента**. COA мог бы располагаться у FA, т.е. COA — это IP-адрес FA. Таким образом, FA является конечной точкой туннеля и направляет пакеты к MN. Если один FA используют много MN, то его COA используется совместно как общий COA.
  - **Сопредельный COA**. COA называется сопредельным, если MN временно получает дополнительный IP-адрес, который действует как COA. Этот адрес сейчас топологически правилен, и конечная точка туннеля находится у MN. Сопредельные адреса можно запрашивать через службу типа DHCP (см. раздел 9.2). С этим подходом связана одна проблема — если слишком много MN будут запрашивать COA, потребуется много дополнительных адресов. Это не всегда хорошо, принимая во внимание дефицит адресов IPv4.
- **Базовый агент (home agent — HA)**. HA предоставляет некоторые услуги MN и располагается в базовой сети. Туннель для пакетов, направляющихся к MN, начинается с HA. Кроме того, HA содержит реестр расположения, т.е. он знает о текущем расположении MN через COA. Имеются три альтернативных способа реализации HA.
  - HA можно реализовать на маршрутизаторе, который отвечает за базовую сеть. Очевидно, что это наилучшее расположение, так как без оптимизации протокола Mobile IP все пакеты для MN должны в любом случае идти через маршрутизатор, как упоминалось ранее.
  - Если программное обеспечение маршрутизатора изменить нельзя, точно так же можно реализовать HA на произвольном узле подсети. Недостатком такого решения является двойное прохождение пакета через маршрутизатор, если MN находится в чужой сети. Пакет для MN входит через маршрутизатор; HA отсылает его через туннель, который снова пересекает этот маршрутизатор.
  - И наконец, базовая сеть не нужна вовсе. HA можно снова расположить на "маршрутизаторе", но теперь он будет действовать как система управления для MN, который принадлежит виртуальной базовой сети. При таком решении все MN находятся в чужой сети.

Пример сети на рис. 9.1 иллюстрирует следующую ситуацию: CN, базовая и чужая сети подключаются через маршрутизатор к Internet. HA реализуется на маршрутизаторе, соединяющем базовую сеть с Internet, а FA находится на маршрутизаторе чужой сети. В данный момент MN находится в чужой сети. Туннель для пакетов, направляющихся к MN, начинается с HA и заканчивается на FA, поскольку в этом примере COA находится у FA.

9.1.3. Доставка пакетов IP

На рис. 9.2 показана доставка пакетов к MN и обратно в сети, показанной на рис. 9.1. Корреспондентский узел CN хочет отослать IP-пакет MN. Одно из требований к протоколу Mobile IP заключалось в скрытии мобильности MN. Следовательно, CN не нужно знать текущее расположение MN, и пакет отсылается обычным образом на IP-адрес MN (этап 1). Это означает, что CN отсылает IP-пакет по адресу пункта назначения MN и с адресом источника CN. Internet, не имея информации о текущем расположении MN, направляет пакет к маршрутизатору, отвечающему за базовую сеть MN. Это осуществляется с помощью стандартного механизма маршрутизации в Internet.

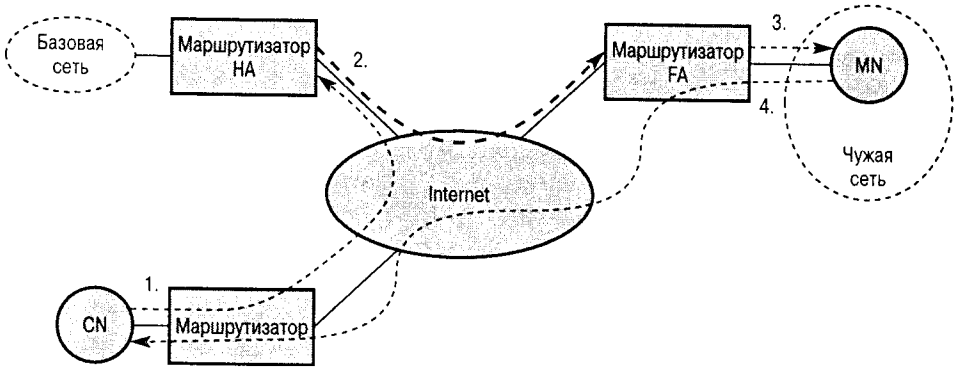

Рис. 9.2. Отправка пакетов к мобильному узлу и обратно

Теперь пакет перехватывает HA, который знает, что MN в данный момент не находится в базовой сети. Значит пакет не переправляется в подсеть, как обычно, а инкапсулируется и пересылается по туннелю на COA. Это происходит путем добавления нового заголовка перед старым IP-заголовком, который показывает COA как новый пункт назначения и HA как источник инкапсулированного пакета (этап 2). (Передача по туннелю и инкапсуляция более подробно описываются разделе 9.1.6.) Чужой агент теперь декапсулирует пакет, т.е. убирает дополнительный заголовок и переправляет исходный пакет (где как источник указан CN) по адресу пункта назначения — MN (этап 3). И снова узлу MN мобильность не видна. Он принимает пакет с теми же адресами отправителя и получателя, как если бы это происходило в базовой сети.

На первый взгляд отправка пакета от MN к CN выглядит просто, но все же имеются и проблемы, которые обсуждаются в разделе 9.1.8. MN отсылает пакет как обычно, с собственным фиксированным IP-адресом, указанным как адрес источника, и адресом CN, указанным как адрес получателя (этап 4). Маршрутизатор с FA, действуя как обычный маршрутизатор, переправляет пакет таким же образом, как для любого другого узла в чужой сети. Если CN является стационарным узлом, в остальной части Internet изменений не наблюдается. Если же CN является мобильным узлом, который находится в чужой сети, то в обратном направлении включается механизм, описанный на этапах 1–3.

В следующих разделах представляются дополнительные механизмы, которые необходимы для работы протокола Mobile IP, некоторые типы усовершенствований протокола и определенные проблемы, связанные с эффективностью и безопасностью.

9.1.4. Объявление и открытие агента

Следующей проблемой для MN после перемещения является обнаружение чужого агента. Как MN может осуществлять обнаружение, если он при этом движется? Для этого чужие и базовые агенты периодически объявляют о своем присутствии посредством специальных информационных сообщений. Такие информационные сообщения можно рассматривать как широковещание в подсети. Для них используется протокол ICMP в соответствии с RFC 1256 [Deering, 1991] с некоторыми дополнениями для поддержки мобильности. Маршрутизаторы в стационарной сети также используют этот стандарт для периодического объявления о своих службах маршрутизации в присоединенных каналах.

Пакет объявления агента, соответствующий RFC 1256, с поддержкой мобильности показан на рис. 9.3. Верхняя часть представляет собой пакет ICMP, а нижняя часть поддерживает мобильность. На этом рисунке не показаны поля, необходимые на нижних уровнях при объявлении агента. Ясно, что мобильные узлы должны иметь подходящие адреса канального уровня. Более того, поле TTL IP-пакета устанавливается на 1 для всех объявлений, чтобы избежать их передачи. IP-адрес пункта назначения, соответствующий стандартному объявлению маршрутизатора, может устанавливаться или как 224.0.0.1, что является групповым адресом для всех систем канала [Deering, 1989], или как широковещательный адрес 255.255.255.255.

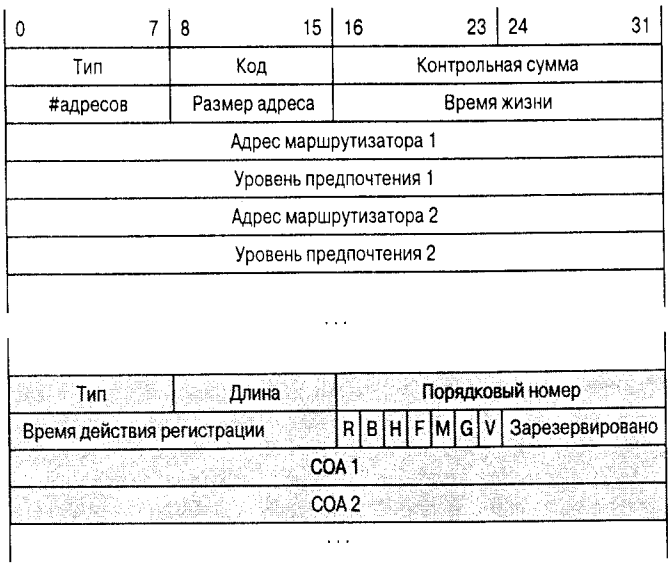

Рис. 9.3. Пакет объявления агента (RFC 1256 + поддержка мобильности)

Поля в части ICMP задаются следующим образом. **Тип** устанавливается на 9, **код** может быть равен 0, если агент направляет информационный поток от немобильных узлов, или 16, если нет другого информационного потока, кроме мобильного. Для переправления пакетов от мобильных узлов требуются, по крайней мере, чужие агенты. Количество адресов, объявленных в этом пакете, находится в поле **#адресов**, за этим полем следуют собственно **адреса**. **Время жизни** обозначает промежуток времени, в течение которого это объявление действительно. Уровни **предпочтения** для каждого адреса помогают узлу выбрать маршрутизатор, который наиболее подходит новому узлу.

Отличие от стандартного объявления ICMP наблюдается в том, *что* находится после адресов маршрутизатора. Это расширение, обеспечивающее поддержку мобильности, имеет сле-

дующие поля: **тип** устанавливается на 16, **длина** зависит от числа адресов COA, указанных в сообщении, причем эта величина равна 6 + 4 * (*количество адресов*). Кроме того, агент показывает общее количество разосланных объявлений непосредственно после инициализации в поле **порядковый номер**. **Время действия регистрации** определяет максимальное время в секундах, в течение которого узел может послать запрос на регистрацию, как объясняется в разделе 9.1.5. Дальнейшую характеристику агента дают такие биты. Бит **R** (регистрация) показывает, что требуется регистрация у этого агента, а не использование сопредельного COA в MN. Если агент сейчас слишком занят, чтобы осуществлять новую регистрацию, он может установить бит **B**. Следующие два бита означают, что агент может предоставлять услуги как базовый агент (**H**) и чужой (**F**) в канале, из которого получено объявление. Биты **M** и **G** задают метод инкапсуляции, используемый в туннеле, о чем будет рассказано в разделе 9.1.6. Хотя обязательной является инкапсуляция IP-IP, **M** может задавать минимальную инкапсуляцию, а **G** — общую маршрутную инкапсуляцию. Бит **V** может задать использование сжатия заголовка, соответствующее RFC 1144 [Jacobson, 1990]. Следующие поля содержат объявленные **COA**. Чужой агент, который устанавливает бит **F**, должен объявить, по крайней мере, один адрес COA. Более подробное описание и некоторые дополнения можно найти в работе [Perkins, 1997].

Мобильный узел в подсети может принимать объявления как от базового агента, так и от чужого. Это одна возможность для MN определить их расположение. Если объявлений агентов нет, и MN не принимает COA каким-либо иным способом, например, через DHCP, (см. раздел 9.2), мобильный узел должен разослать агентам запросы. Эти запросы определены в RFC 1256 для запросов к маршрутизаторам. Следует также обращать внимание на возможность переполнения сети этими запросами, ведь в принципе MN может искать FA, без конца рассылая запросы. Обнаружение нового агента может произойти в любой момент, а не только тогда, когда MN не подключен к нему. Рассмотрим, например, случай, когда MN пытается подыскать лучшее соединение, в то время как передача происходит по старому пути. Это происходит, например, при перемещении через несколько ячеек разных беспроводных сетей.

После реализации этапа объявления запросов и обнаружения агента MN может получить теперь COA. Более того, MN знает свое расположение (в базовой или чужой сети) и возможности агента (если нужно). Следующий этап заключается в регистрации MN и HA, если MN находится в чужой сети, что описывается в разделе 9.1.5.

Использование стандарта, такого как RFC 1256, для целей, которые отличаются от первоначального объявления маршрутизатора, может создать некоторые проблемы. Достаточно очевидным является существование минимального интервала (три секунды) между двумя объявлениями. Это абсолютно обосновано в проводных сетях, поскольку их топология меняется довольно медленно (некоторое время занимает отключение бездействующих маршрутизаторов и т.п.). Однако, в высоко динамичных беспроводных сетях с перемещающимися MN и, вероятно, с приложениями, требующими непрерывных потоков пакетов, три секунды — это слишком много. MN пришлось бы постоянно ждать, по меньшей мере, три секунды, перед тем как выяснится, что агент недоступен. Следовательно, для переключения на другой агент MN придется ждать еще дольше. Рассылка запросов — это не лучший выход, поскольку существует возможность переполнения подсети.

9.1.5. Регистрация

После получения COA MN должен зарегистрироваться у HA. Целью регистрации является сообщение HA своего текущего расположения для правильной пересылки пакетов. Регистрацию можно осуществить двумя разными способами, в зависимости от расположения COA.

- Если COA находится у FA, регистрация происходит, как показано на рис. 9.4. MN отсылает запрос на регистрацию, содержащий COA (см. рис. 9.6), FA, который переправляет запрос HA.

Тогда HA настраивает **привязку мобильности**, содержащую базовый IP-адрес мобильного узла и его текущий адрес COA. Кроме того, привязка мобильности включает в себя время действия регистрации, которое определяется в ходе процесса регистрации. Регистрация аннулируется и исключается автоматически по окончании ее срока действия, таким образом, перед аннуляцией MN должен перерегистрироваться. Это необходимо для того, чтобы избежать появления неиспользуемых привязок мобильности. После установления привязки мобильности HA отсылает ответное сообщение к FA, который, в свою очередь, отсылает его MN.

- Если COA является сопредельным адресом, регистрация выглядит проще, как можно видеть из рис. 9.5. MN отсылает запрос непосредственно HA, и наоборот. Между прочим, так же выглядит процедура регистрации MN в базовой сети. Здесь мобильные узлы регистрируются непосредственно у HA.

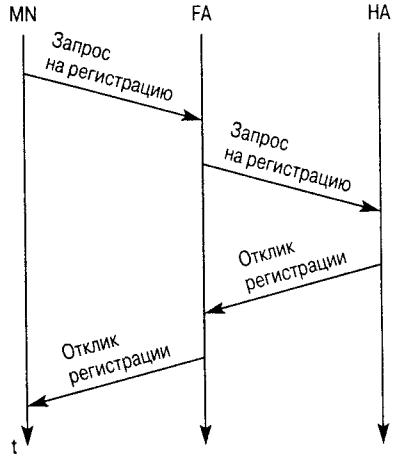

Рис. 9.4. Регистрация мобильного узла через чужой агент

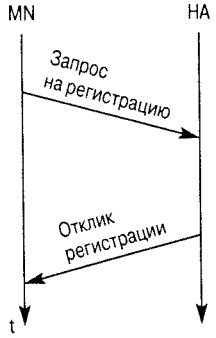

Рис. 9.5. Регистрация мобильного узла через свой агент

Для запросов на регистрацию применяются пакеты UDP. IP-адрес источника пакетов устанавливается равным адресу интерфейса MN, а IP-адрес назначения — адресу FA или HA (в зависимости от расположения COA). Порт назначения UDP устанавливается равным 434. Протокол UDP применяется из-за малых служебных издержек и более высокой производительности по сравнению с возможностями TCP в беспроводных средах (см. главу 10). Поля, связанные с регистрацией Mobile IP, следуют как данные UDP (см. рис. 9.6). Поля определяются следующим образом.

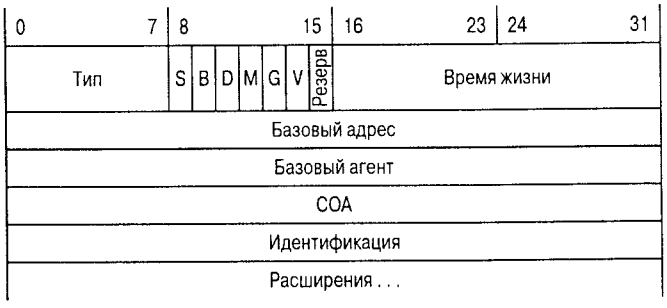

Рис. 9.6. Запрос на регистрацию

1 в первом поле (**тип**) означает запрос на регистрацию. Бит **S** указывает, должен ли HA сохранить прежние привязки мобильности. Это нужно для одновременного связывания. Следующие биты определяют требуемое поведение при пересылке пакетов. Установка бита **B**, как правило, означает, что MN также хочет принимать широковещательные пакеты, принятые HA в своей сети. Более подробное описание механизма фильтрации мобильным узлом широковещательных сообщений можно найти в работе [Perkins, 1997]. Если MN использует сопредельный адрес COA, то он также занимается декапсуляцией пакетов в конечной точке туннеля, что указывается битом **D**. Как определялось для объявления агентов, биты **M**, **G** и **V** задают минимальную инкапсуляцию, специальную маршрутную инкапсуляцию и сжатие заголовка соответственно.

Время жизни задает время в секундах, в течение которого регистрация будет действительна. **Базовый адрес** соответствует фиксированному IP-адресу MN, **базовый агент** — это IP-адрес HA, а **COA** представляет собой конечную точку туннелирования. 64-битное поле **идентификация** создается MN для идентификации запроса и согласования его с регистрационным ответом. Кроме того, это поле применяется для защиты от атак регистрации.

9.1.6. Туннелирование и инкапсуляция

Ниже описываются механизмы, используемые для переправки пакетов между HA и COA, как показано на рис. 9.2, этап 2. **Туннель** устанавливает виртуальный канал для пакетов данных между входом туннеля и конечной точкой туннеля. Пакеты поступают в туннель, пересылаются по нему и покидают его неизмененными. Туннелирование, т.е. пересылка пакетов по туннелю, осуществляется с использованием инкапсуляции.

Инкапсуляция — это механизм приема пакета, состоящего из заголовка и информации, и помещение его в информационную часть нового пакета. Обратная операция, извлечение пакета из информационной части другого пакета, называется **декапсуляцией**. Чаще всего процедуры инкапсуляции и декапсуляции выполняются при передаче пакета от протоколов более высокого уровня к протоколам более низкого уровня и обратно. Здесь же эти функции используются на одном и том же уровне.

Описанный механизм показан на рис. 9.7, где точно обозначено, что HA делает на входе туннеля. HA берет исходный пакет с целевым адресом MN и помещает его в область данных нового пакета, для которого создается новый заголовок IP, после чего полученный пакет отсылается на адрес COA. По очевидным причинам новый заголовок еще называют **внешним**. Кроме него имеется **внутренний заголовок**, который может быть идентичен исходному, как в случае инкапсуляции IP-IP, или же внутренний заголовок может вычисляться в ходе инкапсуляции.

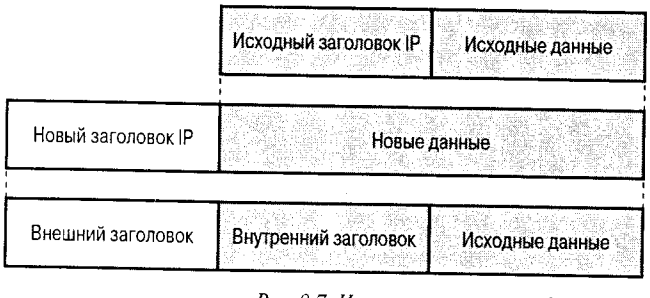

Рис. 9.7. Инкапсуляция пакета IP

9.1.6.1. Инкапсуляция IP-IP

Существуют разные способы осуществления процедуры инкапсуляции, необходимой при передаче пакетов между HA и COA. Для протокола Mobile IP в качестве обязательной определена **инкапсуляция IP-IP** [Perkins, 1996б). На рис. 9.8 показан пакет в туннеле. Поля соответствуют спецификации

IP-протокола в RFC 791 [Postel, 1981]. Поля внешнего заголовка устанавливаются следующим образом. Поле версии (**версия**) устанавливается равным 4, если используется версия IPv4, длина заголовка Internet (**IHL**) задает размер внешнего заголовка в 32-битовых словах. Поле **TOS** копируется из внутреннего заголовка, поле **длина** охватывает полный инкапсулированный пакет. Поля до TTL в протоколе Mobile IP имеют специальное назначение и задаются в соответствии с RFC 791. Значение поля **TTL** должно быть достаточно большим, чтобы пакет мог достичь конечной точки туннеля. Следующее поле, которое здесь обозначено как **IP-IP**, означает тип протокола, применяемый для полезной нагрузки IP. В нашем случае значение поля равно 4, поскольку оба заголовка соответствуют IPv4. **Контрольная сумма** пакета IP рассчитывается традиционным способом. Следующее поле соответствует входу туннеля и представляет собой адрес источника (**IP-адрес HA**), а выход туннеля представляет собой адреса пункта назначения (**COA**).

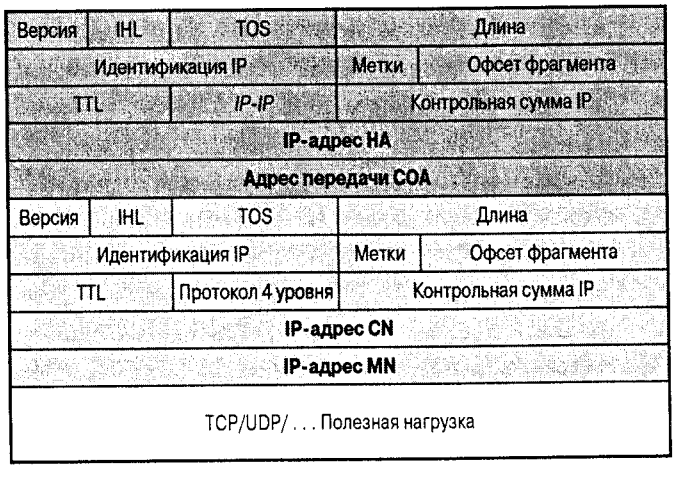

Рис. 9.8. Инкапсуляция IP-IP

Если после внешнего заголовка нет никаких опций, то внутренний заголовок состоит из таких же полей, как было описано выше. Этот заголовок в процессе инкапсуляции остается практически неизменным, указывая, таким образом, исходного отправителя (**CN**) и получателя пакета (**MN**). Изменение касается лишь поля **TTL**, значение которого уменьшается на 1. Это означает, что весь туннель с точки зрения исходного пакета рассматривается как единый маршрутный участок. Это очень важное свойство туннелирования, которое позволяет MN вести себя так же, как и в базовой сети. Не имеет значения, сколько маршрутных участков в действительности проходит пакет в туннеле, логически это единый участок до MN. Наконец, после двух заголовков следует полезная нагрузка.

9.1.6.2. Минимальная инкапсуляция

Как видно на примере инкапсуляции IP-IP, некоторые поля являются избыточными. Например, поле TOS просто копируется, часто нет нужды во фрагментации и т.п. Таким образом, **минимальная инкапсуляция**, показанная на рис. 9.9, — это частный случай инкапсуляции, применяемый в протоколе Mobile IP. Как и ранее, задаются начальная и конечная точки туннеля, но в этом случае поле **тип заголовка** содержит значение 55 для протокола минимальной инкапсуляции. Внутренний заголовок в схеме минимальной инкапсуляции отличается. Остается тип применяемого протокола и адрес MN. Если есть бит **S**, то включается исходный адрес отправителя CN. Во внутреннем заголовке отсутствуют поля для сдвига фрагментации, и, следовательно, минимальная инкапсуляция не работает при уже фрагментированных пакетах.

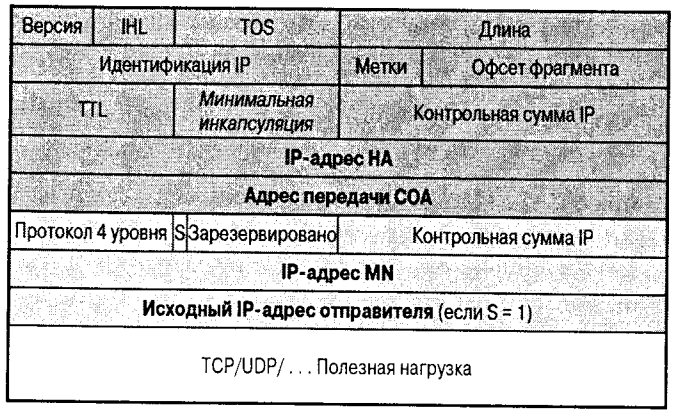

Рис. 9.9. Минимальная инкапсуляция

9.1.6.3. Общая маршрутная инкапсуляция

Если обычная инкапсуляция и минимальная инкапсуляция работают только под протоколом IP, описываемая ниже схема инкапсуляции наряду с IP поддерживает также и другие протоколы сетевого уровня. **Общая маршрутная инкапсуляция** (generic routing encapsulation — GRE) позволяет инкапсуляцию пакетов одного семейства протоколов в полезную нагрузку пакетов другого семейства [Hanks, 1994]. Эта процедура показана на рис. 9.10. К пакету одного семейства протоколов с исходными заголовком и данными прибавляется новый заголовок GRE. Все вместе является новыми данными и новым заголовком пакета. В завершение, в начало полученного пакета добавляется заголовок второго семейства протоколов.

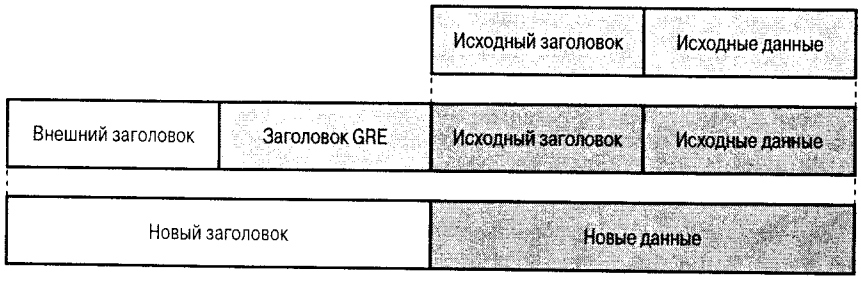

Рис. 9.10. Общая маршрутная инкапсуляция [Hanks, 1994]

На рис. 9.11 показаны поля пакета из туннеля между базовым агентом и COA, использующим GRE в качестве схемы инкапсуляции (RFC 1701). Внешний заголовок — это обычный IP-заголовок с HA в качестве адреса источника и COA в качестве адреса пункта назначения. В этом внешнем IP-заголовке тип протокола равен 47 (GRE). Прочие поля внешнего пакета, такие как **TTL** и **TOS**, можно скопировать с исходного IP-заголовка. Однако, после декапсуляции поле **TTL** нужно уменьшить на 1, чтобы избежать бесконечной пересылки.

Заголовок GRE начинается с нескольких меток, указывающих на наличие или отсутствие полей. Минимальный заголовок GRE занимает всего 4 байта, но при этом он достаточно гибок для включения нескольких механизмов. Бит **C** указывает, используется ли поле контрольной суммы и содержит нужную информацию. Если значение C установлено, то поле **контрольная сумма** содержит правильную контрольную сумму IP-заголовка GRE и полезную нагрузку. Бит **R** показывает на наличие полей сдвига и маршрута и содержит нужную информацию. Поле **сдвиг** означает сдвиг в байтах для первой позиции **маршрутизации** источника. Поле маршрути-

зации (если есть) имеет переменный размер и содержит поля, требуемые для маршрутизации источника. Если установлен бит **C**, также есть поле сдвига, и наоборот, если установлен бит **R**, должно быть поле контрольной суммы. Единственное обоснование такого подхода — размер этих полей должен быть равен 4 байтам. Поле контрольной суммы действительно, только если установлено C, а поле сдвига действительно, если установлено R.

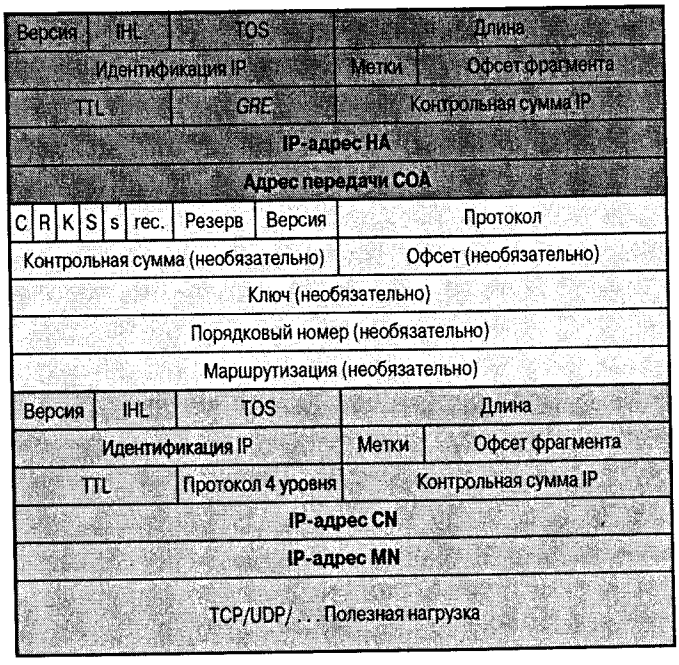

Рис. 9.11. Поля протокола при общей маршрутной инкапсуляции (RFC 1701)

В GRE имеется также поле **ключ**, которое можно использовать для аутентификации. При наличии этого поля устанавливается бит **K**. Однако, алгоритмы аутентификации для GRE не специфицированы. Бит порядкового номера **S** определяет наличие поля **порядкового номера**; если бит **S** установлен, используется маршрутизация только через заданные узлы. Порядковые номера могут использоваться при декапсуляции для восстановления порядка пакетов. Это может оказаться важным, если инкапсулируется протокол, гарантирующий упорядоченную передачу, а затем происходит передача с использованием протокола, не обладающего такой особенностью, например, IP. В таком случае для сохранения характеристик протокола при декапсуляции нужно восстановить прежний порядок.

Поле **контроля рекурсии** (rec.) — это важное поле, которое дополнительно отличает GRE от инкапсуляции IP-IP и минимальной инкапсуляции. Данное поле представляет собой счетчик, который показывает число выполненных рекурсивных операций инкапсуляции. После прибытия пакета производится проверка и определяется, равно ли это поле нулю или нет. Если поле не равно нулю, делается дополнительная инкапсуляция — пакет инкапсулируется, и значение поля уменьшается на единицу. В противном случае пакет с высокой вероятностью будет отброшен. Этот механизм нужен для предотвращения бесконечной рекурсивной инкапсуляции, которая может произойти при некорректном определении туннелей (например, если несколько туннелей образуют цикл). Значение данного поля по умолчанию должно быть равно 0, позволяя таким образом лишь один уровень инкапсуляции.

Последующие **зарезервированные** поля должны быть равны нулю и игнорироваться при приеме. Поле **версия** содержит 0 для версии GRE. Далее 2 байта поля **протокол** представляют протокол пакета, за которым следует заголовок GRE. Определены несколько значений, например, 0x6558 для прозрачной передачи Ethernet с использованием туннелей GRE. Для туннеля Mobile IP поле протокола содержит 0x0800 (= IP).

Стандартный заголовок исходного пакета включает адрес источника, принадлежащий корреспондентскому узлу, и адрес пункта назначения, принадлежащий мобильному узлу.

9.1.7. Оптимизация

Разберем в качестве примера следующий сценарий. Японец и немец встречаются на конференции, организованной на Гавайях. Оба хотят использовать свои портативные компьютеры для обмена информацией и оба используют протокол Mobile IP для поддержки мобильности. Обратимся повторно к рис. 9.2 и представим маршрут пакетов между двумя компьютерами.

Когда японец отсылает пакет немцу, его компьютер посылает данные по базовому адресу (HA) немца, т.е. из Гавайев в Германию. Затем HA в Германии инкапсулирует пакеты и туннелирует их по адресу COA компьютера немца (Гавайи). Это означает, что хотя компьютеры могут находиться всего лишь в нескольких метрах друг от друга, пакеты обходят всю планету! Такое неэффективное поведение неоптимизированного протокола Mobile IP называется **треугольной маршрутизацией**. Треугольник образуется тремя сегментами: от CN к HA, от HA к COA/MN и от MN обратно к CN.

В базовом протоколе Mobile IP все пакеты, направляющиеся к MN, должны проходить через HA. Это может приводить к ненужному переполнению сети между CN и HA, а также и между HA и COA, в зависимости от текущего расположения MN. Более того, как показывает пример, может значительно возрастать задержка. Особо неудачная ситуация складывается, если MA и HA разделены, например, трансатлантическими линиями связи.

Один из способов оптимизации состоит в информировании CN о текущем расположении MN. Следовательно, CN может запоминать расположение с помощью кэширования в **кэше привязки**, который является частью локальной таблицы маршрутизации для CN. Подходящим информатором для CN является HA. Оптимизированному протоколу Mobile IP требуются четыре дополнительных сообщения, которые выглядят следующим образом.

- **Запрос привязки**. Любой узел, желающий знать текущее расположение MN, может направить запрос привязки к HA. HA может проверить, позволяет ли MN разглашать свое текущее расположение. Если HA разрешено сообщать сведения о расположении, он отсылает обратно обновление привязки.
- **Обновление привязки**. Это сообщение несет в себе сведения о текущем расположении MN. В сообщение включается стационарный IP-адрес MN и адрес COA. Обновление привязки может запрашивать подтверждения приема.
- **Подтверждение привязки**. Если требуется, узел возвращает подтверждение получения сообщения об обновлении привязки.
- **Предупреждение привязки**. В конце концов, если узел декапсулировал пакет для MN, но он не является текущим FA для этого MN, узел отсылает предупреждение привязки к HA этого MN. Предупреждение содержит IP-адрес MN и адрес узла, который должен попытаться переслать пакет требуемому MN. Теперь HA должен отослать обновление привязки узлу, который, очевидно, имеет неверный адрес COA для MN.

На рис. 9.12 названные четыре типа сообщений показаны в процессе обмена сообщениями между MN и FA. Прежде всего CN может запросить HA о его текущем расположении. Если MN

это позволяет, HA возвращает соответствующий MN адрес COA, используя сообщение обновления. CN подтверждает это сообщение обновления и сохраняет привязку мобильности. Теперь CN может отсылать свои данные непосредственно к текущему чужому агенту FA_{old}. FA_{old} переправляет пакеты к MN. Этот сценарий показывает адрес COA, расположенный у FA. Инкапсуляция данных для туннелирования по адресу COA теперь осуществляет CN, а не HA.

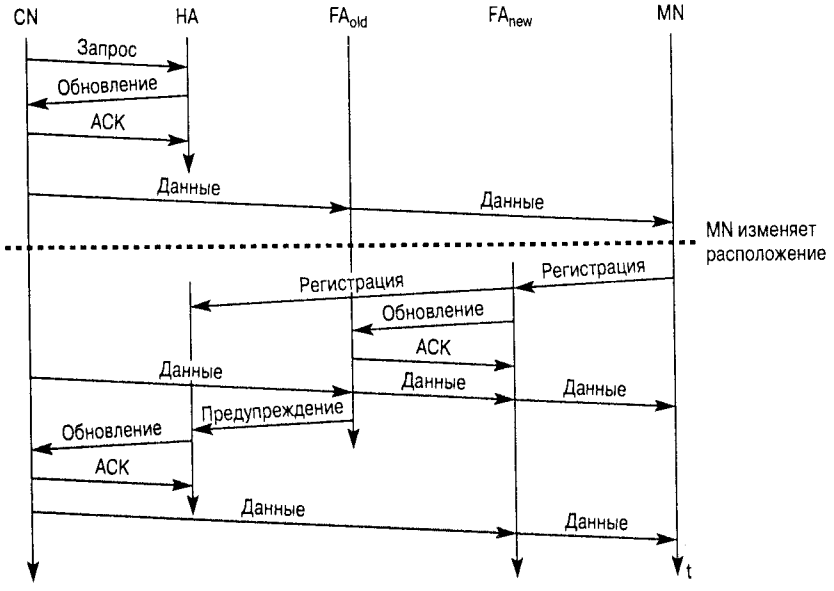

Рис. 9.12. Замена чужого агента в оптимизированном протоколе Mobile IP

Теперь MN может поменять свое расположение и зарегистрироваться у нового чужого агента, FA_{new}. Эта регистрация также переправляется HA для обновления его базы данных расположения. Кроме того, FA_{new} информирует FA_{old} о новой регистрации MN. С этой целью сообщение о регистрации MN содержит адрес FA_{old}. Ответ на регистрацию в этом сценарии не показан. Без информации, которую дает новый FA, старый FA не может ничего узнать о новом расположении MN. В таком случае CN несомненно также ничего не знает о новом расположении. Следовательно, он будет продолжать туннелировать свои пакеты, предназначенные для MN, старому FA, FA_{old}. В этом случае данный FA получает пакеты, направляемые к MN, но он знает о том, что не является текущим FA для MN. FA_{old} может в сложившейся ситуации переправить эти пакеты на новый адрес COA, соответствующий MN и расположенный, в данном случае, у FA_{new}. Такая переправка пакетов является иным способом оптимизации базового протокола Mobile IP, который обеспечивает **плавное переключение**. Без такой оптимизации все пакеты, находящиеся в пути при переключении MN с одного FA на другой, будут утеряны. В частности, если в качестве протокола высокого уровня используется TCP, это может вызвать резкое падение производительности (см. главу 10).

Чтобы сообщить CN об устаревании кэша привязки, FA_{old} отсылает предупреждение HA. Затем HA посылает CN обновление, чтобы информировать его о своем новом расположении. CN подтверждает получение обновления. После этого CN может отсылать пакеты непосредственно FA_{new}, избегая таким образом треугольной маршрутизации.

К несчастью, такая оптимизация протокола Mobile IP с целью устранения треугольной маршрутизации создает некоторые угрозы безопасности (например, атака туннеля), что рассмотрено в работе [Montenegro, 1998].

9.1.8. Обратное туннелирование

На первый взгляд обратный канал от MN к CN, который показан на рис. 9.2, выглядит проще. MN может направлять свои пакеты непосредственно CN, как и в случае традиционного IP. Адрес пункта назначения пакетов является адресом CN. Но с таким простым решением связаны некоторые сложности.

- **Брандмауэры**. Почти все компании и многие другие учреждения защищают свои внутренние сети, имеющие выход в Internet, с помощью брандмауэров. Кроме множества прочих задач брандмауэры могут осуществлять фильтрацию некоторых опасных, с точки зрения администратора, адресов. Чаще всего брандмауэры пропускают пакеты только с топологически правильными адресами. Однако, если MN отсылает пакет со своим стационарным IP-адресом в качестве исходного, то для чужой сети он не будет топологически правильным. Кроме того, брандмауэры часто отфильтровывают пакеты, пришедшие извне и содержащие адреса источников, принадлежащих внутренней сети. Это делается, чтобы избежать использования внутренних адресов другими компьютерами и их возможной маскировки под компьютеры внутренней сети. Впрочем, из этого также следует, что MN не сможет отослать пакет компьютеру, постоянно находящемуся в своей сети. В целом это означает, что из-за соображений безопасности для пересылаемого IP-пакета важен не только адрес пункта назначения, но и адрес источника.
- **Многоадресная передача**. Обратные туннели нужны MN для участия в групповой передаче. Если узел базовой сети состоит в группе, то MN в чужой сети не сможет передавать многоадресные пакеты, поскольку они будут исходить из базовой сети без обратного туннеля.
- **TTL**. Пусть MN отсылает пакеты с определенным временем TTL, оставаясь в своей сети. Время TTL может быть достаточно низким, чтобы пакеты не могли передаваться далее определенной области. Если MN теперь перейдет в чужую сеть, такое TTL может оказаться слишком низким, и пакеты уже не смогут достичь нужного узла. Протокол Mobile IP уже не будет незаметным, если пользователю нужно корректировать TTL во время движения. Следовательно, туннель должен представлять только один маршрутный участок вне зависимости от того, сколько участков в действительности находятся между базовой и чужой сетью.

Указанные соображения привели к созданию RFC 2344 [Montenegro, 1998], где определено обратное туннелирование как расширение протокола Mobile IP. Это расширение совместимо с протоколом Mobile IP и определяет топологически правильное обратное туннелирование, которое необходимо для решения рассмотренных выше проблем. Очевидно, что обратное туннелирование теперь поднимает проблему треугольной маршрутизации уже в обратном направлении. Все пакеты от MN к CN идут через HA. RFC 2344 не предлагает решения этой проблемы, поскольку не ясно, как CN будет декапсулировать пакеты. Следует помнить, что протокол Mobile IP должен работать со всеми традиционными немобильными IP-узлами. Следовательно, нельзя предполагать, что CN может стать конечной точкой туннеля.

Обратное туннелирование дополнительно поднимает несколько вопросов безопасности, на которые еще не найдено решений. Например, туннели, начинающиеся в частной сети компании и простирающиеся в Internet, могут быть взломаны и использованы для пересылки пакетов через брандмауэр. Таким образом, не ясно, как компании смогут провести туннели сквозь брандмауэр без предварительной проверки пакетов. Наиболее вероятной представляется возможность создания компанией специальной виртуальной сети для посещения мобильных узлов вне брандмауэра с полным подсоединением к Internet. Это позволит гостям пользоваться их мобильным оборудованием, в то же время придерживаясь современных стандартов безопасности.

9.1.9. IPv6

Если протокол Mobile IP изначально разрабатывался для IP четвертой версии, IP версии 6 [Deering, 1995] значительно облегчает жизнь. Некоторые механизмы, которые приходилось специально определять для поддержки мобильности, в IPv6 вошли как составляющие [Perkins, 1996г], [Johnson, 1998]. Одним из таких моментов является безопасность, которая теперь есть требуемым свойством всех узлов IPv6. Специальных механизмов для безопасной регистрации в протоколе Mobile IP дополнительно не требуется. Каждый узел IPv6 сам управляет конфигурированием адреса; таким образом, механизм запроса COA уже встроен в IPv6. Как обязательный механизм в спецификацию также включено обнаружение соседа для каждого узла, а следовательно, для предоставления услуг не потребуются специальные чужие агенты. Совмещение функций автоконфигурации и обнаружения соседей означает, что каждый мобильный узел способен создать или рассчитать топологически правильный адрес для текущей точки подключения.

Более того, каждый узел IPv6 может отослать обновление привязки другому узлу, а значит, MN может отсылать свой текущий адрес COA непосредственно к CN или HA. Эти механизмы являются неотъемлемой частью IPv6. Кроме этого, в IPv6 возможно мягкое переключение. MN сообщает свой новый адрес COA старому маршрутизатору, обслуживающему MN на старом COA, и старый маршрутизатор инкапсулирует все поступающие для MN пакеты и переправляет их по новому адресу COA.

В сумме протокол Mobile IP в сетях IPv6 требует меньше дополнительных механизмов от CN, MN и HA. FA в этом протоколе больше не нужны. От CN требуется лишь уметь проводить обновление привязки, т.е. создавать или обновлять данные в кэше маршрутизации. MN должен сам уметь декапсулировать пакеты, определять время получения нового адреса COA и сроки отправки обновлений привязки HA и CN. А HA должен уметь инкапсулировать пакеты.

9.2. Протокол динамической конфигурации хоста

Протокол динамической конфигурации хоста (Dynamic Host Configuration Protocol — DHCP) (RFC 2131, [Drohms, 1997]) в основном применяется для упрощения установки и эксплуатации сетевых компьютеров. Если к сети подключается новый компьютер, DHCP может обеспечить его всей необходимой информацией для полной интеграции системы в сеть, например, адресами серверов DNS и стандартных маршрутизаторов, маской подсети, именем домена и IP-адресом. В частности, последняя возможность — предоставление IP-адреса, делает DHCP очень интересным инструментом для протокола Mobile IP в качестве источника адресов передачи. Хотя базовые механизмы DHCP достаточно просты, в них имеется много дополнительных возможностей настройки, которые определены в RFC 2132 [Alexander, 1997].

DHCP основан на модели клиент/сервер, как показано на рис. 9.13. Клиенты DHCP отсылают серверу (в данном случае DHCPDISCOVER) запрос, на который сервер отвечает. Клиент отсылает запросы, используя широковещание MAC. Для переправки запросов на сервер DHCP может потребоваться ретранслятор DHCP.

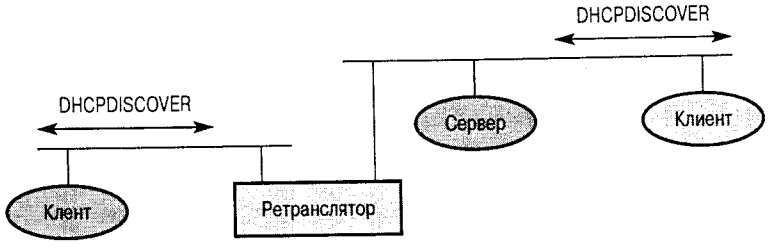

Рис. 9.13. Базовая конфигурация DHCP

На рис. 9.14 показана типичная процедура установки клиента DHCP. На этом рисунке изображен один клиент и два сервера. Как упоминалось выше, клиент использует широковещательную передачу и передает сообщение DHCPDISCOVER по подсети. Иногда для передачи такого широковещательного сообщения может применяться ретранслятор. В данном примере показано, что два сервера принимают это широковещательное сообщение и определяют конфигурацию, которую смогут предоставить клиенту. Далее должна произойти проверка доступных IP-адресов и выбор одного из них для клиента. Серверы отвечают на запрос клиента через сообщение DHCPOFFER и предоставляют список настроечных параметров. После этого клиент может выбрать себе один из предоставленных вариантов настройки. После принятия варианта настройки клиент связывается с серверами и уведомляет об отклонении прочих настроек через сообщение DHCPREQUEST. Если сервер принимает сообщение DHCPREQUEST, то он освобождает зарезервированные настройки для использования их другими возможными клиентами. Сервер подтверждает принятие конфигурации клиентом через сообщение DHCPACK. На этом фаза инициализации завершается.

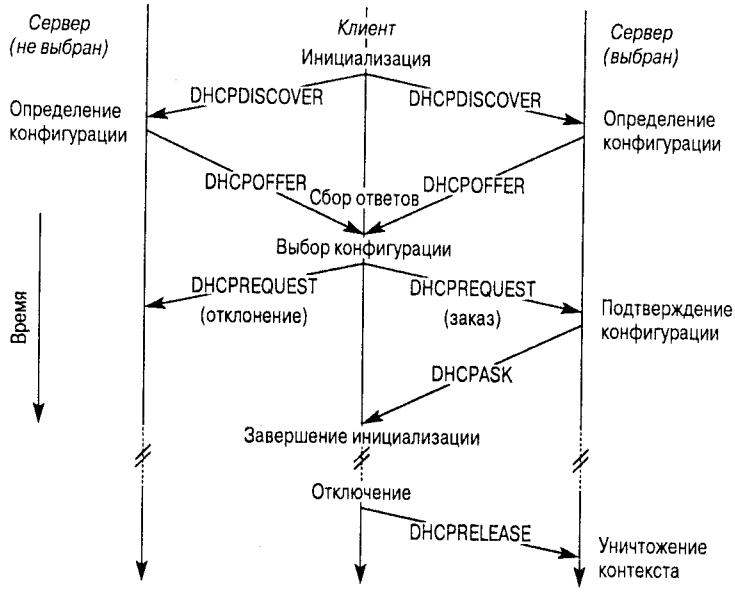

Рис. 9.14. Инициализация клиента посредством DHCP

Если клиент покидает подсеть, он должен освободить конфигурацию, полученную от сервера, передав сообщение DHCPRELEASE. Теперь сервер может освободить среду, сохраняемую для клиента, и снова предоставить конфигурацию. Клиент получает конфигурацию от сервера лишь на некоторое время, именуемое временем аренды. Следовательно, время от времени клиент должен подтверждать конфигурацию. В противном случае сервер будет считать конфигурацию освободившейся. Такое ограничение времени действия конфигурации облегчает жизнь, если узлы повреждаются или покидают сеть без освобождения среды.

DHCP является хорошим кандидатом для поддержки получения адресов передачи мобильных узлов. То же самое справедливо и для других необходимых параметров, таких как адреса и стандартные маршрутизаторы, серверы DNS и т.п. Следовательно, сервер DHCP должен располагаться в подсети точки доступа мобильного узла, или, по крайней мере, ретранслятор DHCP должен обеспечивать переправку сообщений.

К несчастью, с использованием DHCP связаны определенные проблемы. Одной из важнейших является безопасность. Пока нет системы аутентификации сообщений DHCP, это означает, что мобильный узел не может доверять серверу DHCP, а сервер DHCP не может доверять мобильному узлу. Кроме того, нет протокола для конфигурации сервер/сервер, т.е. один сервер DHCP не может связаться с другим сервером DHCP и обменяться текущими используемыми конфигурациями. Следовательно, конфигурирование серверов должно производиться вручную. Например, пространства адресов, которые серверы DHCP используют для настройки клиентов, должны быть непересекающимися. Администратору следует позаботиться, чтобы каждый сервер DHCP имел свое собственное пространство адресов клиентов. Как правило, это приводит к фрагментации пространства адресов.

9.3. Эпизодические сети

Поддержка мобильности, описываемая в разделах 9.1 и 9.2, существенно зависит от наличия хоть какой-то инфраструктуры. Протокол Mobile IP, например, требует наличия базового агента, туннелей и стандартных маршрутизаторов. DHCP требует наличия серверов и возможности распространения широковещательных сообщений для всех участников или ретрансляторов к серверам. Сотовые телефонные сети (см. главу 4) требуют базовых станций, инфраструктурных сетей и т.п.

В то же время существуют ситуации, когда пользователи сети не могут надеяться на инфраструктуру, инфраструктура слишком дорога или ее нет вообще. В такой ситуации единственной альтернативой будут эпизодические сети. Ниже приводятся примеры использования эпизодических сетей.

- **Мгновенная инфраструктура**. Незапланированные встречи, спонтанное общение между людьми и т.п. не могут полагаться на какую-либо инфраструктуру. Инфраструктура требует планирования и администрирования, но создание такой инфраструктуры будет слишком долгим, а значит, следует воспользоваться возможностью эпизодического соединения.
- **Зоны бедствия**. В зонах бедствия инфраструктура чаще всего бывает разрушена. Ураганы обрывают телефонные линии и линии электропередач, наводнения разрушают базовые станции, пожары уничтожают серверы. Таким образом, аварийные спасательные команды полагаются только на ту инфраструктуру, которую они создадут сами. Заранее спланировать ничего не удастся, поэтому настройка должна быть максимально быстрой и надежной. То же самое можно сказать и по отношению ко многим военным акциям.
- **Труднодоступные районы**. Даже если инфраструктуру можно спланировать заранее, иногда оказывается слишком дорого создавать ее в малонаселенных районах. В зависимости от структуры коммуникаций, эпизодические сети или спутниковые инфраструктуры могут оказаться вполне подходящими.
- **Эффективность**. Службы существующей инфраструктуры могут быть слишком дороги для некоторых приложений. Если, например, существует только сотовая сеть, ориентированная на установку соединения, а приложение ежеминутно отсылает лишь незначительную информацию о текущем состоянии, то более дешевая эпизодическая сеть с пакетной передачей может оказаться лучшим решением. Кроме того, процедура регистрации может быть слишком длительной, и служебные издержки могут не оправдывать передаваемую информацию. Таким образом, эпизодические сети, подстраиваемые под конкретные задачи, могут предложить лучшее решение.

Причина, по которой эпизодическим сетям посвящен отдельный раздел внутри главы о сетевом уровне, состоит в том, что маршрутизация данных — это один из сложнейших вопросов эпизодических сетей. Общая проблема маршрутизации рассмотрена в разделе 9.3.1, а в последующие разделы включены примеры реализации алгоритмов маршрутизации в

эпизодических сетях. Следует упомянуть, что функции маршрутизации иногда также имеются и на уровне 2, а не только на сетевом (третьем) уровне эталонной модели. Например, HIPERLAN (см. раздел 7.4) предлагает возможности пересылки/маршрутизации на уровне 2, основываясь на адресах MAC эпизодических сетей.

Одной из первых беспроводных эпизодических сетей бала пакетная радиосеть, запущенная ARPA в 1973 году. Она поддерживала до 138 узлов сети, а для передачи информации использовались IP-пакеты. Это позволяло осуществлять легкое подключение к ARPAnet, родоначальнице теперешней Internet. При использовании двадцати радиоканалов в диапазоне 1718,4–1840 МГц предлагались скорости передачи 100 или 400 Кбит/с. В системе использовалась схема DSSS со 128 или 32 элементарными сигналами/бит.

В данной эпизодической сети была использована разновидность дистанционно-векторной маршрутизации [Perlman, 1992]. При таком подходе каждый узел отсылает маршрутное объявление каждые 7,5 с. Эти объявления содержат таблицу соседей плюс список параметров качества связи с каждым из соседей. Основываясь на объявлениях, каждый узел обновляет локальную таблицу маршрутизации в соответствии с алгоритмом дистанционно-векторной маршрутизации. Принятые пакеты также помогают обновить таблицу маршрутизации. Далее отправитель посылает пакет первому соседу по маршрутному участку, используя таблицу соседей. Каждый узел переправляет принимаемые пакеты, основываясь на таблице соседей. Чтобы избежать появления замкнутых маршрутов и разрешить быстро изменять топологию, в этой простой схеме нужно осуществить некоторые модификации. В следующих разделах проблемы маршрутизации и вопросы усовершенствования механизмов маршрутизации в эпизодических сетях обсуждаются более подробно.

9.3.1. Маршрутизация

Если беспроводные сети с инфраструктурной поддержкой базовой станции имеют доступ ко всем мобильным узлам, это не всегда справедливо в случае эпизодической сети. Целевой узел может находиться вне досягаемости исходного узла, передающего пакеты. Таким образом, для нахождения пути между источником и пунктом назначения, а также для правильной пересылки пакетов необходима маршрутизация. В беспроводных сетях, использующих инфраструктуру, определены ячейки. Внутри ячейки базовая станция может связаться со всеми узлами посредством широковещания, не используя маршрутизацию. В ситуации же с эпизодической сетью каждый узел должен быть способен переправлять данные для других узлов. Из-за этого возникает много новых проблем, которые рассматриваются ниже.

На рис. 9.15 изображен простой пример эпизодической сети. В определенный момент времени t_1 топология сети может выглядеть, как показано в левой части рисунка. Пять узлов, от N_1 до N_5, соединены в зависимости от текущих характеристик передачи между ними. На этом снимке сеть N_4 может хорошо принимать данные, посылаемые N_1, но N_1 может принимать сообщения от N_4 лишь по слабой линии связи. Таким образом, линии связи необязательно должны иметь одинаковые характеристики в обоих направлениях. Причиной могут быть, например, различные характеристики антенн или разная излучаемая мощность. N_1 не может принимать передач от N_2 вовсе, а N_2 может только принимать сигналы N_1.

Данная ситуация может очень быстро измениться, как показано на снимке в момент времени t_2. N_1 больше вообще не может принимать сигналы от N_4, N_4 принимает сигналы от N_1, но через слабую линию связи. Теперь между N_1 и N_2 существует асимметричная двунаправленная линия связи, которой раньше не было.

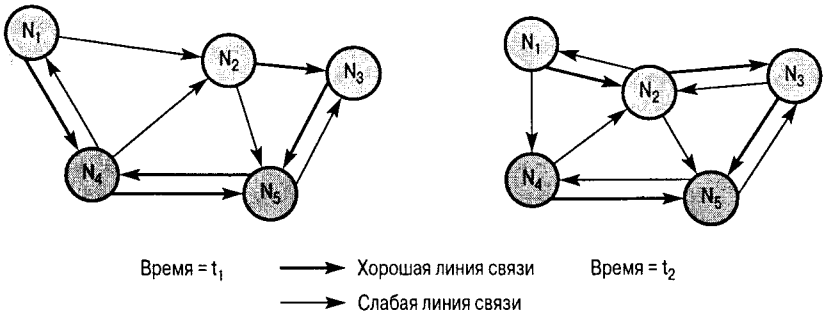

Рис. 9.15. Пример эпизодической сети

Этот очень простой пример показывает фундаментальное отличие эпизодической беспроводной сети от проводной сети с точки зрения маршрутизации.

- **Асимметричные линии связи**. Если узел А принимает сигнал от узла Б, то это еще ничего не говорит о качестве связи в обратном направлении. Б может вообще ничего не принимать, иметь слабую связь или же иметь даже лучшую связь, чем в противоположном направлении. Таким образом, маршрутная информация, собранная по одному направлению, практически бесполезна для другого. Впрочем, многие алгоритмы маршрутизации для проводных сетей основываются на симметричном сценарии.
- **Избыточные линии связи**. Проводные сети также имеют избыточные линии связи для защиты от сбоев. В то же время в проводных сетях имеется лишь некоторая избыточность, которая к тому же контролируется сетевым администратором. В эпизодических сетях никто не контролирует избыточность, так что в данном случае может быть много избыточных линий связи, вплоть до появления решетчатой топологии. Алгоритмы маршрутизации могут хорошо справляться с небольшой избыточностью, но сильная избыточность может создавать слишком большую вычислительную нагрузку по обновлению таблиц маршрутизации.
- **Интерференция**. В проводных сетях связь существует лишь при наличии проводников, а соединения планируются сетевыми администраторами. Для беспроводных эпизодических сетей это не так. Линии связи появляются и исчезают в зависимости от характера передачи, а одна передача может интерферировать с другой, узлы могут перехватывать сообщения от других узлов. Таким образом, интерференция создает новые проблемы появления "незапланированных" линий связи между узлами: если два отдельных узла осуществляют две передачи, то эти передачи могут интерферировать и подавить друг друга. С другой стороны, интерференция может облегчить маршрутизацию. Узел может изучить топологию с помощью пакетной служебной информации.
- **Динамическая топология**. Наиболее серьезную проблему для маршрутизации представляет высокодинамичная топология. Мобильные узлы могут перемещаться, как показано на рис. 9.15, или могут изменяться характеристики среды. Это приводит к частому изменению топологии, так что данное расположение будет действительным в течение очень короткого времени. В эпизодических сетях таблицы маршрутизации должны так или иначе отображать эти частые изменения в топологии и подстраивать алгоритмы маршрутизации. Алгоритмы маршрутизации, используемые в проводных сетях, или будут реагировать слишком медленно, или будут порождать слишком много обновлений, пытаясь отобразить все изменения топологии. Таблица маршрутизации в проводной сети обновляется, например, каждые 30 секунд. Такая частота обновления может оказаться слишком низкой, чтобы ее можно было применить в эпизодических сетях. Кроме того, некоторые алгоритмы зависят от пол-

ной схемы всей сети. Это работает в проводной сети, где обновления довольно редки, но это абсолютно неприемлемо в эпизодических сетях. Топология меняется во время распространения "текущего" состояния, и фиксация этого состояния бессмысленна.

Вернемся к примеру сети, изображенному на рис. 9.15, и предположим, что узел N_1 хочет отослать данные узлу N_3 и нуждается в подтверждении. Если N_1 имеет полную картину сети (которая не совсем является эпизодической) на момент t_1, он должен выбрать путь N_1, N_2, N_3, так как это потребует лишь двух маршрутных участков (если в качестве метрики рассматриваются маршрутные участки). Подтверждение не может придти тем же путем, N_3 выбирает путь N_3, N_5, N_4, N_1. Это потребует трех маршрутных участков, и можно видеть, что маршрутизация сильно влияет на функции более высоких уровней. Например, TCP при вычислениях, связанных с периодом кругового обращения пакета, предполагает одинаковый путь в обоих направлениях. Очевидно, что в рассматриваемом примере это не так, что приведет к неверному толкованию измерений и неэффективности (см. главу 10).

В следующее мгновение (момент t_2) топология изменяется. Теперь N_3 больше не может использовать тот же путь для отправки подтверждения обратно к N_1, тогда как N_1 все еще доступен старый путь к N_3. Хотя сложившаяся ситуация уже является более сложной, чем стационарная сеть, предполагается, что узлам известна текущая топология сети. Оптимальной информацией для каждого узла было бы описание существующей на данный момент связности всех узлов, ожидаемых информационных потоков, пропускной способности линий связи, задержек на каждой из линий и вычислительной мощности каждого узла. Отметим, что даже в стационарных сетях информационные потоки трудно предсказуемы, а в эпизодических сетях к тому же неизвестны пропускные способности линий связи. Пропускная способность каждой линии связи может меняться от 0 до максимальной величины, которую допускает технология передачи. В реальных эпизодических сетях ни одному узлу не известны все эти факторы, и сбор самой свежей информации о текущем состоянии почти невозможен.

Использование в эпизодических сетях мобильных компьютеров порождает новые проблемы, связанные с ограничениями аппаратного обеспечения. Использование стандартных протоколов маршрутизации с периодическим обновлением приводит к трате энергии источниками питания даже без передачи пользовательских данных и исключает использование режима ожидания. Кроме того, периодическое обновление расходует пропускную способность — и без того скудный ресурс беспроводных линий связи.

Еще одной проблемой, отсутствующей в проводных сетях, является интерференция между двумя и более передачами, которые не используют одинаковые узлы в ходе пересылки. Если, например, вторая передача из узла N_4 в узел N_5 (см. рис. 9.15) происходит одновременно с передачей от N_1 к N_3, то они могут интерферировать. Интерференция может возникнуть в N_2, когда он принимает сигналы от N_1 и N_4, или в N_5, когда он примет сигналы от N_4 и N_2. Интерференции можно избежать, если правильно экранировать два проводника.

Рассматривая все дополнительные трудности, возникающие в сравнении с проводными сетями, можно сделать следующие выводы относительно эпизодических сетей с подвижными узлами.

- Традиционные алгоритмы маршрутизации, применяемые в проводных сетях, работают неэффективно (например, дистанционно-векторные алгоритмы, такие как RIP [Hendrik, 1988], [Malkin, 1994], сходятся слишком медленно) или целиком непригодны (например, алгоритм маршрутизации, при котором путь выбирается на основании информации о состоянии канала, такой как OSPF [Moy, 1994], требует полного знания состояния сети). Эти алгоритмы разрабатывались без учета высокодинамичной топологии асимметричных линий связи или интерференции.

- Маршрутизация в беспроводных эпизодических сетях не может полагаться только на информацию третьего уровня. Информация с нижних уровней, касающаяся связности или интерференции, может значительно облегчить задачу алгоритмов маршрутизации по отысканию хорошего пути.
- Централизованные подходы не работают, поскольку они требуют слишком много времени на сбор сведений о текущем состоянии и передачу этой информации обратно. За это время топология может поменяться.
- Многим узлам нужны возможности маршрутизации. Хотя некоторые узлы смогут без этого обойтись, но, по крайней мере, один маршрутизатор должен быть в пределах досягаемости каждого узла. Алгоритмы должны экономно расходовать ограниченную мощность источников питания этих узлов.
- Понятие соединения с некоторыми характеристиками не может использоваться адекватно. Эпизодические сети будут работать без установления соединения, в быстро меняющемся окружении невозможно поддерживать соединение и пересылать по этому соединению данные. Узлам нужно принимать локальные решения для пересылки и отправки пакетов приблизительно в направлении пункта назначения.
- Последней альтернативой для пересылки пакетов при неизвестной топологии является лавинная маршрутизация. Эта модель всегда работает при низкой загрузке, но она очень неэффективна. Чтобы избежать зацикливания, к каждому пакету прикрепляется счетчик маршрутных участков, причем должен быть известен диаметр эпизодической сети, т.е. максимальное число маршрутных участков. (Как верхнюю оценку диаметра сети можно использовать количество узлов.)
- Может оказаться полезной иерархическая кластеризация узлов. Если удастся идентифицировать несколько групп узлов, связанных вместе, можно создать кластеры. Тогда как отдельные узлы могут перемещаться быстрее, целый кластер может быть более стационарным. Следовательно, маршрутизация между кластерами может оказаться проще и менее динамичной (см. раздел 9.3.4.).

В последующих разделах приводится несколько примеров алгоритмов маршрутизации, используемых в эпизодических сетях, также рассматриваются удобные метрики, отличные от обычного подсчета маршрутных участков.

9.3.2. Дистанционно-векторная маршрутизация с нумерацией и оценкой пути к пункту назначения

Дистанционно-векторная маршрутизация с нумерацией и оценкой пути к пункту назначения (destination sequence distance vector — DSDV) — это модификация дистанционно-векторной маршрутизации для использования в эпизодических сетях [Perkins, 1994]. Дистанционно-векторная маршрутизация используется как протокол RIP в проводных сетях. При определенных изменениях в сети эта схема работает очень плохо в связи с проблемой "подсчета бесконечности". Каждый узел периодически обменивается с близлежащими узлами таблицей соседей. Таким образом, сведения об изменении в одном узле медленно распространятся по всей сети (шаг за шагом по мере обмена). Методы борьбы, используемые в стационарных сетях (обратное заражение/расщепление горизонта [Perlman, 1992]), не помогают в эпизодической сети из-за быстрого изменения топологии. По этой же причине в сети могут возникать циклы и недосягаемые области.

DSDV вводит в дистанционно-векторный алгоритм два усовершенствования.

- **Порядковые номера**. Каждое объявление маршрутизации приходит со своим порядковым номером. В эпизодической сети объявления могут распространяться множеством путей.

Порядковые номера призваны помочь правильному толкованию объявлений. Это позволяет избегать циклов подобно исходной дистанционно-векторной маршрутизации.

- **Демпфирование**. Краткие переходные процессы, возникающие при изменении топологии, не должны дестабилизировать механизмы маршрутизации. Следовательно, объявления, содержащие информацию о текущей топологии, не распространяются дальше. Узел откладывает распространение, если изменения с большой вероятностью неустойчивы. Время ожидания зависит от времени между первым и лучшим объявлением пути к определенному пункту назначения.

Таблица маршрутизации для N_1 из рис. 9.15 может выглядеть подобно табл. 9.1.

Таблица 9.1. Часть таблицы маршрутизации для DSDV

Пункт назначения	Следующий маршрут	Метрика	Порядковый номер	Время запуска
N_1	N_1	0	$S_1 - 321$	$T_4 - 001$
N_2	N_2	1	$S_2 - 218$	$T_4 - 001$
N_3	N_2	2	$S_3 - 043$	$T_4 - 002$
N_4	N_4	1	$S_4 - 092$	$T_4 - 001$
N_5	N_4	2	$S_5 - 163$	$T_4 - 002$

Для каждого узла в N_1 хранится информация о следующем маршруте к этому узлу, метрике (в данном случае количестве маршрутов), порядковом номере последнего объявления для этого узла и времени первого введения в действие этого пути. Кроме того, в таблице содержатся метки и время установления сигнала, которые помогают решить, когда же путь будет считаться стабильным. Объявление маршрутизатора от N_1 теперь содержит данные из первого, третьего и четвертого столбцов: адрес пункта назначения, метрику и порядковый номер. Помимо отсутствия циклов схема DSDV малотребовательна к памяти и быстро сходится при использовании присоединенных процедур, инициируемых при заданных условиях.

9.3.3. Динамическая маршрутизация от источника

Представим, что произойдет в эпизодической сети, где узлы время от времени обмениваются пакетами, т.е. сеть лишь слегка загружена, а для обновления таблиц маршрутизации применяется DSDV или другой алгоритм, используемый для обновления таблиц маршрутизации. Хотя передаются лишь некоторые пользовательские данные, узлы все же обмениваются маршрутной информацией для отслеживания топологии. Эти алгоритмы поддерживают маршруты между всеми узлами, хотя на данный момент может и не быть данных для передачи, что приводит к возникновению ненужных информационных потоков и не позволяет узлам экономить энергию.

Динамическая маршрутизация от источника решает задачу маршрутизации как две отдельных проблемы [Johnson, 1996].

- **Обнаружение маршрута**. Узел лишь пытается обнаружить маршрут к пункту назначения, если у него есть что пересылать по этому адресу и маршрут сейчас неизвестен.
- **Поддержание маршрута**. Если узел постоянно отсылает пакеты по маршруту, это означает, что маршрут еще действенен. Как только узел обнаруживает проблемы с текущим маршрутом, он начинает поиск альтернативного маршрута.

Основной принцип маршрутизации сообщений от источника также используется и в стационарных сетях, например, сетях Token Ring. Динамическая маршрутизация от источника исключает периодические обновления маршрутов и работает следующим образом.

Когда узлу нужно обнаружить маршрут, он производит широковещательную передачу, передавая запрос на маршрут с уникальным идентификатором и адресом пункта назначения как параметрами. Любой узел, который примет запрос на маршрут, делает следующее.

- Если узел уже получал этот запрос (который однозначно определяется с помощью идентификатора), он отбрасывает пакет запроса.
- Если узел узнает свой собственный адрес в качестве адреса пункта назначения, то запрос достигает цели.
- В противном случае узел добавляет в пакет свой собственный адрес к списку пройденных маршрутных участков и опять с помощью широковещательной передачи отправляет такой обновленный запрос дальше.

Используя эту модель, запрос на маршрут собирает список адресов, представляющих возможные пути к определенному пункту назначения. Таким образом, как только запрос достигнет места назначения, его можно вернуть вместе с содержащимся в нем списком получателям, которые смогут использовать этот список в обратном порядке. Однако, здесь имеется одно условие — линия связи должна быть двунаправленной. Если это не так, и узел пункта назначения не поддерживает обратный маршрут к инициатору запроса, то он начинает самостоятельный поиск маршрута. Пункта назначения могут достичь несколько списков, содержащих несколько разных путей от инициатора. Это должно дать инициатору возможность использовать либо первый путь, либо наилучший путь или несколько путей сразу.

Использование алгоритма обнаружения маршрута для примера, показанного на рис. 9.15, при передаче от узла N_1 до узла N_3 дает следующее.

- N_1 по широковещательной передаче распространяет запрос ((N_1), id = 42, цель = N_3), N_2 и N_4 принимают этот запрос.
- N_2, используя широковещательную передачу, переправляет запрос ((N_1, N_2), id = 42, цель = N_3), N_4 переправляет сообщение ((N_1, N_4), id = 42, цель = N_3). N_3 и N_5 получают запрос от N_2, а N_1, N_2 и N_5 — от N_4.
- N_3 обнаруживает себя в качестве пункта назначения, N_5 пересылает запрос ((N_1, N_2, N_5), id = 42, цель = N_3). N_3 и N_4 получают сообщение от N_5. N_1, N_2 и N_5 отбрасывают пакет от N_4, поскольку узнают уже полученный запрос на маршрут (N_5 получает сообщение от N_2 до прихода сообщения от N_4).
- N_4 отбрасывает сообщение от N_5, N_3 узнает запрос (N_1, N_2, N_5) в качестве альтернативного, но более длинного маршрута.
- Теперь N_3 возвращает N_1 путь (N_1, N_2, N_5). Это предполагает наличие двунаправленной линии связи. N_3 может передавать информацию согласно обратному порядку списка.

Предположение о двунаправленной линий связи оправдывается для многих эпизодических сетей. Однако, если линии связи не являются двунаправленными, алгоритм может оказаться сложнее, и его придется повторить снова, но уже в обратном направлении, если пункт назначения в данный момент не поддерживает путь к источнику запроса на маршрут.

- N_3 по широковещательной передаче распространяет запрос ((N_3), id = 17, цель = N_1), этот запрос принимает только N_5.
- Тогда N_5, используя широковещательную передачу, переправляет запрос ((N_3, N_5), id = 17, цель = N_1), N_3 и N_4 получают сообщение от N_5.

- N_3 узнает уже отправленный идентификатор id и отбрасывает сообщение, N_4 пересылает запрос ((N_3, N_5, N_4), id = 17, цель = N_1). N_5, N_2 и N_1 получают это сообщение.
- N_5 выбрасывает пакет запроса, N_1 узнает себя в качестве цели, и N_2 пересылает запрос ((N_3, N_5, N_4, N_2), id = 17, цель = N_1). N_3 и N_5 получают его сообщение.
- N_3 и N_5 отбрасывают пакет запроса.

Теперь у N_3 имеется список путей от N_1 к N_3, т.е. (N_1, N_2, N_3), и N_1 известен путь от N_3 к N_1, а именно, (N_3, N_5, N_4, N_1). Но N_1 все еще не знает, как отсылать данные N_3! Единственным решением представляется отправка N_3 широковещательного сообщения со списком (N_1, N_2, N_3) в обратном направлении. Этот пример показывает, насколько проще реализуется маршрутизация при симметричных линиях связи.

Оптимизацию базового алгоритма обнаружения маршрута можно осуществлять по-разному.

- Чтобы избежать слишком многих широковещательных передач, каждый запрос на маршрут может содержать счетчик. Каждый узел, осуществляющий повторную передачу запроса, увеличивает счетчик на единицу. Зная максимальный диаметр сети (можно взять количество узлов, если больше ничего не известно), узлы могут отбрасывать запрос после достижения счетчиком этого значения.
- Узлы могут кэшировать фрагменты путей от недавних запросов. Эти фрагменты можно использовать для быстрого ответа на запросы (если фрагменты все еще отвечают топологии!).
- Узлы могут обновлять кэш из заголовков пересылаемых пакетов.
- Если узел случайно получит передачу от других узлов, он может воспользоваться также и этой информацией для сокращения маршрутов.

После обнаружения маршрута он поддерживается, пока по нему узел посылает пакеты. В зависимости от механизмов второго уровня можно использовать различные модели.

- Если для подтверждения используется канальный уровень, как в сети IEEE 802.11, узел может интерпретировать это подтверждение как доказательство существования маршрута.
- Если возможно, узел также должен прислушиваться к передаче пакета следующим узлом, что является косвенным подтверждением приема предыдущего пакета.
- Кроме того, узел должен запрашивать явное подтверждение.

Опять же, ситуация усложняется при асимметричных линиях связи. Если узел обнаруживает проблемы со связью, он должен информировать отправителя пакета, инициируя таким образом новое обнаружение маршрута, которое начинается с отправителя. Или же узел может попытаться самостоятельно обнаружить новый маршрут.

Таким образом, динамическая маршрутизация сообщений от источника по сравнению с другими алгоритмами дает определенные выгоды за счет более эффективного использования пропускной способности, хотя при наличии асимметричных линий связи и высокодинамичной топологии появляются проблемы.

9.3.4. Иерархические алгоритмы

Оба алгоритма маршрутизации, дистанционно-векторная маршрутизация с нумерацией и оценкой пути к пункту назначения и динамическая маршрутизация от источника, могут работать лишь при небольшом количестве узлов и сильно зависят от мобильности узлов. Для больших сетей лучшим подходом может оказаться кластеризация узлов и использование разных алгоритмов маршрутизации между кластерами и внутри них. Это объясняется свойством локаль-

ности, т.е. если можно создать кластер, узлы остаются в нем, пока кластер не изменится. Таким образом, если топология внутри кластера изменится, об этом будут знать лишь узлы этого кластера. Узлам прочих кластеров достаточно знать лишь то, как достичь кластера. Модель, по сути, скрывает все мелкие подробности удаленных кластеров.

Время от времени каждому узлу нужны сведения о топологии. И опять же обновления из удаленных кластеров будут рассылаться реже, по сравнению с локальным обновлением. Кластеры можно объединять в суперкластеры и т.п., что дает иерархическую структуру. Используя этот подход, один или более узлов могут функционировать как центр кластера, представляя собой маршрутизатор для всех информационных потоков в кластер и из него. Все узлы кластера и все центры других кластеров используют его в качестве шлюза этого кластера. На рис. 9.16 показана эпизодическая сеть с подключением к Internet через базовую станцию. Эта базовая станция пересылает данные к центрам кластеров и обратно. В данном примере один из центров кластеров также играет роль центра суперкластера, маршрутизируя таким образом все информационные потоки в суперкластер и обратно.

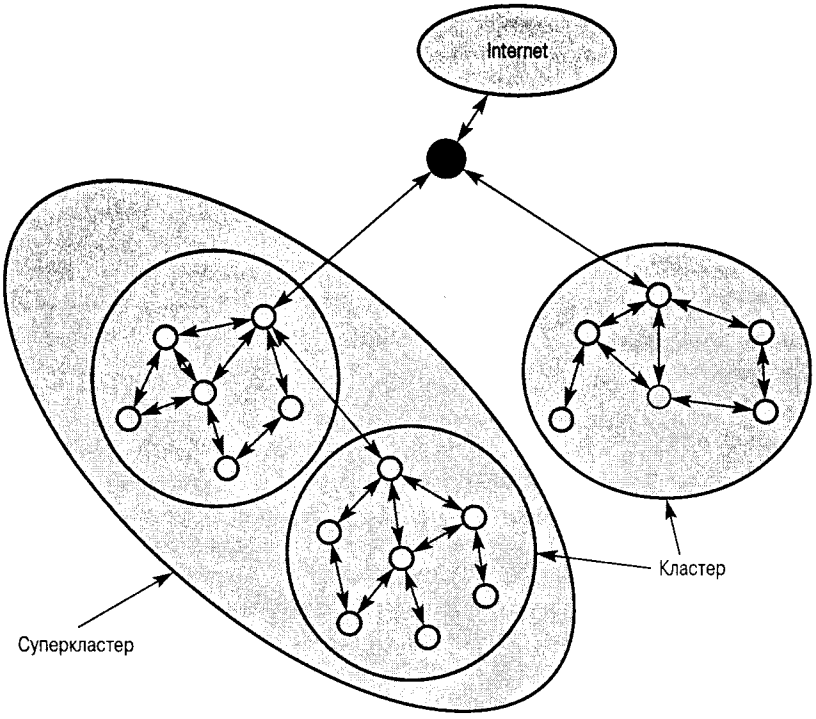

Рис. 9.16. Иерархическая структура в эпизодической сети

9.3.5. Альтернативные метрики

Во всех примерах, использовавшихся в этой главе, в качестве метрики маршрутизации всегда рассматривалось количество маршрутных участков. Несмотря на свою простоту, в особенности в беспроводных и эпизодических сетях, такой выбор не всегда является наилучшим. Даже в стационарных сетях в качестве метрики маршрутизации может, например, служить пропускная способность. При изменении качества линии связи и интерференции разных передач более удобными могут оказаться другие метрики.

Одна из таких метрик, называемая **маршрутизация с минимальной интерференцией** (least interference routing — LIR), позволяет учитывать интерференцию. На рис. 9.17 изображена топология эпизо-

дической сети. Отправитель S_1 хочет отослать пакет получателю R_1, а отправитель S_2 — получателю R_2. Используя в качестве метрики счетчик количества маршрутных участков, S_1 может выбрать три разных пути с тремя маршрутными участками, являющиеся минимальными при используемой метрике. Возможными путями являются (S_1, N_3, N_4, R_1), (S_1, N_3, N_2, R_1), (S_1, N_1, N_2, R_1). S_2 может воспользоваться лишь одним доступным путем из трех маршрутных участков (S_2, N_5, N_6, R_2). Если теперь учесть интерференцию, то картина изменится. Для расчета возможной интерференции пути каждый узел делает это для своих маршрутных участков. В данном случае под интерференцией понимается число соседей, которые смогут подслушать передачу. Следовательно, каждому узлу для расчета интерференции необходима локальная информация.

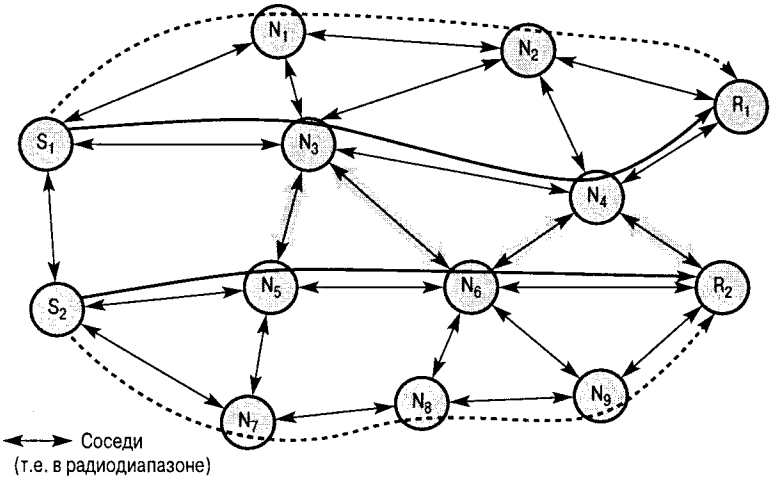

Рис. 9.17. Пример маршрутизации с минимальной интерференцией

В этом примере интерференция узла N_3 равна 6, узла N_4 равна 5 и т.д. Расчет стоимости возможных путей между S_1 и R_1 дает следующее:

$$C_1 = \text{стоимость } (S_1, N_3, N_4, R_1) = 16;$$
$$C_2 = \text{стоимость } (S_1, N_3, N_2, R_1) = 15;$$
$$C_3 = \text{стоимость } (S_1, N_1, N_2, R_1) = 12.$$

Все три пути состоят из одинакового количества маршрутных участков, но последний путь имеет более низкую интерференцию. Следовательно, S_1 выбирает путь (S_1, N_1, N_2, R_1). S_2 также рассчитывает стоимости разных путей, т.е. C_4 = стоимость (S_2, N_5, N_6, R_2) = 16 и C_5 = стоимость (S_2, N_7, N_8, N_9, R_2) = 15. Следовательно, S_2 должен выбрать путь (S_2, N_7, N_8, N_9, R_2), хотя этот путь состоит из большего количества маршрутных участков, чем первый выбранный.

Если одновременно ведется две передачи, между ними возникает интерференция, как показано на рис. 9.17. Этого можно избежать, используя маршрутизацию с минимальной интерференцией. Учитывая лишь локальную среду и не думая, по каким путям пойдут другие передачи, эта схема безусловно может снизить вероятность интерференции. Полностью устранить интерференцию можно лишь в том случае, если все отправители будут знать обо всех других передачах (и обо всей топологии в целом) и учитывать это при маршрутизации.

Кроме того, маршрутизация может учитывать несколько метрик одновременно и взвешивать их. Метриками могут быть количество маршрутных участков h, интерференция i, надежность r, частота появления ошибок e и другие. Стоимость пути тогда будет определяться следующим образом.

$$\text{Стоимость} = \alpha h + \beta i + \gamma r + \delta e + \ldots$$

Подобрать весовые коэффициенты α, β, γ, δ, ..., чтобы получить желаемые характеристики маршрутизации, не так легко (если вообще возможно).

9.4. Резюме

Поддержка мобильности на сетевом уровне имеет особую важность, поскольку сетевой уровень с помощью обычного протокола IP связывает Internet воедино. Несмотря на различия, лежащие в основе используемых проводных и беспроводных технологий, все узлы должны иметь возможность обмена информацией. Следовательно, протокол Mobile IP (как расширение классического IP) разрабатывался с целью обеспечения мобильность в Internet, не затрагивая при этом существующие проводные системы. В то же время протокол Mobile IP оставляет неразрешенными несколько вопросов, в особенности касающихся безопасности, эффективности потока пакетов и поддержки качества обслуживания. Многие вопросы решены в протоколе IP version 6 (IPv6), который несколько упрощает работу с мобильными устройствами.

DHCP предлагает полностью автоматический механизм обеспечения узла всей необходимой информацией для интеграции в сеть, поддерживая, таким образом, процесс подключения новых компьютеров и интеграцию мобильных компьютеров в сеть. DHCP является главным источником адресов передачи, необходимых протоколу Mobile IP. Но у DHCP также имеются некоторые проблемы, в основном это отсутствие механизмов обеспечения безопасности и протокола сервер/сервер, что требует ручной настройки сервера DHCP.

Наконец, эпизодические сети предлагают абсолютно новый путь организации мобильной связи при отсутствии инфраструктуры. В этих сетях главным вопросом является маршрутизация, поскольку здесь нет базовой станции, имеющей доступ ко всем узлам через широковещательную связь, как в сотовых сетях. Традиционные алгоритмы маршрутизации оказываются неэффективны в высокодинамичной среде эпизодических сетей, а значит, потребуется модификация существующих или привлечение совершенно новых алгоритмов. Похоже, что для больших групп узлов единственным решением проблемы маршрутизации может оказаться иерархический подход, а такие алгоритмы, как DSDV, не позволяют провести хорошее масштабирование. Важным отличием беспроводных сетей является потребность в информации о характеристиках второго уровня. При нахождении хорошего маршрута в качестве метрики могут использоваться сведения об интерференции и подтверждениях.

Многие вопросы, касающиеся мобильного сетевого уровня, остаются открытыми. Важно видеть, как много различных подходов можно скомбинировать. Например, расширение возможностей чужих агентов, допускающих множественные маршруты к мобильным узлам, можно объединить с протоколом Mobile IP в эпизодических сетях. В последнее время появилось множество работ по созданию эпизодических вычислительных сетей с протоколом Mobile IP. Группа развития мобильных эпизодических сетей (mobile ad hoc networking group — MANET) предложила новые механизмы инкапсуляции и протоколы маршрутизации [Broch, 1998], [Corson, 1998], [Perkins, 1998], [Corson, 1999]. В общем, сетевой уровень — это обширная открытая и очень интересная область исследований.

9.5. Упражнения

1. Объясните маршрутизацию в стационарных сетях с протоколом IP. Назовите проблемы и последствия использования IP совместно со стандартными протоколами маршрутизации в мобильной связи.
2. Какими могут быть быстрые "решения" и почему они не работают?

3. Назовите требования к протоколу Mobile IP и обоснуйте их. Удовлетворяет ли протокол всем им?
4. Составьте список объектов протокола Mobile IP и опишите передачу данных от мобильного узла к стационарному и наоборот. Зачем и где требуется инкапсуляция пакетов?
5. Как осуществляется регистрация на третьем уровне при работе мобильного узла?
6. Укажите все этапы, нужные для переключения с одного чужого агента на другой, включая уровни 2 и 3.
7. Опишите поток пакетов между двумя мобильными узлами, если оба они находятся в чужой сети. Какими дополнительными маршрутами воспользуются пакеты, если потребуется обратное туннелирование?
8. Объясните, каким образом работает туннелирование в общем и, в частности, в протоколе Mobile IP при минимальной и обычной инкапсуляции IP-IP, а также специальной маршрутной инкапсуляции соответственно. Выясните преимущества и недостатки этих трех методов.
9. Перечислите слабые стороны протокола Mobile IP, рассмотрев передачу данных от корреспондентского узла к мобильному. Как можно оптимизировать этот процесс и какие дополнительные проблемы это вызовет?
10. Назовите преимущества использования IPv6 в мобильной связи? Какими теперь будут объекты протокола Mobile IP?
11. Каковы основные проблемы протокола Mobile IP, касающиеся безопасности и поддержки качества обслуживания.
12. Какова основная задача DHCP? Перечислите объекты DHCP.
13. Как можно использовать DHCP в мобильной связи и для поддержки протокола Mobile IP?
14. Назовите основные нерешенные вопросы, касающиеся DHCP.
15. Перечислите основные отличия эпизодических сетей от прочих сетей. Какие выгоды дает использование эпизодических сетей?
16. В чем заключается сложность маршрутизации в эпизодических сетях, в чем состоят особые трудности?
17. Опишите алгоритмы дистанционно-векторной маршрутизации и маршрутизации на основе оценки локального состояния в стационарных сетях. Почему оба эти алгоритма сложно применить в эпизодических сетях?
18. В чем состоит различие между обычной дистанционно-векторной маршрутизацией и дистанционно-векторной маршрутизацией с нумерацией и оценкой пути к пункту назначения? В чем необходимость такой модификации?
19. Как осуществляется маршрутизация в алгоритме динамической маршрутизации от источника? В чем состоит смысл использования динамической маршрутизации от источника по сравнению с другими алгоритмами маршрутизации в стационарных сетях?
20. Как влияет симметрия линий связи на предлагаемые алгоритмы маршрутизации?
21. Представьте эпизодическую сеть с быстро перемещающимися узлами, например, автомобили в городе. Какие возникнут проблемы, если даже используются алгоритмы маршрутизации, приспособленные к эпизодическим сетям? Какой будет ситуация на автостраде?

9.6. Рекомендуемая литература

Alexander, S.; Drohms, R. (1997) *DHCP option and BOOTP vendor extensions*, RFC 2132, March.

Brewer, E. A.; Katz, R. H.; Chawathe, Y.; Gribble, S. D.; Hodes, T.; Nguyen, G.; Stemm, M.; Henderson, T.; Amit, E.; Balakrishnan, H.; Fox, A.; Padmanabhan, V.; Seshan, S. (1998) "A network architecture for heterogeneous mobile computing", *IEEE Personal Communications*, 5, (5).

Broch, J.; Johnson, D. B.; Maltz, D. A. (1998) *The dynamic source routing protocol for mobile ad hoc networks*, draft-ietf-manet-dsr-01.txt (в разработке), December.

Corson, S.; Macker, J. (1998) *Mobile Ad hoc Networking (MANET): Routing protocol performance issues and evaluation considerations*, draft-ietf-manet-issues-02.txt (в разработке), October.

Corson, S.; Papademetriou, S.; Papadopulos, P.; Park, V.; Qayyum, A. (1999) *An Internet MANET Encapsulation Protocol (IMEP) specification*, draft-ietf-manet-spec-01.txt (в разработке), February.

Deering, S. (1989) *Host extensions for IP multicasting*, RFC 1112, August.

Deering, S. (1991) *ICMP router discovery messages*, RFC 1256, September.

Deering, S.; Hinden, R. (1995) *Internet Protocol, version 6 (IPv6) Specification*, RFC 1883, December.

Drohms, R. (1997) *Dynamic Host Configuration Protocol*, RFC 2131, March.

Hanks, S.; Li, T.; Farinacci, D.; Traina, P. (1994) *Generic Routing Encapsulation (GRE)*, RFC 1701, October.

Hendrick, C. (1988) *Routing Information Protocol*, RFC 1058, June.

Jacobson, V. (1990) *Compressing TCP/IP headers for low-speed serial links*, RFC 1144, February.

Johnson, D. B.; Maltz, D. A. (1996) *Dynamic source routing in ad hoc wireless networks*, в Mobile Computing (Imielinski/Korth), Kluwer Academic Publishers.

Johnson, D. B.; Perkins, C. (1998) *Mobility support in IPv6*, draft-ietf-mobileip-ipv6-07.txt (в разработке), November.

Malkin, G. (1994) *RIP version 2 carrying additional information*, RFC 1723, November.

Montenegro, G. (1998) *Reverse Tunneling for Mobile IP*, RFC 2344, May.

Moy, J. (1994) *OSPF version 2*, RFC 1583, March.

Perkins, C. (1996a) *IP Mobility Support*, RFC 2002, October.

Perkins, C. (1996b) *IP Encapsulation within IP*, RFC 2003, October.

Perkins, C. (1996c) *Minimal Encapsulation within IP*, RFC 2004, October.

Perkins, C.; Johnson, D. B. (1996d) *Mobility support in IPv6*, proceedings of ACM Mobicom 96, November.

Perkins, C. (1997) *Mobile IP: Design Principles and Practice*. Addison Wesley Longman Publishing Company, Reading, Massachusetts, USA.

Perkins, C. (1998) *Mobile ad hoc networking terminology*, draft-ieft-manet-term-01.txt (в разработке), November.

Perlman, R. (1992) *Interconnections: Bridges and Routers*, Addison-Wesley Longman Publishing Company, Reading, Massachusetts, USA.

Postel, J. B. (1981) *Internet Protocol*, RFC 791, September.

Stevens, W. R. (1994) *TCP/IP Illustrated, Volume1: The Protocols*. Addison Wesley Longman Publishing Company, Reading, Massachusetts, USA.

Мобильный транспортный уровень 10

Поддержки мобильности только на нижних уровнях, вплоть до сетевого уровня, недостаточно для обеспечения поддержки мобильности прикладных систем. Большинство приложений зависят от транспортного уровня, такого как TCP или UDP в случае Internet. У Internet-приложений транспортный уровень имеет две функции — использование контрольных сумм в пользовательских данных и уплотнение/разуплотнение данных, поступающих от прикладных систем или передаваемых этим системам. Тогда как сетевой уровень снабжает адресом лишь хост, порты в TCP или UDP допускают обращение к отдельным приложениям. Это все, что предлагает UDP без соединения, следовательно, все перечисленное ниже сосредоточивается в TCP. Тогда как UDP работает без установления соединения и не дает никаких гарантий надежной доставки информации, TCP значительно сложнее, а значит, ему требуются специальные механизмы для работы в мобильной среде. Для работы UDP вполне достаточно поддержки мобильности на уровне IP (например, использования Mobile IP).

Главное различие между UDP и TCP состоит в том, что TCP обеспечивает соединение между двумя приложениями. В соединении TCP может выполнять несколько функций, таких как упорядоченная доставка или надежная передача данных с использованием методов повторной передачи. Кроме того, TCP имеет встроенные механизмы для удобной работы именно в сети. Если, к примеру, TCP обнаружит потерю пакета, он предполагает внутрисетевую перегрузку и снижает скорость передачи. Это также одна из причин использования протоколов, подобных TCP, и отказа от более простого UDP. UDP требует, чтобы приложению предоставлялась надежность, упорядоченная доставка и т.п. Кроме того, UDP не является удобным сетевым протоколом, т.е. он не инициирует замедления в случае перегрузки.

В последующих разделах предлагается обзор механизмов TCP, которые играют важную роль при использовании TCP в мобильных системах. Основной проблемой многих механизмов является то, что они разрабатывались на основе идей, абсолютно отличных от лежащих в основе мобильных сетей. Учитывая эти проблемы, которые могут привести к полному сбою потока данных TCP, был разработан целый комплекс решений. Некоторые предложения представлены в разделах 10.2 и 10.8; каждое из предложений имеет свои специфические сильные и слабые стороны.

10.1. Традиционный TCP

В этом разделе рассматриваются несколько механизмов протокола управления передачей (transmission control protocol — TCP) [Postel, 1981], которые влияют на эффективность TCP в мобильной среде. Очень подробное описание TCP дается также в работе [Stevens, 1994].

10.1.1. Управление перегрузкой

Протокол транспортного уровня, подобный TCP, разрабатывался для стационарных сетей со стационарными конечными системами. Передача данных осуществляется с использованием

сетевых адаптеров, волоконно-оптических кабелей, медных проводников, специального оборудования для маршрутизаторов и т.п. Как правило, эта аппаратура функционирует без введения ошибок передачи. Более того, если программное обеспечение достаточно давно и успешно используется, не происходит потери пакетов или инверсии битов. Следовательно, если в стационарной сети происходит потеря пакета при пересылке между отправителем и получателем, то чаще всего это происходит не по вине аппаратных или программных сбоев. Таким образом, причиной высоких потерь пакетов в стационарной сети, вероятно, является временная перегрузка в некоторой точке маршрута передачи, т.е. состояние перегрузки определенного узла.

Перегрузка может появляться время от времени даже в очень тщательно разработанных сетях. Пакетные буферы маршрутизаторов заполняются, и маршрутизаторы не могут достаточно быстро переправлять пакеты, так как сумма входных скоростей пакетов, предназначенных на один выход, больше пропускной способности этого выходного канала связи. Единственное, что может сделать в сложившейся ситуации маршрутизатор, так это отбросить некоторые пакеты. Отброшенные пакеты потеряны для передачи, и получатель замечает прореху в потоке пакетов. Теперь получатель не сообщает отправителю прямо, какие именно пакеты были утеряны, но продолжает подтверждать прием всех последовательных пакетов, вплоть до утерянного.

Отправитель замечает отсутствие подтверждения о получении потерянных пакетов и полагает, что пакеты утеряны вследствие перегрузки. Повторно передать утерянные пакеты и продолжить работу при полной скорости передачи было бы неразумно, это вызвало бы лишь усиление перегрузки. Хотя нет гарантии, что все пакеты соединения TCP проходят по одинаковому пути сквозь сеть, это предположение справедливо для большинства пакетов. Для смягчения перегрузки TCP сильно снижает скорость передачи данных. Все остальные соединения TCP при такой перегрузке поступают точно так же, и вскоре перегрузка исчезает. Такое взаимодействие соединений TCP в Internet является одной из главных причин, по которой Internet все еще работает. Применение UDP — это не выход, поскольку только вначале его пропускная способность выше, чем при соединении TCP. Как только UDP начинают использовать все, перегрузка становится ординарным событием, и передача данных происходит медленнее, чем предвиделось. Даже при очень сильной загрузке TCP гарантирует, по крайней мере, совместное использование полосы пропускания.

10.1.2. Плавный старт

Реакция TCP на отсутствие подтверждения выглядит довольно резкой, но является необходимой для достаточно быстрого устранения перегрузки. Поведение TCP после обнаружения перегрузки называется **плавным стартом** (slow start).

Отправитель всегда рассчитывает **окно перегрузки** для приемника. Начальный размер окна перегрузки равен одному сегменту (пакет TCP). Теперь отправитель посылает один пакет и ждет подтверждения. Если это подтверждение приходит, отправитель увеличивает окно перегрузки еще на один сегмент и теперь отсылает два пакета (окно перегрузки = 2). После прибытия двух соответствующих подтверждений отправитель снова увеличивает окно перегрузки на 2, по единице на каждое подтверждение. Теперь окно перегрузки равно 4. В этой схеме окно перегрузки удваивается после каждого прибытия подтверждений, что занимает один период кругового обращения пакета (round trip time — RTT). Это называют степенным увеличением окна перегрузки в механизме плавного старта.

Постоянное удвоение окна перегрузки является очень опасным, так как данный шаг может быть слишком большим. Следовательно, степенное увеличение прекращается при достижении **порога перегрузки**. Дальнейшее увеличение скорости передачи осуществляется линейно, прибавлением к окну перегрузки 1 после каждого возвращения подтверждений.

Линейное увеличение продолжается, пока отправитель не обнаружит истечение времени ожидания отсутствующего подтверждения или промежуток в передаваемых данных из-за непрерывных подтверждений для одного и того же пакета. В любом случае отправитель устанав-

ливает порог перегрузки равным половине текущего окна перегрузки. Само окно перегрузки устанавливается на один сегмент, и отправитель начинает передачу с одного сегмента. Степенное увеличение, описываемое выше, начинается еще раз, до момента достижения нового порога перегрузки, и далее окно увеличивается линейно.

10.1.3. Быстрая повторная передача/быстрое восстановление

Как упоминалось ранее, существуют две причины снижения порога перегрузки. Одной из них является возникновение ситуации, когда отправитель непрерывно принимает подтверждения о приеме одного и того же пакета. Это говорит отправителю о двух вещах. Первое, приемник получил все пакеты, вплоть до последнего подтвержденного пакета последовательности. В TCP приемник отправляет подтверждения лишь в том случае, если он получает какой-либо пакет от отправителя. Следовательно, получение подтверждений от приемника дополнительно показывает, что получатель что-то получает от отправителя. Значит, промежуток в потоке пакетов возник не из-за резкой перегрузки, а просто пакет был утерян вследствие ошибок передачи. Теперь отправитель может повторить передачу утраченного пакета (пакетов) до истечения времени ожидания. Такой режим называется **быстрой повторной передачей**.

Кроме того, получение подтверждений показывает, что перегрузка отсутствует, а значит, плавный старт оправдан. Отправитель может продолжать работу с имеющимся окном перегрузки. В таком случае отправитель производит **быстрое восстановление** утерянных пакетов. Этот механизм может сильно повысить эффективность TCP.

Другой причиной инициации плавного старта, упоминаемого в предыдущем разделе, является истечение времени ожидания отсутствующих подтверждений. Это единственная ситуации, которую TCP, использующий быструю повторную передачу или быстрое восстановление, интерпретирует как перегрузку в сети и включает механизм плавного старта.

10.1.4. Влияние мобильности

Хотя плавный старт является одним из наиболее полезных механизмов в стационарных сетях, он чрезмерно снижает эффективность TCP при наличии мобильных отправителей или получателей. Причина: принятие решения об использовании плавного старта при неверных предпосылках. При отсутствии подтверждения схема плавного старта делает заключение о перегрузке, хотя это также может произойти в сети с мобильными и беспроводными конечными системами, где причиной не обязательно будет потеря пакета.

Частота появления ошибок в беспроводных каналах выше, чем в стационарных каналах на основе оптоволоконных или медных кабелей. Таким образом, потеря пакета является значительно более распространенным явлением и ее не всегда можно компенсировать повторными передачами на втором уровне. Если попытка повторной передачи на втором уровне слишком затягивается, то она может, например, инициировать повторную передачу TCP. Теперь второй уровень сталкивается с проблемой передачи одного и того же пакета дважды по возможно ненадежной линии связи. Определение дубликата на втором уровне — это не выход, поскольку все большее и большее количество соединений используют сквозное шифрование, исключающее возможность просмотра содержимого пакета.

Кроме того, потеря пакетов может вызываться собственно мобильностью. Существует множество ситуаций, когда мягкое переключение между двумя точками доступа невозможно для мобильных конечных систем. При использовании, например, протокола Mobile IP при переходе мобильного узла к новому чужому агенту в процессе передачи может оставаться некоторое количество пакетов, предназначенных старому чужому агенту. Возможно, старый чужой агент не сможет переправить эти пакеты новому или даже сохранить их в буфере, если разрыв соединения длится слишком долго. Такая потеря пакетов не связана с беспроводным доступом, а вызывается проблемами изменения маршрута потока данных.

Механизм TCP обнаруживает отсутствие подтверждений посредством истечения времени ожидания и решает, что потеря пакетов происходит из-за перегрузки, не различая возможные ситуации. Реакцией TCP является плавный старт, который не помогает при появлении ошибок передачи в беспроводном канале и который совсем не работает при переключении. Такое поведение приведет к резкому падению производительности, если на беспроводных линиях или мобильных узлах будет использоваться обычный TCP.

Однако, не стоит целиком изменять TCP лишь для поддержки мобильных пользователей или беспроводных каналов. Все аргументы, которые приводились при рассмотрении IP, сохраняют свою силу и в случае TCP. Установленная компьютерная база, использующая TCP, уже очень велика, чтобы ее менять, и, что более важно, механизм плавного старта поддерживает функционирование Internet. Таким образом, любое дополнение к TCP должно быть совместимым со стандартным TCP и не должно подвергать риску предупредительное поведение TCP при перегрузке.

10.2. Косвенный TCP

Существуют две причины, приведшие к разработке косвенного TCP (indirect TCP — I-TCP) [Bakre, 1995]. Одна — неэффективность TCP на беспроводных линиях связи, а другая — невозможность изменения TCP в стационарных сетях. Таким образом, I-TCP разделяет соединение TCP на две части — стационарную и беспроводную. На рис. 10.1 показан пример подключения мобильного хоста через беспроводный канал и точку доступа к "проводной" Internet, где находится корреспондентский хост. Корреспондентский узел также может использовать беспроводный доступ. Тогда все последующее справедливо и для линии доступа корреспондентского хоста.

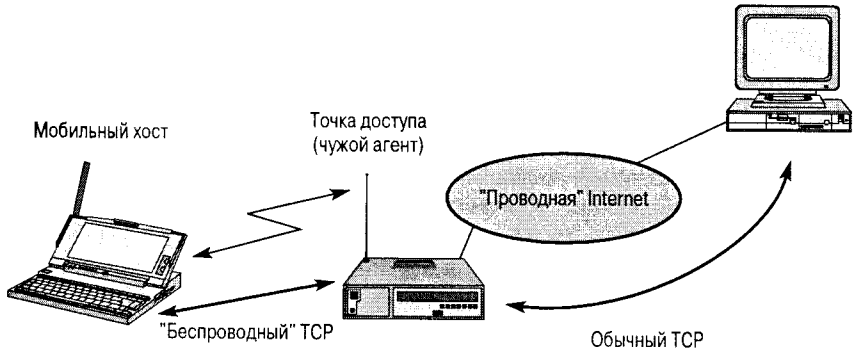

Рис. 10.1. Разделение соединения TCP на две части посредством косвенного TCP

Между стационарным компьютером и точкой доступа используется стандартный протокол TCP. Таким образом, ни один из компьютеров в Internet не видит никаких изменений относительного стандартного TCP. Вместо мобильного хоста конечной точкой соединения TCP теперь будет точка доступа, действующая как агент. Это означает, что для стационарного хоста точка доступа будет выглядеть как мобильный хост, а для мобильного хоста — как стационарный хост. Между точкой доступа и мобильным хостом используется специальный TCP, адаптированный к беспроводным линиям связи. Однако, изменение TCP для применения в беспроводных линиях связи необязательно. Даже обычный TCP может выиграть за счет меньшего периода кругового обращения, а следовательно, повторная передача сможет начаться быстрее. Удачным местом для сегментации соединения между мобильным и корреспондентским хостами является чужой агент протокола Mobile IP. Во всяком случае чужой агент контролирует передвижения мобильного хоста и может также переключать соединение на следующий чужой агент при перемещении мобильного хоста. В то же время можно также представить себе разделение соединения TCP на специальном сервере, например, в точке входа в мобильную телефонную сеть.

Корреспондентский узел в стационарной сети не замечает беспроводный канал связи или сегментацию соединения. Чужой агент действует как посредник и переправляет все данные в обоих направлениях. Если корреспондентский хост отсылает пакет, чужой агент подтверждает этот пакет. Далее чужой агент пытается переправить пакет мобильному хосту. Если мобильный хост получает пакет, он подтверждает получение пакета. Однако, это подтверждение используется лишь чужим агентом. Если вследствие ошибки передачи пакет теряется на беспроводной линии, корреспондентский хост не заметит этого. Тем не менее, чужой агент попытается повторить передачу пакета локально, чтобы поддержать надежность связи.

Подобным образом, если пакет отсылает мобильный хост, чужой агент подтверждает получение пакета и пробует переправить его на корреспондентский хост. Если пакет теряется на беспроводной линии, мобильный хост может узнать об этом значительно быстрее (из-за меньшего периода кругового обращения пакета) и может непосредственно повторить передачу пакета. Потерей пакета в проводной сети теперь занимается чужой агент.

При переключении I-TCP должен выполнить несколько действий. Как показано на рис. 10.2, с помощью, например, протокола Mobile IP нужно перенаправить не только пакеты. В этом примере показано, что точка доступа действует как посредник, обеспечивая буферизацию пакетов для повторной передачи. После переключения старый агент должен переслать буферизованные данные новому, поскольку он уже подтвердил их получение. Как объяснялось в главе 9, после регистрации с новым чужим агентом этот новый чужой агент может сообщить старому свое расположение, что позволит осуществить переправку пакетов. Помимо содержимого буфера к новому чужому агенту, расположенному в точке доступа, тоже должны перейти сокеты посредника. Сокеты отражают текущее состояние соединения TCP, т.е. порядковый номер, адреса, порты и прочее. Мобильный хост не может создавать никаких новых соединений, а корреспондентский хост не должен увидеть никаких изменений в состоянии соединения.

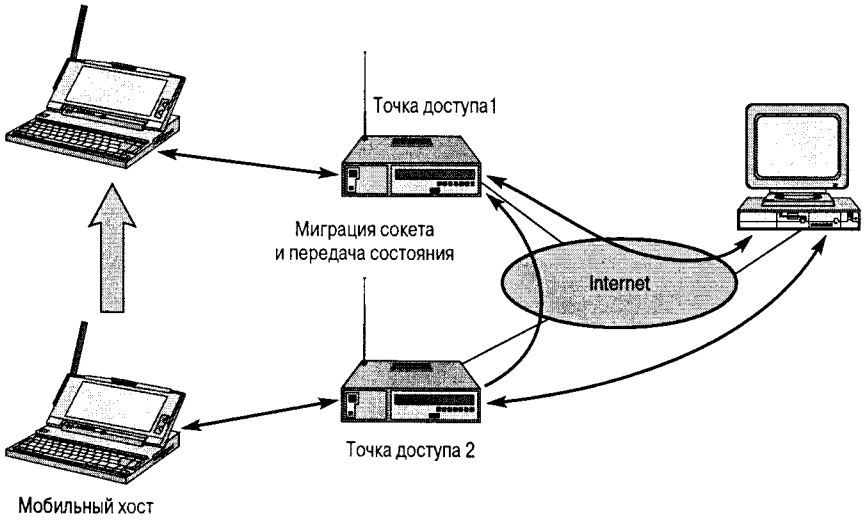

Рис. 10.2. Сокет и состояние миграции после переключения мобильного хоста

I-TCP дает несколько **преимуществ**.

- I-TCP не требует никаких изменений в протоколе TCP, который применяется хостами стационарной сети или другими хостами беспроводной сети, где не проводилась описанная оптимизация. Следовательно, все текущие модификации TCP будут работать лишь между чужим агентом и корреспондентским хостом.

- Из-за резкого разделения на два соединения ошибки передачи беспроводных каналов, т.е. потерянные пакеты, не могут распространяться в стационарную сеть. Без разделения повторная передача утраченных пакетов должна была бы осуществляться между мобильным хостом и корреспондентским через всю сеть. А в данной ситуации от чужого агента передаются лишь последовательности пакетов без пропусков.
- Всегда опасно вводить новые механизмы в большую сеть, такую как Internet, без четкого понимания последствий. Однако, для повышения производительности TCP требуются новые механизмы (например, отключение плавного старта в некоторых случаях), но при I-TCP лишь между мобильным хостом и чужим агентом. Можно проверить разные решения или использовать их одновременно, не подвергая риску дестабилизации Internet. Кроме того, оптимизация новых механизмов выглядит проще, так как они охватывают лишь один маршрутный участок.
- Между мобильным хостом и чужим агентом существует короткая задержка, которую можно определить и которая не зависит от других информационных потоков. Следовательно, оптимизированный TCP может использовать точное время ожидания для максимально быстрой гарантированной повторной передачи. Даже стандартный TCP выигрывает от меньшего времени кругового обращения, что дает возможность проводить быстрое восстановление после потери пакетов.
- Разделение на два соединения также позволяет использовать разные протоколы транспортного уровня между чужим агентом и мобильным хостом или применить сжатие заголовков и т.д. Чужой агент теперь может действовать как шлюз для преобразования между разными протоколами.

При всем сказанном, идея сегментации в I-TCP имеет и определенные **недостатки**.

- Потеря сквозной семантики TCP может вызвать проблемы при сбоях чужого агента, разделяющего соединение TCP. Если отправитель принимает подтверждение, то считается, что приемник получил пакет. Получение подтверждения теперь означает (для мобильного и корреспондентского хостов) лишь то, что чужой агент получил пакет. Корреспондентский узел ничего не знает о разделении, а значит, сбой узла доступа также может привести к срыву приложения, запущенного на корреспондентском узле и предполагающего надежную сквозную доставку.
- В прикладных задачах увеличение задержки переключения может быть более проблематичным. Все пакеты, посланные корреспондентским хостом, буферизуются чужим агентом перед пересылкой на мобильный хост (если соединение TCP разделяется именно чужим агентом). Чужой агент убирает пакеты из буфера только после получения соответствующего подтверждения. Если теперь мобильный хост переключается на другой чужой агент, он делает это до того, как старый чужой агент сможет переправить буферизованные данные новому чужому агенту. За это время могут поступить еще пакеты. Все эти пакеты нужно переправить новому чужому агенту до того, как он сможет начать их передачу.
- Чужой агент должен быть высоконадежным объектом, поскольку соединение TCP заканчивается в этой точке. Если пользователи применяют сквозное шифрование, например, согласно стандарту RFC 1827 [Atkinson, 1995], чужой агент нужно интегрировать во все механизмы обеспечения безопасности.

10.3. Подслушивающий TCP

Одним из недостатков I-TCP является сегментация единого соединения TCP на два соединения, вследствие чего теряется исходная сквозная семантика TCP. Та модификация TCP, о ко-

торой пойдет речь ниже, является полностью независимой, причем она не затрагивает сквозное соединение TCP. Основная идея этой модернизации состоит в буферизации данных вблизи от мобильного хоста с целью осуществления быстрой повторной передачи при потере пакета. Как и ранее, удачным местом введения улучшения в контексте протокола Mobile IP может быть чужой агент (см. рис. 10.3).

Рис. 10.3. Подслушивающий TCP как независимое расширение TCP

При предлагаемом подходе чужой агент буферизует все пакеты, **адресованные мобильному хосту**, и вдобавок "прослушивает" все потоки пакетов, которые движутся в обоих направлениях, в поисках подтверждений [Balakrishnan, 1995], [Brewer, 1998]. Целью буферизации всех пакетов, направляющихся на мобильный узел, является возможность осуществления быстрой повторной передачи в случае обнаружения утери пакета в беспроводном канале связи. Чужой агент буферизует каждый пакет, пока не получит подтверждение от мобильного хоста. Если чужой агент не получает подтверждения от мобильного хоста в течение определенного времени, это означает, что пакет или подтверждение были утеряны. В качестве альтернативного варианта чужой агент может принять дублирующееся подтверждение ACK, что также указывает на утерю пакета. Теперь чужой агент переправляет пакет непосредственно из буфера, осуществляя более быструю повторную передачу по сравнению с корреспондентским хостом. Время ожидания подтверждения можно сделать очень коротким, поскольку оно определяется только задержкой одного маршрутного участка и временем обработки сигнала.

Чтобы чужой агент был независимым, он не должен подтверждать данные корреспондентскому хосту. Если это сделать, корреспондентский хост будет считать, что мобильный хост получил данные, что уничтожит сквозную семантику при сбое чужого агента. Однако, чужой агент может фильтровать копии подтверждений, чтобы избежать ненужных повторных передач от корреспондентского хоста. Теперь при сбое чужого агента повторная передача будет инициирована по истечении времени ожидания корреспондентского хоста. Кроме того, чужой агент может отказаться от дублирования пакетов, которые уже были переданы локально и подтверждены мобильным хостом. Это позволяет избежать передачи ненужной информации по беспроводному каналу.

Передача данных от мобильного хоста **корреспондентскому хосту** происходит следующим образом. Чужой агент прослушивает поток пакетов с целью определения промежутка в порядковых номерах TCP. Как только чужой агент обнаруживает отсутствие пакета, он отправляет отрицательное подтверждение (NACK) мобильному хосту. Тогда мобильный хост может повторить передачу утраченного пакета. Упорядочение пакетов происходит автоматически с использованием TCP корреспондентского хоста. Расширение функций чужого агента до "прослушивания" TCP имеет несколько **положительных сторон**.

- Основным преимуществом этой модели является сохранение сквозной семантики TCP. Не имеет значения, когда чужой агент претерпит сбой (если именно он является местом буферизации и прослушивания), и мобильный и корреспондентский хосты не будут использо-

вать несовместимые протоколы TCP, что возможно в случае I-TCP. Модель автоматически вернется к стандартному TCP, если расширенная схема прекратит работу.

- Исчезает необходимость изменения корреспондентского хоста, большая часть модификаций затрагивает чужой агент. Поддержка потока пакетов лишь между корреспондентским и мобильным хостом даже не требует изменений на мобильном хосте.
- Плюсом этой модели является отсутствие необходимости в переключении при перемещении мобильного хост к другому чужому агенту. Все, что при этом происходит, — это превышение времени ожидания на корреспондентском хосте и повторные передачи пакетов, возможно, на новый адрес передачи.
- Кроме того, не имеет значения, использует следующий чужой агент модифицированный протокол TCP или нет. Если нет, то модель автоматически возвращается к стандартной схеме. Это одна из проблем I-TCP — старый чужой агент уже может успеть передать на корреспондентский хост подтверждение о корректном приеме данных и теперь должен начать переправку пакетов на мобильный хост через новый чужой агент.

Тем не менее, простота такой схемы приводит и к своим **недостаткам**.

- Подслушивающий TCP не обеспечивает такого хорошего отделения поведения беспроводной линии связи, как I-TCP. Например, представим, что (в связи с проблемами в беспроводном канале связи) на успешную повторную передачу чужим агентом пакетов из буфера уходит некоторое время. Хотя после истечения времени ожидания корреспондентский хост активирует повторную передачу, время ожидания для чужого агента может оказаться меньше, чем для корреспондентского хоста. Следовательно, проблемы беспроводного канала теперь видимы корреспондентскому хосту и не являются полностью изолированными. Качество обособления, которое обеспечивает подслушивающий TCP, сильно зависит от качества беспроводного канала связи, времени ожидания и прочих характеристик информационного потока.
- Использование отрицательных подтверждений между мобильным хостом и чужим агентом требует дополнительных механизмов мобильного хоста. Таким образом, эта модель уже является зависимой от мобильных хостов.
- И наконец, все попытки прослушивания и буферизации данных могут быть тщетными, если применяются некоторые сквозные схемы шифрования между мобильным и корреспондентским хостами. Используя инкапсуляцию информации, относящейся к безопасности [RFC 1827, Atkinson, 1995], можно зашифровать заголовок протокола TCP, после чего прослушивание порядковых номеров больше ничего не даст. Более того, повторная передача данных от чужого агента может оказаться невозможной, поскольку многие схемы безопасности предотвращают атаки воспроизведения, а повторная передача данных от чужого агента может ложно интерпретироваться как атака воспроизведения[1]. Сквозное шифрование используется многими приложениями, а значит не ясно, как в будущем можно будет использовать эту схему.

10.4. Мобильный TCP

Отбрасывание пакетов при переключении или вследствие высокой частоты появления ошибок — это не единственное следствие мобильности и использования беспроводных каналов связи. Кроме этого имеются еще длительные и/или частые разрывы соединений. Довольно часто мобильный узел вообще не может установить соединение. Примеры: изолированные беспроводные ЛВС в отдельных зданиях, не покрывающие территорию всего учебного заведения. Что происходит со стандартным TCP при разъединении?

[1] В этом контексте чужой агент также называется "подслушивающий посредник"

Отправитель TCP пытается повторить передачу данных под управлением таймера повторных передач, который удваивается после каждой неудачной попытки повторной передачи, пока не достигнет максимального значения — одной минуты. Это означает, что отправитель каждую минуту пробует повторить передачу неподтвержденного пакета и прекращает это делать после 12 попыток. Что произойдет, если соединение восстановится раньше? Даже в этом случае в течение одной минуты не будет произведено успешной передачи данных! Время ожидания еще не истекло, и отправитель должен ждать. Более того, отправитель инициирует плавный старт, поскольку предполагает перегрузку.

А что происходит, если связь разрывается у мобильного узла, использующего I-TCP? Посредник должен буферизовать все больше и больше данных, а значит, чем больше период разъединения, тем большего объема потребует буфер. К тому же, если после разъединения следует переключение (что чаще всего и происходит), то новому посреднику нужно отослать больше состояний. При использовании метода прослушивания разъединение также доставляет неприятности. Мобильный узел не сможет отослать подтверждение ACK, а значит, прослушивание ничего не даст.

Модель мобильного TCP (Mobile TCP — M-TCP)[2] имеет те же цели, что и I-TCP и прослушивающий TCP: предотвратить уменьшение окна отправителя при появлении ошибок или разъединения, но не при перегрузке. M-TCP настроен на повышение общей производительности, снижение задержек, поддержку сквозной семантики TCP и осуществление более эффективного переключения. Кроме того, M-TCP специально адаптирован к сложностям, возникающим при длительных или частых разъединениях [Brown, 1997].

В M-TCP соединение TCP точно так же разделяется на две части, как это делается в I-TCP. TCP без каких-либо модификаций применяется на участке между обычным хостами и **координирующим хостом (supervisory host — SH)**, а оптимизированный TCP используется в соединении между координирующим и мобильным хостом (mobile host — MH). Координирующий хост отвечает за обмен данными между обоими участками соединения подобно посреднику в M-TCP (см. рис. 10.1). Модель M-TCP предполагает относительно низкую частоту появления ошибок в беспроводном канале. Следовательно, в ней не выполняется кэширование/повторная передача данных посредством хоста SH. Если пакет теряется на беспроводном участке, его повторную передачу должен осуществлять исходный отправитель. Это поддерживает сквозную семантику TCP.

Хост SH отслеживает все пакеты, отсылаемые на мобильный хост, и возвращаемые подтверждения ACK. Если SH не получает подтверждения в течение некоторого времени, то он считает, что мобильный хост отсоединился. Тогда он перекрывает канал отправителя, выставляя размер окна отправителя равным 0. Это вынуждает отправителя перейти в **устойчивый режим**, т.е. состояние отправителя не изменится независимо от того, как долго получатель будет разъединен. Это означает, что отправитель не будет пытаться повторять передачу данных. Как только хост SH (старый или новый) определит, что соединение восстановлено, он снова восстановит окно отправителя до прежнего размера. Таким образом, отправитель сможет продолжить передачу на полной скорости. Этот механизм не требует изменений TCP отправителя.

С беспроводной стороны используется адаптированный TCP, который может быстрее восстанавливаться после потерь пакетов. Такой модифицированный TCP не использует плавного старта, а значит, M-TCP нужна **служба контроля пропускной способности** для реализации беспристрастного совместного использования беспроводного канала связи.

Достоинства M-TCP.

- M-TCP поддерживает сквозную семантику TCP. SH сам не отсылает никаких подтверждений ACK, вместо этого он переправляет ACK от мобильного хоста.

---

[2] Читатель должен понимать, что Mobile TCP не имеет того же статуса, что Mobile IP, являющийся стандартом (RFC) Internet

- Если мобильный хост отсоединяется, M-TCP избегает ненужных повторных передач, плавных стартов или обрывов соединений, просто выставляя нулевой размер окна отправителя.
- Поскольку в M-TCP данные не буферизуются в SH, как это делается в I-TCP, отпадает необходимость пересылки буферизованных данных новому SH. Потерянные пакеты будут автоматически повторно переданы новому SH.

Отсутствие буферов и изменение TCP в беспроводной части канала связи имеет также свои **недостатки**.

- Поскольку SH не действует как посредник I-TCP, потери пакетов в беспроводном канале, возникающие вследствие битовых ошибок, распространяются к отправителю. В M-TCP предполагается, что частота появления битовых ошибок мала, а такое допущение справедливо не всегда.
- Модифицированный TCP в беспроводном канале связи требует не только изменения в программном обеспечении мобильного хоста, но и новых элементов в сети, таких как служба контроля пропускной способности.

10.5. Быстрая повторная передача/быстрое восстановление

Как было сказано в разделе 10.1.4, перемещение к новому чужому агенту может приводить к потере пакетов или истечению времени ожидания мобильного или корреспондентского хоста. TCP воспринимает это как перегрузку и инициирует плавный старт, хотя на самом деле никакой перегрузки нет. В разделе 10.1.3 показаны механизмы быстрого восстановления/быстрой повторной передачи, которые может использовать хост после получения дублирующегося подтверждения, приходя, таким образом, к заключению, что пакет был утрачен не из-за перегрузки.

Идея, предложенная в работе [Caceres, 1995], заключается в искусственном вызове быстрой повторной передачи со стороны мобильного и корреспондентского хостов. Как только мобильный хост регистрируется у нового чужого агента, используя Mobile IP, он сразу начинает отсылать дублирующиеся подтверждения на корреспондентские хосты. Предлагается отсылать три копии. Это вынуждает корреспондентский хост перейти в режим быстрой повторной передачи, а не использовать плавный старт, т.е. корреспондентский хост продолжает передачу с той же скоростью, что и ранее, до перемещения мобильного хоста к другому чужому агенту.

Поскольку мобильный хост также может инициировать плавный старт после перемещения к новому чужому агенту, предлагаемый подход дополнительно переводит мобильный хост в режим быстрой повторной передачи. Таким образом, мобильный хост повторно передает все подтвержденные пакеты, используя текущий размер окна перегрузки без перехода к плавному старту.

Преимущество этой модели заключается в ее простоте. Повышение производительности достигается лишь минимальными изменениями программного обеспечения мобильного хоста. Никакого вмешательства в чужой агент и корреспондентский хост не требуется.

Главный же **недостаток** этой схемы заключается в недостаточной изоляции потерь пакетов. Вынужденная быстрая повторная передача повышает эффективность, но передаваемые пакеты должны пересечь всю сеть между корреспондентским и мобильным хостами. Кроме того, если переключение от одного чужого агента к другому занимает много времени, корреспондентский хост успевает начать повторную передачу. Описанный подход в основном сосредоточен на потерях вследствие переключения, поэтому потери вследствие проблем беспроводного канала не рассмотрены. И наконец, эта модель требует большей кооперации между уровнями Mobile IP и TCP, что усложняет задачу изменения одного протокола, не затронув другой.

10.6. Замораживание передачи/времени ожидания

Представленные выше подходы могут успешно использоваться при коротких разрывах соединений, вызванных переключением или ошибками передачи в беспроводном канале связи, но кроме них также существуют модели, разработанные для длительных перерывов связи. Примеры: мобильный хост в автомобиле, въезжающем в туннель, из-за чего происходит разрыв соединения, например, со спутником (впрочем, многие тоннели, а также метрополитены позволяют поддерживать связь по мобильному телефону) или пользователь, который переходит в ячейку, у которой превышена пропускная способность. В этом случае система мобильного телефона прерывает связь. Реакцией со стороны TCP (даже с описанными выше расширениями) будет отсоединение после истечения времени ожидания.

Достаточно часто на уровне MAC проблемы с соединением отмечаются еще до того, как соединение действительно прервется с точки зрения TCP. К тому же на уровне MAC известна настоящая причина разрыва и предполагается перегрузка, как это происходит на уровне TCP. Следовательно, уровень MAC может уведомлять уровень TCP о приближающейся потере соединения или текущем разрыве, вызванном не перегрузкой. TCP не может остановить отправку и "замораживает" текущее состояние с его окном перегрузки и даже таймеры ожидания. Если на уровне MAC наступление разъединения замечается достаточно рано, то мобильный и корреспондентский хосты будут предупреждены. Если в беспроводном канале связи происходит быстрое прерывание, в точке доступа нужен дополнительный механизм уведомления корреспондентского хоста о причине прерывания. В противном случае корреспондентский хост инициирует плавный старт, считая, что произошла перегрузка, и в конце концов оборвет соединение.

Как только на уровне MAC снова фиксируется наличие соединения, на уровень TCP поступает сообщение, что можно продолжить работу с того же места, где она была вынужденно остановлена. Для TCP время просто стоит на месте, а значит, время таймеров не истекает.

Преимущество этой модели в том, что она позволяет продолжать соединение TCP даже после длительных разрывов. Более того, она не зависит от прочих механизмов TCP, таких как подтверждение или порядковая нумерация, а значит, ее можно использовать совместно с шифрованными данными. Однако, у такой схемы имеются и свои **недостатки**. Потребуются не только изменения программного обеспечения мобильного хоста. Для большей эффективности должен меняться корреспондентский хост. Все механизмы зависят от способности уровня MAC к обнаружению будущих прерываний. И наконец, замораживание состояния TCP не сможет помочь при шифровании с использованием зависящих от времени случайных чисел. Эта схема требует повторной синхронизации после прерывания.

10.7. Выборочная повторная передача

Очень удобным расширением TCP является выборочная повторная передача. Подтверждения в TCP являются куммулятивными, т.е. они подтверждают последовательный прием пакетов вплоть до некоторого пакета. Если теряется один пакет, отправителю приходится повторять всю передачу, начиная с утерянного пакета (возвратная повторная передача). Очевидно, что при этом неэффективно используется пропускная способность, причем это справедливо не только в мобильном случае, но и в любой сети.

Используя стандарт RFC 2018 [Mathis, 1996], TCP может косвенно запрашивать выборочную повторную передачу пакетов. Получатель может подтверждать отдельные пакеты, а не только последовательности упорядоченных пакетов. Теперь отправитель может точно определять, какой пакет нужен, и повторно отсылать его.

Преимущество этой схемы очевидно: отправитель повторяет передачу лишь утерянных пакетов. Это снижает требования к пропускной способности, особенно в случае медленных беспроводных линий связи. Выигрыш в эффективности достигается не только в беспроводных линиях связи и мобильных средах. Использовать выборочную повторную передачу выгодно и в других сетях. Однако, есть незначительный **недостаток**, который сводится к усложнению программного обеспечения со стороны получателя, так как потребуется буфер большего размера для повторного упорядочения данных и ожидания заполнения промежутков. Здесь стоит отметить, что объемы памяти и производительность процессоров неуклонно растут, а полоса пропускания радиоинтерфейса остается практически неизменной. Следовательно, высокая сложность — это уже не такой большой недостаток, каким она была в ранний период существования TCP.

10.8. TCP, ориентированный на транзакции

Представим приложение, запущенное на мобильном хосте, время от времени посылающем на сервер короткий запрос, на который отсылается короткий же ответ. Если приложение требует надежной доставки пакетов, оно может использовать TCP (многие приложения такого рода используют UDP и решают вопрос надежности на более высоком уровне, ориентированном на конкретное приложение).

Использование TCP теперь требует передачи нескольких пакетов по беспроводному каналу связи. Прежде всего TCP использует трехстороннее квитирование для установления соединения. По крайней мере, один пакет нужен для передачи запроса, и в завершение TCP нужно еще три пакета для закрытия соединения посредством трехстороннего квитирования. Если считать, что соединение происходит в условиях интенсивного информационного обмена или продолжается очень долго, эти служебные издержки минимальны. В случае же одного пакета данных TCP потребуются все семь пакетов.

Данные служебные издержки стали причиной разработки TCP, ориентированного на транзакции (T-TCP, RFC 1644 [Braden, 1994]). T-TCP может комбинировать пакеты для установления и разрыва соединения с пакетами пользовательских данных, что позволяет снизить число пакетов до двух вместо семи.

Очевидной **выгодой** для некоторых прикладных систем является уменьшение служебных издержек, которые в стандартном TCP нужны для установления и разрыва соединения. Однако, T-TCP — это не исходный TCP, так что ему нужны изменения в мобильном и корреспондентском хостах, что, безусловно, является серьезным **недостатком**. Такое решение более не скрывает мобильности.

10.9. Резюме

В этой главе обрисованы проблемы TCP как протокола, ориентированного на соединение в мобильной среде. Основные предположения, сделанные при разработке TCP, абсолютно не согласуются с реальностью использования мобильных хостов. В частности, механизмы, делающие TCP эффективным в сети, и следовательно поддерживают целостность Internet, приводят к определенным проблемам эффективности. При отсутствии своевременного подтверждения в TCP считается, что возникла перегрузка.

Однако, в беспроводных линиях связи частота появления ошибок выше по сравнению, например, с витой парой или волоконно-оптическим кабелем, из-за чего частота утраты пакетов может оказаться более значительной. Кроме того, собственно мобильность, т.е. переключение между различными точками доступа может привести к потере пакетов без перегрузки в сети. В любом случае TCP переходит в режим плавного старта, сильно снижая скорость передачи.

Для повышения эффективности TCP в мобильной среде были предложены несколько решений. В табл. 10.1 дается обзор таких механизмов, а также перечисляются их преимущества и недостатки. Модели не являются взаимоисключающими и допускают комбинирование. Например, выборочную повторную передачу можно объединить с другими методами и даже использовать в стационарных сетях.

Таблица 10.1. Обзор нескольких улучшений TCP, поддерживающих мобильность

Подход	Механизм	Преимущества	Недостатки
Косвенный TCP	Разделение соединения TCP на два соединения	Изоляция беспроводной линии связи, простота	Утрата сквозной семантики TCP, высокие задержки при переключении
Подслушивающий TCP	"Подслушивание" данных и подтверждений, локальные повторные передачи	Незаметен для сквозного соединения, возможна интеграция с MAC	Проблематичен при использовании шифрования, недостаточная изоляция беспроводного канала связи
M-TCP	Разделение соединения TCP, подавление отправителя через размер окна	Поддержка сквозной семантики, эффективная работа при длительных и частых разъединениях	Плохая изоляция беспроводного канала связи, служебные издержки на обработку вследствие управления пропускной способностью
Быстрая повторная передача/быстрое восстановление	Избегает плавного старта после роуминга	Простота и эффективность	Смешение уровней, отсутствие независимости
Замораживание передачи/времени ожидания	При разрыве соединение TCP замораживается и продолжается после восстановления соединения	Независимое содержание, работа при длительных разрывах соединения	Нужны изменения в TCP, зависит от MAC
Выборочная повторная передача	Повторно передаются лишь утраченные данные	Очень эффективен	Незначительное усложнение программного обеспечения получателя, нужен буфер большего размера
TCP, ориентированный на транзакции	Совмещается установление/разрыв соединения и передача данных	Эффективен в некоторых приложениях	Требует изменений в TCP, отсутствует независимость

Неизмененный TCP порождает даже больше проблем при использовании на спутниковых линиях связи или в общих каналах связи с космическим аппаратом (от аппаратов LEO до зондов дальней космической разведки). Главной проблемой является очень большое время на передачу и подтверждение приема, подверженность ошибкам линий связи, ограниченная пропускная способность, прерывистая связь и асимметричные каналы (отношение потоков до 1000:1). Асимметричные каналы, например, с широкой полосой пропускания от космического аппарата до наземной станции, ограничивают пропускную способность из-за ограниченной пропускной способности, выделенной

подтверждениям в обратном направлении. Был представлен комплекс модификаций TCP [Durst, 1997], в первую очередь возможность **выборочного отрицательного подтверждения (selective negative acknowledgement — SNACK)**, которая адаптирует TCP к требованиям космической связи. Набор протоколов, разработанных для космической связи, известен под названием **стандартов протоколов космической связи (space communication protocol standards — SCPS)**, а расширение TCP называется **транспортным протоколом SCPS (SCPS-TP)**.

Поскольку в беспроводных линиях связи пропускная способность, как правило, весьма ограничена, то низкая доля служебных издержек является очень важной. Использование совместно с TCP схемы туннелирования, как в протоколе Mobile IP (см. раздел 9.1), даст в итоге заголовки протокола размером в 60 байт в случае IPv4 и 100 байт в случае IPv6, что связано с большим адресным полем. В работе [Degermark, 1997] была предложена схема **сжатия заголовков**, которая особенно удачно подходит к каналам связи, подверженным ошибкам, и для широковещательной связи. Приложения, особо чувствительные к задержкам, как например, интерактивные игры с большим количеством небольших пакетов, очень выигрывают от сжатия заголовков.

10.10. Упражнения

1. Сравните различные типы ошибок передачи, которые могут появиться в проводных и беспроводных линиях связи. Какие дополнительные факторы связаны с мобильностью?
2. Какова реакция стандартного TCP на потерю пакета? В какой ситуации такая реакция имеет смысл и почему очень часто она проблематична в контексте беспроводных сетей и мобильности?
3. Можно ли решить проблему использования TCP путем замены TCP на UDP? Где это будет удобно и почему это может довольно сильно угрожать стабильности сети?
4. Как и почему I-TCP изолирует проблемы беспроводного канала связи? Каков главный недостаток такого решения?
5. Покажите взаимодействие протокола Mobile IP и стандартного TCP. Опишите поток пакетов от стационарного хоста к мобильному хосту через чужой агент. Повторите то же самое для переключения. Какими будут последующие действия протокола Mobile IP и реакция TCP?
6. Покажите шаги, которые необходимы при переключении в схемах, предложенных в этой главе. Каковы состояние и функция чужого агента, базового агента, корреспондентского хоста, мобильного хоста и адреса передачи до, в течение и после переключения? Какую информацию, какому объекту необходимо переслать для поддержания согласованности соединения TCP?
7. Как шифрование влияет на предложенные схемы? Рассмотрите систему безопасности IP, которая может шифровать полезную нагрузку, например, пакет TCP.
8. Перечислите последующие оптимизации TCP относительно протокольных служебных издержек, что представляет особую важность для узкополосных соединений.

10.11. Рекомендуемая литература

Atkinson, R. (1995) *IP Encapsulation Security Payload (ESP)*, RFC, 1827.

Bakre, A.; Badrinath, B. (1995) "I-TCP: Indirect TCP for mobile host", proceedings of the 15th International Conference on Distributed Computing Systems (ICDCS), Vancouver, Canada.

Balakrishnan, H; Seshan, S; Katz, R. H. (1995) "Improving reliable transport and handoff performance in cellular wireless networks", *Wireless Networks*, J. C. Baltzer, 1.

Braden, R. (1994) *T-TCP — TCP extensions for transactions functional specification*, RFC 1644.

Brewer, E. A.; Katz, R. H.; Chawathe, Y.; Gribble, S. D.; Hodes, T.; Nguyen, G.; Stemm, M.; Henderson, T.; Amit, E.; Balakrishnan, H.; Fox, A.; Padmanabhan, V.; Seshan, S. (1998) "A network architecture for heterogeneous mobile computing", *IEEE Personal Communications*, 5, (5).

Brown, K.; Singh, S. (1997) "M-TCP: TCP for mobile cellular networks", *ACM Computer Communications Review*, 27, (5).

Caceres, R.; Iftode, L.; (1995) "Improving the performance of reliable transport protocols in mobile computing environments", *IEEE Journal on Selected Areas in Communications*, 13, (5).

Degermark, M.; Engan, M.; Nordgren, B.; Pink, S. (1997) "Low-loss TCP/IP header compression for wireless networks", *Wireless Networks*, J. C. Baltzer, no. 3.

Durst, R. C.; Miller, G. J.; Travis, E. J. (1997) "TCP extensions for space communications", *Wireless Networks*, J. C. Baltzer, no. 3.

Mathis, S.; Mahdavi, J.; Foyd, S.; Romanow, A. (1996) *TCP selective acknowledgement options*, RFC 2018.

Postel, J. (1981) *Transmission Control Protocol*, RFC 793.

Stevens, W. R. (1994) *TCP/IP Illustrated, Volume 1: The Protocols*. Addison Wesley Longman Publishing Company, Reading, Massachusetts, USA.

Поддержка мобильности 11

Одной лишь передачи данных от отправителя к одному или нескольким получателям недостаточно. Только приложения делают сеть связи полезной. Однако, чтобы использовать хорошо известные приложения стационарных сетей в системах мобильной и беспроводной связи, нужны некоторые дополнительные компоненты. В качестве примеров можно назвать файловые системы, базы данных, безопасность, механизмы учета и выписывания счетов. Поскольку мобильность пользователя должно поддерживать обычное аппаратное обеспечение, важным вопросом является потребляемая мощность.

В этой главе внимание сфокусировано на двух аспектах: файловых системах и доступе к **World Wide Web (WWW)**. Несколько лет назад темы многих научно-исследовательских проектов касались задач **распределенных файловых систем**. Некоторые были связаны с поддержкой мобильных устройств, линиями беспроводной связи с низкой пропускной способностью и работой без соединения. Основной проблемой распределенных, слабосвязанных файловых систем является поддержание согласованности. Все ли представления файловой системы одинаковы? Что случится, если разъединенный пользователь изменит данные? В разделе 11.1 рассматривается несколько таких проблем и представлены некоторые исследовательские проекты.

Впрочем, успех WWW смещает акценты многих проектов. В настоящее время многие исследовательские проекты направлены на поддержку просмотра Web-ресурсов мобильными пользователями, поскольку Web — это приложение, направляющее Internet. Итак, в разделе 11.2 объясняются некоторые основные свойства Web и представляется краткий обзор протокола передачи гипертекстовых файлов (hypertext transfer protocol — HTTP) и языка гипертекстовой разметки (hypertext markup language — HTML). В разделе демонстрируются фундаментальные проблемы HTTP и HTML при использовании в мобильной сети, доступ к которой возможен только с низкой пропускной способностью. Web разрабатывалась для обычных компьютеров и стационарных сетей. В нескольких системных архитектурах были предприняты попытки устранить такие проблемы. Эти архитектуры являются также хорошими примерами сценариев клиент/сервер в беспроводной среде.

Раздел 11.3, представляющий **протокол беспроводных приложений** (wireless application protocol — WAP), является основной частью главы. WAP — это результат общей работы многих компаний и организаций по формированию основы для беспроводного и мобильного доступа к Web с использованием множества различных транспортных систем. Примерами этих систем являются GSM, GPRS и UMTS, описанные в главе 4 WAP интегрирует в себе несколько уровней связи, предлагающих механизм безопасности, протоколы, ориентированные на транзакции, и поддержку приложений. В современной WWW эти опции не интегрированы в систему, а добавлены к ней. Кроме того, WAP сочетает телефонную сеть и Internet, интегрируя приложения телефонии в Web, используя свой собственный язык разметки для беспроводных приложений (wireless markup language — WML) и язык подготовки сценариев (WMLScript). В разделе 11.3 приведено большое количество примеров. Чита-

тель, однако, должен осознавать, что все это — свежие разработки, и в настоящее время по-прежнему развиваются новые приложения и механизмы.

11.1. Файловые системы

Основной задачей файловых систем является поддержка эффективного и согласованного доступа к файлам независимо от того, где размещены клиент, запрашивающий файлы, или сервер(ы), предлагающий их. **Эффективность** в особенности важна для беспроводных систем, поскольку для них характерна низкая производительность и, следовательно, служебная информация протоколов, операции обновления и т.д. должны быть минимальными. **Независимость от размещения** связана с проблемой зависимых от расположения представлений файловой системы. Для поддержки мобильности файловая система должна обеспечивать одинаковые представления каталогов, имен файлов, прав доступа и т.д. независимо от текущего положения. Главной же является проблема **согласованности**, которая более детально рассмотрена в разделе 11.1.1.

Общими проблемами являются ограниченность ресурсов портативных устройств и низкая пропускная способность беспроводного доступа. Файловые системы не могут полагаться на большой кэш в конечной системе или проводить много обновлений посредством беспроводной линии связи. Кроме того, портативные устройства могут быть разъединены в течение длительного периода времени. Аппаратные и программные компоненты портативных устройств часто не соответствуют стандартным компьютерным архитектурам или операционным системам соответственно. Мобильные телефоны, PDA и другие устройства имеют свои собственные операционные системы, аппаратные средства и программное обеспечение. Портативные устройства не так надежны, как настольные системы или традиционные файловые серверы.

Стандартные файловые системы, такие как сетевая файловая система (network file system — NFS), очень неэффективны, почти непригодны для использования в мобильной и беспроводной среде [Honeyman, 1995]. Традиционная файловая система не ожидает разъединения, соединения с низкой пропускной способностью и высокой латентности. Для поддержки работы при разрывах портативные устройства могут дублировать файлы или единичные объекты. Это можно делать заблаговременно путем предварительной выборки данных или выборки с упреждением (кэширования). Основной проблемой в такой ситуации является согласованность копий и исходных данных. В следующем разделе представлено еще несколько проблем, а также их решения, касающиеся согласованности.

11.1.1. Согласованность

Основной проблемой распределенных файловых систем, допускающих дублирование данных для поддержания производительности, является согласованность дублирующихся объектов (файлов, частей файлов, элементов структур данных и т.д.). Что произойдет, например, если два портативных устройства содержали копии одного и того же объекта, затем одно устройство изменило значение объекта, а после этого оба устройства прочли это значение? Без дополнительных механизмов одно портативное устройство прочтет старое значение.

Для избежания несогласованности многие традиционные системы применяют механизмы поддержания постоянного согласованного представления для всех пользователей файловой системы. Эта **сильная согласованность** достигается путем элементарных обновлений, подобно системам баз данных. Редактор объекта блокирует объект, изменяет объект и разблокирует объект после изменения. Если объект заблокирован, то никакое другое устройство не может редактировать его. Кэшированные объекты становятся недействительными после изменения. Поддержка сильной согласованности не только очень дорого обходится (с точки зрения обмена обновлениями по беспроводной линии связи), но иногда просто невозможна. Предположим, что устройство с несколькими объектами в кэше временно отсоединено. В этом случае невозможно

обновить объекты или признать их недействительными. Более того, блокирование кэшированных объектов может быть невидимым для других пользователей.

Одним из решений является запрещение доступа к разъединенным объектам. По сути, это препятствует работе любых реальных приложений, основанных на файловой системе. Соответственно, мобильные системы должны использовать модель **слабой согласованности** для файловых систем. Слабая согласованность предполагает определенные периоды несогласованности, которые должны быть допустимыми для поддержания производительности. В то же время общая файловая система должна оставаться согласованной, и, следовательно, для повторной интеграции требуются стратегии разрешения конфликтов. **Повторная интеграция** — это процесс слияния объектов различных пользователей, дающий в результате одну согласованную файловую систему. Пользователь может сохранить копию объекта, отсоединиться от сети, изменить объект и подсоединиться снова. Тогда измененный объект нужно повторно интегрировать. Может произойти **конфликт**, например, если объект был изменен двумя пользователями, работающими с двумя копиями. Во время повторной интеграции файловая система может заметить, что обе копии различны, после чего стратегия разрешения конфликтов должна определить, какую копию использовать или как действовать дальше. Система может выявлять конфликты, основываясь на отметках времени, нумерации версий, сопоставлении и т.д.

Пусть, например, несколько человек пишут статью. Каждый работает над своим разделом, используя свой портативный компьютер. Пока каждый из них остается в пределах своего раздела, повторная интеграция проста. Как только кто-то делает копию другого раздела и начинает вносить изменения, повторная интеграция становится сложной и зависящей от содержания. В примерах следующего раздела показаны различные методы разрешения возникшей ситуации. Эти решения различаются по степени кэширования и предварительной выборки (файлы, каталоги, поддеревья, разделы диска), по местоположению объекта, поддерживающего мобильность (стационарная сеть и/или мобильный компьютер), и по стратегиям разрешения конфликтов.

11.1.2. Примеры

Предшественницей многих распределенных файловых систем, которая может использоваться в мобильных системах, является файловая система AFS [Howard, 1988]. Ниже приводятся несколько примеров основных исследовательских систем.

11.1.2.1. Coda

Coda — это преемница AFS, предлагающая два различных типа дублирования: серверное дублирование и кэширование на клиентах. Разъединенные клиенты работают только с кэшем, т.е. приложения используют кэшированные дублированные файлы. На рис. 11.1 показан кэш между приложением и сервером. Coda — это очевидное расширение клиентской схемы управления кэшем.

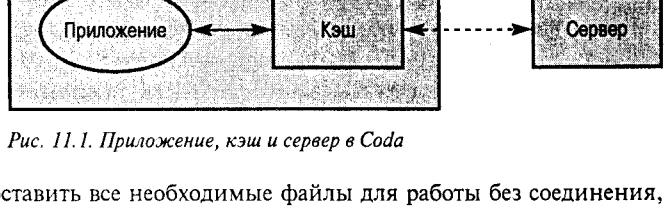

Рис. 11.1. Приложение, кэш и сервер в Coda

Чтобы предоставить все необходимые файлы для работы без соединения, Coda использует расширенные механизмы предварительной выборки файлов при наличии соединения, называемые **накоплением** [Kistler, 1992]. Если клиент связан с сервером посредством устойчивого со-

единения (см. рис. 11.2), накопление предварительно выбирает файлы, используемые в данный момент. Этот автоматический сбор данных необходим, поскольку обычный пользователь не может знать все файлы, используемые в данный момент. Хотя стандартные программы и данные приложений могут быть знакомы пользователю, он обычно ничего не знает о множестве дополнительно необходимых малых системных файлов (например, профили, библиотеки совместного доступа, драйверы, шрифты).

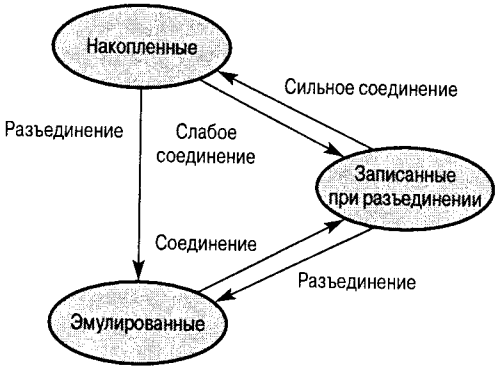

Рис. 11.2. Состояния клиента в Coda

Более того, пользователь может предопределить список файлов, для которых Coda должна сделать явную выборку. Кроме того, пользователь может установить приоритеты для определенных программ. Теперь Coda принимает решение относительно текущего содержимого кэша, используя список и стратегию наиболее позднего использования (last recently used — LRU).

Как только клиент разъединен, приложения начинают работать с копиями (см. рис. 11.2, **эмуляция**). Coda использует оптимистичный подход и, следовательно, позволяет доступ для чтения и записи всех файлов. Coda хранит запись измененных файлов, но не поддерживает историю изменения для каждого файла. Таким образом, кэш всегда имеет только одну копию (возможно, измененную). После повторного соединения Coda согласует копии с файлами на сервере, как описано в [Kistler, 1992]. Если Coda замечает, что два различных пользователя изменили файл, повторная интеграция этого файла сбоит, и Coda сохраняет измененный файл в виде копии на сервере, чтобы оставить возможность ручной повторной интеграции.

Оптимистичный подход Coda является очень крупномасштабным, поскольку она работает с целыми файлами. Успех Coda связан с тем, что в UNIX файлы редко редактируются более чем одним пользователем. Большинство файлов только читаются, и лишь некоторые файлы изменяются. Опыт работы с Coda показывает, что только 0,72 % всех обращений к файлам порождают конфликты записи [Satyanarayanan, 1993]. Если рассматривать только файлы пользователей, то эта величина снижается до 0,3 %. Однако, такой низкий уровень конфликтов не относится к произвольным совместно используемым файлам, например, применяющимся при совместной работе на базе компьютеров (CSCW). Следовательно, для автоматического разрешения конфликтов после неудачной повторной интеграции был разработан ориентированный на приложения арбитр (application specific resolver — ASR) [Kumar, 1993]. Общей проблемой этих средств является то, что они могут работать только постфактум. Это означает, что механизмы должны реконструировать историю изменений, основываясь на копиях, поскольку Coda не записывает каждое единичное изменение.

Еще одной проблемой Coda является определение конфликтов. Она обнаруживает только конфликты записи, т.е. если два или более пользователя изменяют файл. Рассмотрим два файла f_1 и f_2. Один клиент использует для вычислений чего-либо значения файлов f_1 и f_2 и сохра-

няет изменения в файле f_1. Другой клиент использует значения файлов f_1 и f_2 для вычисления чего-либо еще и сохраняет результаты в файле f_2. Coda не обнаружила бы наличия каких-либо проблем при повторной интеграции файлов. Однако, результаты могут не отражать правильных значений. Порядок выполнения играет важную роль. Для разрешения этой проблемы в Coda был введен простой механизм транзакций, так называемые только разделяющие транзакции (isolation-only transaction — IOT), [Lu, 1994]. IOT позволяет сгруппировать отдельные операции и проверить их на последовательное исполнение.

Если изначально Coda просто различала два состояния, "накопленных" при наличии связи и "эмулированных" при разъединении, позже интегрируются и слабосвязанные состояния, **записанные при разъединении** (см. рис. 11.2) [Mummert, 1995]. Если клиент только слабосвязан, Coda решает, стоит произвести выборку файла посредством такой связи или подождать, пока не будет установлено лучшее соединение. Иными словами, Coda моделирует устойчивость пользователя и взвешивает ее в смысле стоимости проведения выборки файла, который требуется пользователю.

На рис. 11.2 показано три состояния клиента в Coda. Пока существует сильная связь с сервером, клиент только выполняет накопление. Если связь прерывается совсем, клиент прибегает к эмуляции и использует только кэшированные копии. Если клиент теряет сильную связь и остается только слабая, он не выполняет накопления, а решает, следует ли проводить выборку файла при сбое кэша, принимая во внимание срочность требуемой информации и тип файла. Слабое соединение, однако, не используется для повторной интеграции файлов.

11.1.2.2. Little Work

Распределенная файловая система Little Work, как и Coda, является расширением AFS [Huston, 1993], [Honyman, 1995]. Little Work требует только изменений клиентской программы управления кэшем и определяет конфликты записи в процессе повторной интеграции. Little Work не имеет специальных инструментов для повторной интеграции и не предлагает служб транзакции.

В то же время Little Work для поддержания согласованности использует больше состояний клиента.

- **Соединение**. Клиент работает в нормальном режиме, т.е. от Little Work не требуется никаких специальных механизмов. Этот режим требует постоянной высокой пропускной способности, подобной доступной в типичной офисной среде, использующей, например, сеть WLAN.
- **Частичное соединение**. Если клиенту доступна только низкая пропускная способность, но при этом еще сохраняется возможность непрерывного соединения, то клиент называется частично соединенным. Примером такого типа сети является сеть пакетной радиосвязи. Эти сети обычно загружаются на основе объема информационных потоков, а не на длительности соединения. Такое состояние клиента позволяет использовать протоколы согласованности кэша подобно нормальному состоянию, однако, с задерживаемой записью на сервер, чтобы снизить стоимость связи, если клиент снова изменит файл. Это позволяет избежать проблем согласованности, хотя соединение с высокой пропускной способностью недоступно.
- **Только выборка**. Если доступна только сеть, предлагающая связь по требованию, клиент переходит в режим "только выборка". Примером сетей такого типа является сотовая сеть, такая как GSM, с платой за вызов. Теперь клиент использует копии в кэше оптимистичным образом, но при этом производит выборку файлов посредством линии связи, если эти файлы не доступны в кэше, что дает пользователю возможность доступа ко всем файлам сервера, при этом делается попытка минимизировать связь, работая с копиями и повторной интеграцией после повторного соединения, используя линию связи с постоянно высокой пропускной способностью.
- **Разъединение**. Наконец, в отсутствие какой-либо сети клиент разъединен. Если происходит сбой кэша, Little Work прекращает работу, в противном случае используются копии.

11.1.2.3. Ficus

В отличие от Coda, Ficus — это распределенная файловая система, не основанная на подходе клиент/сервер [Popek, 1990], [Heidemann, 1992]. Ficus позволяет оптимистично использовать копии, определять конфликты записи и решать конфликты на уровне каталогов. Ficus использует так называемые **протоколы слухов** (gossip protocols), идея которых позже использовалась во многих других системах. Мобильному компьютеру не обязательно нужна прямая связь с сервером. С помощью других мобильных компьютеров он может распространять обновления по сети, пока не достигнет стационарных сети и сервера. Таким образом, изменения в файлах распространяются по сети поэтапно. Более того, Ficus пытается минимизировать обмен файлов, которые действительны только в течение короткого промежутка времени, т.е. временных файлов. Крайне важный вопрос для протоколов слухов — как быстро информация может дойти до клиента, который в ней нуждается, и какой излишний информационный поток порождает распространение информации к клиентам, которым она не нужна.

11.1.2.4. MIo-NFS

Система мобильной интеграции NFS (mobile integration of NFS — MIo-NFS) является расширением системы NFS [Guedes, 1995]. В отличие от многих других систем MIo-NFS использует пессимистический подход с маркерами, управляющими доступом к файлам. Только обладатель маркера определенного файла может изменить этот файл. Таким образом, MIo-NFS избегает конфликтов записи. Как представлено в разделе 11.1.2.1, конфликтов чтения/записи избежать нельзя. MIo-NFS поддерживает три различных режима.

- **Соединение**. Сервер, как обычно, управляет любым доступом к файлам.
- **Слабое соединение**. Клиент использует локальные копии, обмениваясь маркерами по сети и обновляя с ее помощью файлы.
- **Разъединение**. Клиент использует только локальные копии. Запись допустима, если клиент является обладателем маркера.

11.2. World Wide Web

В этом разделе обсуждаются некоторые проблемы, с которыми сталкиваются Web-приложения при использовании в мобильных и беспроводных средах. Читатель должен быть знаком с основными концепциями World Wide Web, протоколом HTTP и языком HTML. В разделе 11.3 представлена завершенная структура, протокол беспроводных приложений WAP, который позволяет устранить многие рассмотренные проблемы; таким образом, настоящий раздел служит основой для этой структуры. Подходы, упомянутые в данном разделе, обсуждаются кратко, основное внимание обращается на развернутое описание структуры WAP. В первых двух подразделах дан краткий обзор HTTP и HTML, а также связанных с ними проблем, которые возникают в беспроводных средах. Затем представлено несколько подходов к улучшению HTML и HTTP. В последнем подразделе вводятся различные системные архитектуры, используемые для доступа к Web-ресурсам, каждая из которых направлена на улучшение классического сценария клиент/сервер.

11.2.1. Протокол передачи гипертекстовых файлов

Протокол передачи гипертекстовых файлов (Hypertext Transfer Protocol — HTTP) — это не меняющий состояния, упрощенный протокол уровня приложений, предназначенный для передачи данных между серверами и клиентами [Berners-Lee, 1996]. **Транзакция** HTTP состоит из **запроса** HTTP, производимого клиентом, и **отклика** HTTP, даваемого сервером. Постоянность состояния означает, что все транзакции HTTP не зависят друг от друга. HTTP не "помнит" ни-

каких транзакций, запросов или откликов. Это приводит к крайне простой реализации без нужды в сложных конечных аппаратах.

Примерный запрос может обрабатываться следующим образом. GET запрашивает следующий далее источник, / указывает на индексный файл в корневом каталоге (index.html). Кроме того, указывается протокол HTTP и его версия 1.0. Поскольку данная книга — это не введение в HTTP, все возникающие вопросы адресуются к большому количеству литературы о Web и ее протоколах. В то же время каждый может попробовать отослать (используя, например, telnet) следующую строку на порт 80 своего Web-сервера:

```
GET / HTTP/1.0
```

Сервер может ответить чем-то подобным приведенному ниже (отклик):

```
HTTP/1.1 200 OK
Date: Fri, 06 Nov 1998 14:52:12 GMT
Server: Apache/1.3b5
Connection: close
Content-Type: text/html
<HTML>
<HEAD>
<TITLE>Institut f&uuml;r Telematik</TITLE>
</HEAD>
<BODY BGCOLOR="#ffffff">
<img src="icons/uni/faklogo_de.gif">
ALT=" [Universit&auml;t Karlsruhe, Fakul&auml;t
f&uuml;r Informatik] ">
...
```

Первая строка содержит код состояния (200), который показывает, что все в порядке (английское OK) (но не OK в открытом тексте, все это отправляется также с кодами ошибки, указывая только на то, что все работает). Далее следует заголовок HTTP, содержащий информацию о дате, времени, версии сервера, соединении и типе следующей далее информации (тело отклика). Здесь содержание — это (усеченный) код HTML Web-страницы index.html в корневом каталоге Web-сервера.

HTTP предполагает надежный базовый протокол, в Internet обычно используется TCP. Версия HTTP 1.0 устанавливает новое соединение для каждого запроса. В версию HTTP 1.1 были интегрированы некоторые усовершенствования (см. раздел 11.2.3). Это означает, например, что если Web-страница содержит пять пиктограмм, два рисунка и некоторое количество текста, то будет установлено всего восемь TCP соединений — одно для страницы в целом, включая текст; пять для выборки пиктограмм; два для рисунков. Типичный **метод запроса** HTTP — это, как уже было показано, команда GET, которая возвращает затребованный ресурс. Команда GET может стать условной, если к заголовку добавлено If-Modified-Since, что позволяет выбирать более новое содержимое. HTTP дополнительно позволяет запрос только заголовка без тела, используя запрос HEAD. Если клиент хочет передать данные функции на сервере, он может использовать метод POST.

Сервер может ответить с помощью различных **кодов состояния**. Примером является упомянутый выше код 200, указывающий на принятие запроса. Кроме того, сервер может перенаправить клиент в другое место, может показать, что для определенного ресурса, который отказывается выполнять запрос, нужна аутентификация пользователя, или что в текущий момент невозможно обработать запрос.

HTTP поддерживает также механизмы простого **кэширования**. Кэширование позволяет избежать излишних повторных передач содержимого, которое не изменилось со времени последнего доступа. Кэш может размещаться где угодно между сервером и клиентом. Обычно каждый клиент поддерживает кэш локально, чтобы минимизировать задержку при переходах/возвратах по Web-страницам. Может также поддерживаться кэш для организации в целом, университета, региона и т.д. Обычно многие пользователи запрашивают одни и те же станицы. Таким образом, имеет смысл кэшировать те страницы, которые ближе клиентам. Кэширование поддерживает различная информация в заголовках. Например, можно присвоить странице дату окончания срока действия. Это означает, что по окончании указанного времени приложение уже не должно кэшировать страницу. Запись "no-cache" в заголовке отключает функцию кэширования, что может быть полезно для страниц с динамическим содержанием. Дополнительную информацию по механизмам кэширования в HTTP можно найти в работе [Berners-Lee, 1996].

Уже в стационарных сетях HTTP стал причиной **множества проблем**, но еще больше их возникает в беспроводных сетях.

- **Пропускная способность и задержка**. HTTP разрабатывался без оглядки на соединения с низкой пропускной способностью или большой задержкой. Исходной средой были объединенные в сеть рабочие станции, запускающие TCP/IP в проводных сетях с пропускной способностью в несколько Мбит/с. Заголовки протокола HTTP довольно длинные и избыточные. Многие информационные поля передаются снова и снова при каждом запросе, поскольку HTTP не меняет своего состояния в процессе исполнения. Заголовки удобочитаемы для людей и передаются в понятном коде ASCII. Более того, серверы передают содержимое несжатым, т.е. если приложения не сжимают содержимое (как в случае изображений, закодированных GIF или JPEG), сервер не будет проводить никакого сжатия. Поскольку обычно для каждого элемента Web-страницы (пиктограммы, изображения и т.д.) используется новое соединение TCP, каждый элемент — это огромный объем служебных издержек. Представьте пиктограмму размером в 50 байт, для которой должно устанавливаться соединение TCP, включая трехстороннее квитирование, передачу данных и надежный разрыв соединения. Как отмечалось в главе 10, это может приводить к обмену клиента и сервера семью PDU! TCP не разрабатывался для такой транзакции, как схема запрос/отклик с обменом лишь некоторыми данными. Как было показано и в главе 10, механизм плавного старта, встроенный в TCP, может вызвать дополнительные проблемы. TCP может быть слишком предусмотрительным в начале передачи, но, перед тем как он сможет использовать доступную полосу пропускания, передача закончится. Иными словами, TCP никогда не прекращает плавный старт указанным образом, порождая слишком большую задержку. Другая проблема вызывается поиском DNS, необходимым для многих элементов Web-страницы, что еще больше снижает пропускную способность и увеличивает задержку. Каждый раз, когда броузер читает гипертекстовую ссылку на новый сервер, он должен разрешить логическое имя в IP-адрес перед выборкой элемента с сервера. Это требует дополнительного запроса на сервер DNS по беспроводной линии связи, добавляя к задержке период кругового обращения пакета.
- **Кэширование**. Хотя кэширование во многих случаях оказывается полезным, оно довольно часто блокируется поставщиками содержимого. Многие компании желают поместить рекламу на Web-страницах и нуждаются в обратной связи, например, в виде количества посещений этой страницы для оценки количества потенциальных покупателей. При наличии кэша между сервером и клиентом компании не смогут получить реалистичную обратную связь. Либо кэшы нуждаются в дополнительных механизмах для создания профилей использования, либо, что используется чаще всего, кэширование блокируется изначально, используя в заголовке ключевое слово `no-cache`. Поставщикам сети нужно, чтобы кто-то платил за страницы и, следо-

вательно, они поддерживают требование отсутствия кэша. В этом случае пользователи страдают от многократной загрузки одного и того же содержимого с сервера. Более того, многие существующие на данный момент страницы содержат динамические объекты, которые не могут быть кэшированы. Примерами являются счетчики доступа, время, дата или другие пользовательские элементы. Это содержание изменяется со временем или при каждом доступе; иногда, по крайней мере, часть страницы остается неизменной и может быть кэширована. Многие современные компании организовывают заказные страницы для каждого пользователя (посредством CGI, ASP и т.п.). Невозможно сохранить закладку на точку, находящуюся на более низких иерархических уровнях линии связи. Вместо этого пользователи должны всегда заходить на страницы компаний с заглавных страниц. Заказанные элементы сохраняются в фрагментах cookie[1]. Это, в большей или меньшей степени, препятствует кэшированию, поскольку названия каналов связи также генерируются динамически, а алгоритмы кэширования не могут выявить доступ к одинаковому содержимому, если каналы связи различны. Кроме того, заглавные страницы компаний часто создаются динамически, в зависимости от типа броузера, аппаратного обеспечения клиента, его локализации и т.д. Даже если бы кэш мог сохранять некоторое постоянное содержимое, его часто невозможно соединить с динамическим остатком страницы. Мобильность также довольно часто препятствует кэшированию, поскольку способ доступа к Web-серверам изменяется со временем вследствие изменения точек доступа. Однако, кэши в точках доступа мобильных сетей могут сэкономить некоторую полосу пропускания и время. Многие механизмы безопасности также препятствуют кэшированию. Аутентификация чаще происходит между клиентом и сервером, а не между клиентом и его кэшем. С ключами аутентификации соотнесено определенное время действия, после которого они больше не действительны, следовательно, кэширование содержимого для защищенных таким образом транзакций бесполезно.

- **Почтовое отправление**. Отправка содержимого от клиента к серверу может вызвать дополнительные проблемы, если клиент в текущий момент разъединен. Запрос `POST` не может быть удовлетворен в разъединенном состоянии, следовательно, сервер нельзя имитировать, принимая почтовое сообщение посредством дополнительных процессов. Однако очевидно, что это вызывает дополнительные проблемы, например, если реальный сервер не принимает почтовое сообщение или если сервер не может принять отложенное сообщение.

11.2.2. Язык гипертекстовой разметки

Язык **HTML** широко используется при описании содержимого Web-страниц в World Wide Web [Raggett, 1998]. Не имеет значения, какая версия используется, все они имеют общие свойства: HTML разрабатывался на стандартных настольных компьютерах, соединенных с Internet посредством стационарных проводников. Эти компьютеры также имеют много общего: относительно высокая пропускная способность (в особенности по сравнению с портативными устройствами), цветные дисплеи с высоким разрешением (24-битовое кодирование, разрешение 1 2000 × 1 024 пикселей), мышь, звуковая система и жесткие диски большой емкости.

Что же могут предложить стандартные портативные устройства? Вследствие ограничений мощности и размеров (все-таки они должны оставаться "портативными") эти устройства имеют довольно маленькие дисплеи, часто лишь черно-белые, с низким разрешением, очень ограниченными пользовательскими интерфейсами (сенсорные экраны, многофункциональные программируемые клавиатуры, голосовые команды и т.д.), а также низкопроизводительные процессоры.

---

[1] Обычно cookie представляется как позиция файла, хранящая специфическую информацию о пользователе для Web-серверов со стороны клиента. Компании могут хранить информацию в cookie и извлекать эту информацию при повторном посещении пользователем Web-страницы.

Более того, сетевое соединение настольных компьютеров часто состоит из ЛВС со скоростью передачи 100 Мбит/с, нескольких соединений ADSL со скоростью 6 Мбит/с или, по крайней мере, соединения ISDN со скоростью 64 Кбит/с. В трансатлантических линиях связи двусторонняя задержка (время распространения в прямом и обратном направлениях) составляет порядка нескольких миллисекунд, возможно, нескольких сотен миллисекунд. Что могут предложить современные беспроводные соединения в глобальных сетях? 9 600 бит/с для стандарта GSM, порядка 100 Кбит/с в будущем, например, в GPRS. Двусторонняя задержка в этом случае составляет порядка нескольких секунд.

Web-страницы, используя текущий HTML, часто игнорируют эти различия конечных систем. Страницы разрабатываются преимущественно для красивого представления содержимого, а не для его эффективной передачи. HTML сам по себе не предлагает никаких способов оптимизации страниц для различных клиентов или различных технологий передачи. Однако, HTML — это не самая большая проблема при доступе к страницам беспроводных портативных клиентов.

Почти все современные Web-страницы, в особенности страницы компаний, "обогащены" специальными "излишествами", некоторые из них используют HTML, а некоторые нет. Эти "излишества" включают в себя анимированные изображения в формате GIF, апплеты Java, управляющие элементы ActiveX, мультимедийное содержимое различных форматов и т.д. Некоторые могут непосредственно интерпретироваться броузером клиента, для других нужны специальные **дополнительные модули** (plug-in). Такие дополнительные форматы содержимого создают некоторые проблемы. Прежде всего, подходящие модули часто доступны только для наиболее распространенных компьютерных платформ, а не для большинства портативных устройств, каждое из которых имеет свою собственную операционную систему. Если бы даже модули и были доступны, то броузер все же имел бы проблемы с отображением, например, при отображении видеоизображения с реалистичным воспроизведением цвета на маленьких черно-белых дисплеях или отображении рисунков GIF с множеством областей, реагирующих на щелчок мыши. Многие Web-страницы для навигации используют именно такие рисунки GIF. Пользователь должен лишь щелкнуть мышью в нужной области — но что, если рисунки GIF нельзя отобразить?

Подходы, использующие отсев содержимого или семантическое сжатие, могут работать с HTML, но каждый из огромного количества дополнительных модулей нуждается в собственном механизме перевода его в формат, пригодный для беспроводных устройств с ограниченными возможностями. Таким образом, без дополнительных механизмов и более интегрированного подхода большие рисунки с высоким разрешением передавались бы на мобильные телефоны с дисплеями с низким разрешением, стоя слишком дорого, поскольку пользователь не знает точно, к чему приведет переход по ссылке. Вообще, Web-страницы обычно игнорируют разнородность конечных систем.

11.2.3. Некоторые подходы, которые могут способствовать беспроводному доступу

Проблемы с HTTP и HTML хорошо известны и привели к появлению многих различных частных решений. На данный момент имеются следующие предложения.

- **Изменение масштаба изображения**. Если страничка содержит рисунок с реалистичным воспроизведением цвета и высоким разрешением, можно уменьшить масштаб этого рисунка до меньшего количества цветов, более низкого разрешения или, наконец, только до названия рисунка. Пользователь может в таком случае решить загрузить рисунок позже. Более того, пользователям, если они интересуются частью картины, могут предлагаться отсечение по границам изображения, изменение масштаба изображения или подробное изучение.
- **Преобразование содержимого**. Многие документы доступны только в определенных форматах, например, PostScript или PDF [Adobe, 1999]. Перед передачей таких документов клиен-

ту, не имеющему подходящей программы чтения, специальные конверторы могут преобразовать этот документ в чистый текст (например, [Fox, 1996a]).

- **Извлечение содержимого/семантическое сжатие**. Наряду с преобразованием содержимого можно проводить также извлечение из документа, к примеру, заголовков или ключевых слов и представлять для использования только их (например, [Bickmore, 1997]). Пользователь может также решить загрузить больше данных, связанных с определенным заголовком или ключевым словом. Можно также попытаться автоматически сгенерировать резюме данного текста. Это называется семантическим сжатием, которое, однако, довольно сложно реализовать для произвольного текста. Извлечение заголовков осуществляется проще, однако, и это иногда невозможно, если заголовки HTML используются для разбивки, а не структурирования документа.
- **Специальные языки и протоколы**. В других подходах пытаются заменить HTML и HTTP другими языками и протоколами, которые лучше адаптированы к беспроводной среде. Примерами являются транспортный протокол портативных устройств (handled device transport protocol — HDTP) и язык разметки для портативных устройств (handled device markup language — HDML) от компании Unwired Planet [King, 1997]. Идеи этих частных решений были интегрированы в расширенный подход (протокол уровня беспроводных приложений), который будет рассмотрен в следующем разделе.
- **Технология принудительной накачки информации**. Вместо извлечения содержимого из сервера, сервер может также принудительно закачивать содержимое клиенту. При этом возможно избежать дополнительных расходов при установлении соединения для каждого элемента, однако, это годится только для определенного содержимого, например, новостей, когда от пользователя не требуется значительной активности.

Обычно многие из таких усовершенствований помещаются в стационарную сеть либо интегрированными в сервер, либо интегрированными в шлюз между стационарной и мобильной сетью. Такие шлюзы приложений уже используются для предоставления содержимого WWW пользователям мобильных телефонов и содержат в себе объекты для сжатия, фильтрации, извлечения содержимого и автоматической адаптации к сетевым характеристикам. В то же время эти частные подходы требуют модернизации стандартных броузеров и в действительности не могут работать с широким спектром разнородных устройств. Однако, и стандартный протокол передачи для содержимого Web, HTTP также был улучшен.

HTTP версия 1.1 (RFC 2068 [Fielding, 1997]) предлагает несколько улучшений.

- **Повторное использование соединения**. Клиент и сервер могут использовать одно и то же соединение TCP для нескольких запросов и откликов. Более того, клиент может отправлять множественные запросы в начале сеанса связи, а сервер может отправить все отклики в том же порядке.
- **Модификации кэширования**. Кэш теперь может также хранить кэшируемые отклики, чтобы снизить время передачи отклика и полосу пропускания для будущих аналогичных откликов. При кэшировании делается попытка достичь независимости от семантики, т.е. кэш не должен оказывать никакого влияния на клиент или сервер, за исключением, возможно, увеличения производительности. Таким образом, увеличилась корректность кэшированных позиций. Чтобы сделать выборку наиболее современной, подходящей версии элемента, может проводиться повторная проверка правильности элемента с помощью исходного сервера. Позиция кэша может рассматриваться как достаточно свежая, может быть включено предупреждение в противном случае, а также может быть указано, что объект не был модифицирован и т.д. Web-страницы могут содержать дополнительную информацию о возможности кэширования и независимости от се-

мантики. Более того, специальные тэги позволяют идентифицировать содержимое и помогают определить, указывают ли различные идентификаторы URI на одинаковое содержимое. Еще несколько тегов дают возможность определить, можно ли кэшировать содержимое, можно ли кэшировать его только в частные кэши и т.д.

- **Разное**. HTTP 1.1 включает в себя дополнительные механизмы передачи сжатых данных, проверки целостности сообщения, аутентификации клиентов, модулей доступа и серверов.

В HTTP с неизменным состоянием могут вводиться некоторые виды состояний, использующие для этой цели маркеры **cookie** [Kristol, 1997]. Cookie могут устанавливать длительный "сеанс связи", сохраняя состояние по запросу. Если сервер запрашивает сохраненное в cookie со стороны клиента, то "инициируется" "сеанс связи". В зависимости от требований сервера cookie могут отображать текущее состояние просмотра, возможности клиента, профили пользователя и т.д. Сеанс связи "возобновляется" путем возвращения cookie серверу. Cookie могут иметь дополнительные атрибуты, например, максимальный возраст. Впрочем, этот механизм в действительности не интегрирован в HTTP и не может заменить реальные сеансы связи механизмами, приостанавливающими сеанс связи по запросу пользователя, возвращающими к определенному состоянию и т.д. Более того, многие пользователи ощущают неудобства при использовании cookie, поскольку совсем не очевидно, что в точности хранят cookie и что они показывают серверам.

11.2.4. Системные архитектуры

Классической базовой системной архитектурой WWW является система клиент/сервер. Клиент, Web-броузер, запускаемый как приложение на компьютере, запрашивает содержимое у сервера, Web-сервера, запускаемого на другом компьютере. Без модернизаций каждый щелчок на гиперссылке инициирует переход к содержимому, на которое указывала ссылка (возможно, чему-то большему, если страница содержит дополнительные ссылки, — броузер делает их выборку автоматически). Для передачи содержимого броузер использует протокол HTTP (см. 11.2.1). Web-страницы описываются с использованием языка HTML (см. 11.2.2) и многих других (частных) форматов.

Как уже говорилось, главная тема в сценарии клиент/сервер — это кэширование. Хотя кэширование полезно также для компьютеров с проводным подключением к сети, поскольку оно снижает задержку доставки ранее запрошенных страниц, кэширование является единственным способом поддержки (частично) разъединенных Web-броузеров. В особенности у мобильных беспроводных клиентов сеанс связи может разрываться или, что бывает довольно часто, иметь очень плохое качество. Таким образом, первым улучшением была интеграция кэширования в Web-броузеры. Для современных броузеров это является стандартом (например, [Netscape, 1999], [Microsoft, 1999]). На рис. 11.3 показан мобильный клиент с запущенным Web-броузером. Броузер модернизирован — в него встроен механизм кэширования. Этот кэш не выполняет автоматической предварительной выборки страниц, а хранит уже переданное содержимое вплоть до определенного предела. Пользователь может тогда перейти в "автономный" режим и при этом просматривать содержимое кэша (страницы, рисунки, мультимедийные объекты и т.д.). Стратегии кэширования очень просты. Пользователь может, например, определить, как выполняется проверка с целью обновления: каждый раз при его обращении к странице, только после перезапуска броузера или как-то иначе (т.е. страница должна обновляться вручную).

На рис. 11.4 показана архитектура для ранних подходов к модернизации доступа к Web для мобильных клиентов. Исходный WebWhacker, например, — это сопутствующее приложение для броузера, которое поддерживает предварительную выборку содержимого, кэширование и службу разъединения [BlueSquirrel, 1999]. В то же время этот подход не является независимым от броузера, так как теперь имеется два различных способа доступа к содержимому (один непосредственно к Web-серверу, другой с помощью дополнительных приложений).

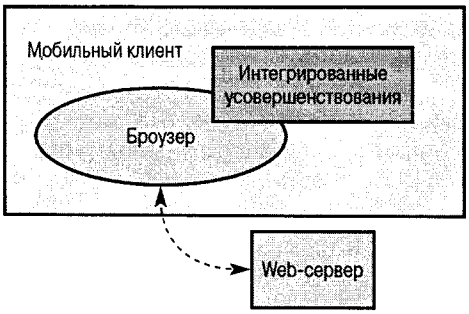

Рис. 11.3. Интегрированная модернизация броузера

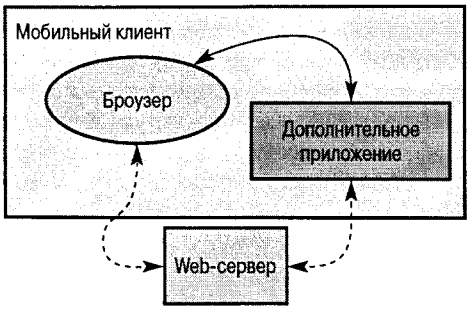

Рис. 11.4. Дополнительные приложения, поддерживающие просмотр

Типичные усовершенствования просмотра Web-ресурсов действуют как независимые модули доступа, как это показано на рис. 11.5. Броузер осуществляет доступ к Web-серверу через модуль доступа клиента, т.е. модуль доступа для броузера действует как сервер, а для Web-сервера — как клиент. Теперь модуль доступа может проводить предварительную выборку и кэширование содержимого согласно различным стратегиям. Как только клиент разъединяется, модуль доступа сохраняет содержимое. Эта схема используется во многих подходах, поскольку не зависит от броузера и, таким образом, допускает другие разработки (например, Caudweb [LoVerso, 1997], TeleWeb [Schilit, 1996], Weblicator [Lotus, 1999], WebWhacker [BlueSquirrel, 1999]).

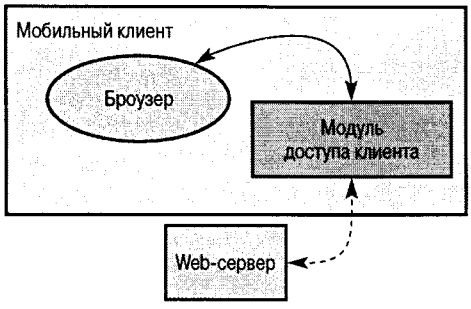

Рис. 11.5. Модуль доступа клиента как поддержка броузера

Примерами стратегий предварительной выборки могут быть: все страницы, на которые ссылается текущая страница; все страницы, включая те, на которые указывали предварительно выбранные страницы (до определенного уровня глубины); страницы, но не рисунки; все страницы с одинаковыми ключевыми словами на одном сервере и т.д.

Модуль доступа может также поддерживать мобильных клиентов со стороны сети (см. рис. 11.6). Этот сетевой модуль доступа может выполнять адаптивное преобразование содержимого (например, семантическое сжатие, извлечение заголовка и т.д., см. раздел 11.2.3) или предварительную выборку и кэширование содержимого. Предварительная выборка и кэширование полезны в беспроводной среде с высокой вероятностью появления ошибок. Подобно модернизациям, произведенным в I-TCP (см. главу 10), разделение доступа к Web на мобильную и стационарную части может значительно улучшить все характеристики системы. Для Web-серверов сетевой модуль доступа выступает как стационарный броузер с проводным доступом. Разъединение мобильного клиента не влияет на Web-сервер. Примерами такого подхода являются TranSend [Fox, 1996a, 1996b], Digestor [Bickmore, 1997].

Можно объединить преимущества клиента и сетевого модуля доступа, что в результате даст системную архитектуру, показанную на рис. 11.7. Данный подход использован, например, в WebExpress [Housel, 1996], [Floyd, 1998]. Модуль доступа клиента и сетевой модуль доступа теперь могут лучше взаимодействовать посредством предварительно выбранных и кэшированных данных. Модуль доступа клиента может проинформировать сетевой модуль доступа относительно поведения пользователя, тогда сетевой модуль доступа, опираясь на эту информацию, может сделать предварительную выборку страниц. В целом, подход остается независимым от Web-сервера и броузера клиента.

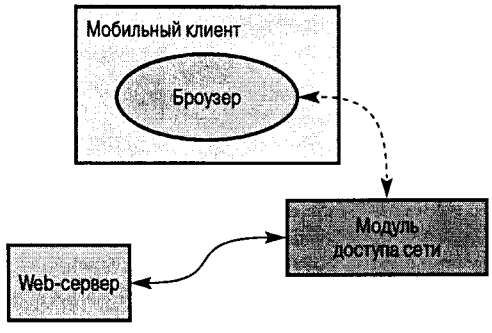

Рис. 11.6. Сетевой модуль доступа как поддержка броузера

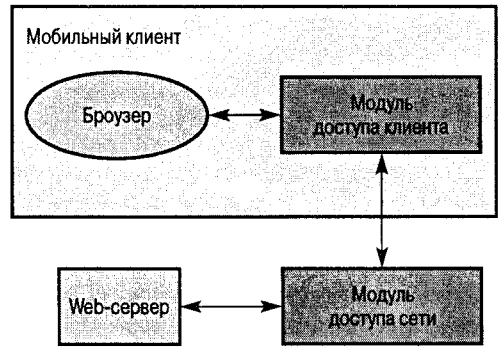

Рис. 11.7. Модуль доступа клиента и сетевой модуль доступа как поддержка броузера

Можно даже продвинуться еще на один шаг дальше и реализовать специализированные сетевые подсистемы, как показано на рис. 11.8. Это решение имеет те же преимущества, что и предыдущее, однако, теперь передаваемое содержимое может подвергаться дальнейшей оптимизации. Примерами являются сжатие в реальном времени и замена протоколов передачи, таких как HTTP и TCP, протоколами, которые лучше адаптированы к мобильности и беспроводному доступу клиента.

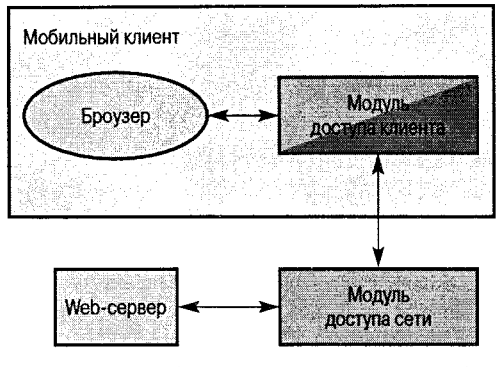

Рис. 11.8. Модуль доступа клиента и сетевой модуль доступа со специальным протоколом передачи

Одним из примеров таких систем является Mowgli [Liljeberg, 1995], [Liljeberg, 1996]. Эта система поддерживает доступ к Web по сотовым телефонным сетям, т.е. сетям с низкой пропускной способностью и относительно высокой задержкой. Система не только заменяет протоколы передачи, но также выполняет дополнительное преобразование содержимого, необходимое для

мобильных телефонов. Броузер по-прежнему использует для модуля доступа клиента протокол HTTP. Затем модуль доступа клиента использует специальную транспортную службу, службу информационного канала Mowgli, для сетевого модуля доступа. Для Web-серверов используются стандартные протоколы. Клиент и сетевой модуль доступа обмениваются сообщениями посредством устойчивых соединений Mowgli избегая, таким образом, плавного старта TCP и выделения на каждый запрос HTTP одного соединения TCP.

Возможны и многие другие усовершенствования. Примерами могут служить расширения сервера, предусматривающие особую адаптацию содержимого к беспроводному доступу и мобильным портативным клиентам. В следующем разделе представлена основа, включающая многие идеи, рассмотренные выше: усовершенствование HTML, поддержка различных системных архитектур и адаптация протоколов передачи к требованиям мобильного беспроводного доступа.

11.3. Протокол беспроводных приложений

Развитие Internet, Internet-приложений и мобильной связи привело к множеству ранних частных решений, обеспечивающих доступ мобильных, беспроводных устройств к Internet-службам. Некоторые проблемы этих частных решений обсуждались в разделе 11.2, поскольку World Wide Web — это наиболее важное и наиболее быстро развивающееся приложение Internet. Чтобы избежать большого количества несовместимых отдельных решений, таких как, например, специальные решения для GSM, IS-136, или определенных производителей, в июне 1997 года фирмами Ericsson, Motorola, Nokia и Unwired Planet был основан **Форум по протоколу беспроводных приложений (wireless application protocol forum — WAP Forum)** [WAP Forum, 1998a].

Основными целями WAP Forum является распространение разнообразного содержимого Internet (например, Web-страницы, службы принудительной накачки информации) и других служб обработки данных (например, фонд библиотечных модулей) на сотовые цифровые телефоны и другие беспроводные терминалы (например, PDA, лэптопы). Кроме того, набор протоколов должен допускать глобальную беспроводную связь различных сетевых беспроводных технологий, например, GSM, CDPD, UMTS и т.д. Таким образом, форум выбирает и расширяет существующие стандарты и технологии во всех областях, где возможна связь с Internet, и создает базу для разработки содержимого и приложений, которые масштабируются на широкий диапазон беспроводных сетей-носителей и типов беспроводных устройств.

Все решения WAP Forum должны быть:

- **способными к взаимодействию**, т.е. для связи с помощью сетей от различных поставщиков допускается использование терминалов и программного обеспечения различных производителей;
- **масштабируемыми**, т.е. протоколы и службы могут расширяться под нужды покупателей и их число;
- **эффективными**, т.е. обеспечение QoS, соответствующего характеристикам беспроводных и мобильных сетей;
- **надежными**, т.е. обеспечение надежной и предсказуемой платформы для внедрения служб;
- **безопасными**, т.е. защита целостности данных пользователя, защита устройств и служб от угроз безопасности [WAP Forum, 1998].

WAP Forum насчитывает на данный момент более 60 членов — производителей аппаратного и программного обеспечения, компьютерных и телекоммуникационных компаний, а также поставщиков сетей. Было учреждено несколько рабочих групп: **рабочая группа по архитектуре WAP (WAP architecture working group — ArchG)** (общая архитектура WAP), **группа по беспроводным протоколам WAP (WAP wireless protocol working group — WPG)** (транспортные протоколы

нейтральных носителей, WSP, WTP, WDP, WCMP), **рабочая группа по беспроводной безопасности WAP (WAP wireless security working group — WSG)** (WTLS) и **рабочая группа по беспроводным приложениям WAP (WAP wireless application working group — WAG)** (форматы содержимого, присваивание имен и т.д., WAE, WML, WMLScript, WTA). Рассматривается возможность создания дополнительных групп (например, по возможности взаимодействия сети). Предметы обсуждения этих групп рассматриваются в следующих разделах. Помимо этого несколько групп обсуждают отношение WAP к работе в консорциуме World Wide Web Consortium (World Wide Web Consortium — W3C) [W3C, 1999], проблемной группе проектирования Internet (Internet Engineering Task Force — IETF) [IETF, 1999], Европейском институте стандартов по телекоммуникациям (European Telecommunications Standards Institute — ETSI) [ETSI, 1999] или консорциуме Bluetooth [Bluetooth, 1999].

WAP Forum опубликовал свой первый набор спецификаций в апреле 1998 года, где были охвачены многие аспекты архитектуры в целом (WAP Forum, 1998a-r). Этот набор составляет основу дальнейших разделов. В разделе 11.3.1 представлена общая архитектура WAP, и сравнивается стандартизация WAP с существующими протоколами и приложениями Internet, в разделах с 11.3.2 по 11.3.9 обсуждаются компоненты архитектуры WAP, а в разделе 11.3.10 представлены примеры конфигураций.

11.3.1. Архитектура

На рис. 11.9 представлен обзор архитектуры WAP, ее протоколов и компонентов и сравнение этой архитектуры с типичной архитектурой Internet при использовании World Wide Web [WAP Forum, 1998a].

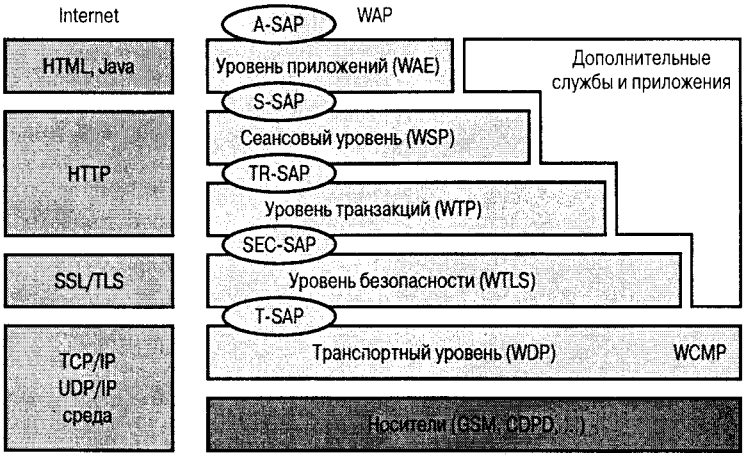

Рис. 11.9. Компоненты и интерфейс архитектуры WAP

Основой для передачи данных являются различные **службы-носители**. WAP не задает службы-носители, а использует существующие информационные службы и будет интегрировать новые службы. Примерами являются службы сообщений, такие как служба передачи коротких сообщений (short message service — SMS) в GSM, передача данных с коммутацией каналов, такая как высокоскоростная передача данных с коммутацией каналов (high-speed circuit switched data — HSCSD) в GSM, или передача данных с коммутацией пакетов, такая как служба всеобщей пакетной радиосвязи (general packet radio service — GPRS) в GSM. Поддерживаются многие другие носители, такие как CDPD, IS-136, PHS. Не задано никаких специальных интерфейсов между службой передачи данных и вышестоящим **транспортным уровнем** с его **беспро-**

водным протоколом дейтаграмм (wireless datagram protocol — WDP) и дополнительным **беспроводным протоколом управляющих сообщений (wireless control message protocol — WCMP)**, поскольку адаптация этих протоколов зависит от носителя. Транспортный уровень предлагает более высоким уровням архитектуры WAP независимую от носителя, последовательную службу, работающую на основе дейтаграмм. Связь осуществляется по одной из доступных служб-носителей. **Точка доступа к службе транспортного уровня (transport layer service access point — T-SAP)** — это общий интерфейс, используемый более высокими уровнями независимо от лежащей в основе сети. Подробнее WDP и WCMP обсуждаются в разделе 11.3.2.

Следующий вышестоящий уровень, **уровень безопасности**, и его **протокол безопасности транспортного уровня в беспроводных сетях (wireless transport layer security — WTLS)** предлагает свои службы в **точке доступа к службе безопасности (SEC-SAP)**. WTLS базируется на безопасности транспортного уровня (transport layer security — TLS, ранее SSL — secure sockets layer, протокол защищенных сокетов), уже известного из WWW. Однако, WTLS был оптимизирован под использование в беспроводных сетях с узкополосными каналами. WTLS может предложить целостность данных, конфиденциальность, аутентификацию и защиту путем отказа от обслуживания. Подробнее этот протокол рассмотрен в разделе 11.3.3.

Уровень транзакций WAP и его **беспроводный протокол транзакций (wireless transaction protocol — WTP)** предлагает облегченную службу транзакции в **точке доступа к службе транзакций (TR-SAP)**. Эта служба эффективно предоставляет надежные или ненадежные запросы и асинхронные транзакции, как объясняется в разделе 11.3.4. С этим уровнем тесно связан следующий вышестоящий уровень, если он используется для служб, ориентированных на установление соединений, как описано в разделе 11.3.5. **Сеансовый уровень** и его **протокол сеансов связи для беспроводной сети (wireless session protocol — WSP)** в настоящее время предлагает две службы в **точке доступа к службам сеансового уровня (S-SAP)**, одну — ориентированную на установление соединения, другую — без установления соединения при использовании непосредственно на вершине WDP. Были определены специальные службы просмотра Web (WSP/B), предлагающие функциональные возможности HTTP/1.1, длительные сеансы связи, откладывание и возобновление сеанса связи, перемещение сеанса связи и другие функции, необходимые при беспроводном мобильном доступе к Web.

На вершине описываемой архитектуры находится **уровень приложений** и **среды беспроводных приложений (wireless application environment — WAE)**, предлагающий основу для интеграции различных приложений WWW и мобильной телефонии. Следовательно, он предлагает множество протоколов и служб со специальными точками доступа к службам, как описано в разделах 11.3.6–11.3.9. Главными вопросами здесь являются языки подготовки сценариев, специальные языки описания, интерфейсы к приложениями телефонии и множество форматов содержания, адаптированных под специальные требования малых портативных беспроводных устройств.

На рис. 11.9 показаны не только общая архитектура WAP, но также и ее связь с традиционной архитектурой Internet для приложений WWW. Транспортный уровень WAP вместе с носителями можно сравнить со службами, предлагаемыми TCP или UDP над уровнем IP и различными средами в Internet. Если носитель в архитектуре WAP уже предлагает службы IP (например, GPRS, CDPD), тогда UDP используется как WDP. Как уже отмечалось, уровень Internet TLS/SSL был также принят для архитектуры WAP с некоторыми изменениями, необходимыми для оптимизации. Функциональные возможности сеансового уровня и уровня транзакций можно грубо сравнить с ролью HTTP в архитектуре Web. В то же время HTTP не предлагает всех дополнительных механизмов, необходимых для эффективного беспроводного мобильного доступа (например, перемещение сеанса связи, откладывание/возобновление). Наконец, уровень приложений предлагает функции, подобные функциям HTML и Java. Как и ранее, были определены специальные форматы и функции, оптимизированные под беспроводный сценарий, а также добавлен телефонный доступ.

WAP не принуждает все приложения постоянно использовать полную протокольную архитектуру. Приложения могут использовать только часть архитектуры, как показано на рис. 11.9. Это означает, что если приложение не требует обеспечения безопасности, а требует надежной передачи данных, оно может использовать службы уровня транзакций и не нуждается в уровне безопасности. Простые приложения могут непосредственно использовать WDP.

Возможны различные сценарии интеграции компонентов WAP в существующие беспроводные и стационарные сети (см. рис. 11.10). Слева показаны различные стационарные сети, такие как традиционная Internet и коммутируемая телефонная сеть общего пользования (public switched telephone network — PSTN). Нельзя изменить протоколы и службы этих существующих сетей и, следовательно, будут реализованы некоторые новые элементы между этими сетями и беспроводными мобильными устройствами, допускающими использование WAP (справа).

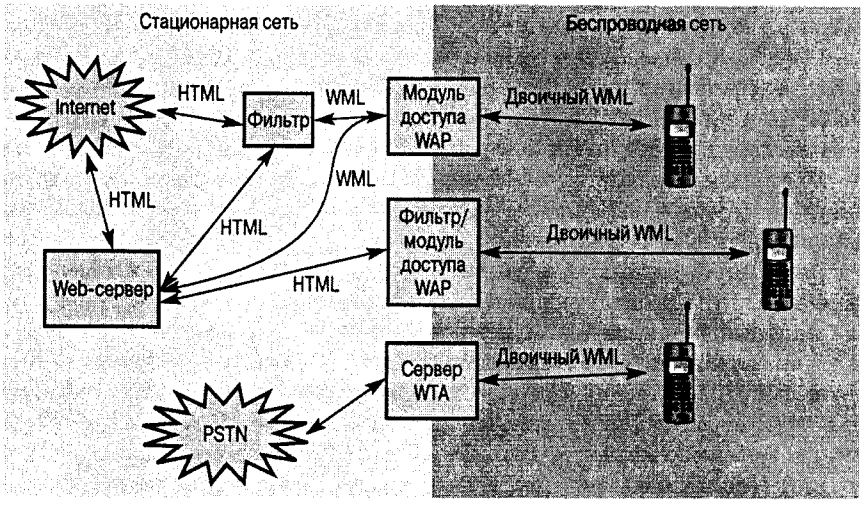

Рис. 11.10. Примеры интеграции компонентов WAP

Современная WWW в Internet предлагает Web-страницы с помощью HTML и Web-серверов. Чтобы иметь возможность просматривать эти или дополнительные страницы посредством портативных устройств, в WAP был определен язык описания для беспроводных систем (wireless markup language — WML). Специальные фильтры стационарных сетей могут преобразовывать HTML в WML; в настоящее время Web-серверы уже могут предоставлять страницы в WML, или же шлюзы между стационарной и беспроводной сетью могут преобразовывать HTML в WML. Эти шлюзы не просто фильтруют страницы, а действуют как модули доступа к Web, что объясняется в следующих разделах. Для более эффективной передачи WML дополнительно конвертируется в двоичный WML.

Подобным образом, используя двоичный WML при доступе к традиционным службам телефонии, будут реализованы специальные шлюзы. Этот сервер приложений беспроводной телефонии (wireless telephone application — WTA) переводит, например, сигнал телефонной сети (входящий вызов и т.д.) в объекты языка WML, отображаемые на портативном устройстве. Важно отметить интегральное представление беспроводного клиента всех различных служб, телефонии и Web посредством WAE (см. раздел 11.3.6).

11.3.2. Протокол дейтаграмм беспроводных сетей

Протокол дейтаграмм беспроводных сетей (wireless datagram protocol — WDP) работает поверх многих различных служб-носителей данных. В точке T-SAP WDP предлагает независимую от

используемых носителей службу последовательной передачи дейтаграмм [WAP Forum, 1998б]. Адаптация, требуемая на транспортном уровне для предоставления этой последовательной службы, может сильно зависеть от служб носителей. Чем ближе службы носителя к IP, тем меньшая нужна адаптация. Если носитель уже предлагает службы IP, как WDP используется UDP [Postel, 1980]. Таким образом, WDP предлагает практически те же службы, что и UDP.

WDP предлагает и **номера портов источника и назначения**, используемые, соответственно, для уплотнения и разуплотнения данных. Примитивом службы для отправки дейтаграмм является **T-DUnitdata.req** с **адресом назначения (AD)**, **портом назначения (DP)**, **адресом источника (SA)**, **портом источника (SP)** и **данными пользователя (UD)** в качестве обязательных параметров (см. рис. 11.11). Адреса назначения и источника являются однозначными адресами для получателя и отправителя данных пользователя. Это могут быть MSISDN (т.е. по сути, телефонные номера), IP-адреса либо другие однозначные идентификаторы. Примитив службы **T-DUnitdata.ind** указывает прием данных. Здесь адрес и порт назначения являются всего лишь дополнительными параметрами.

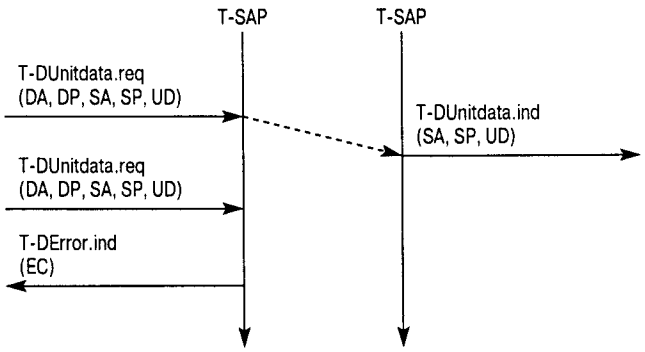

Рис. 11.11. Примитивы службы WDP

Если более высокий уровень запрашивает службу, которую не может предоставить WDP, эта ошибка указывается с помощью примитива службы **T-Derror.ind**, как показано на рис. 11.11. На верхний уровень возвращается **код ошибки (error code — EC)**, указывающий причину ошибки. Однако, этот примитив не должен использоваться WDP для указания на проблемы со службами-носителями, он может указывать только на локальные проблемы, такие как слишком большой размер данных пользователя.

Если происходят какие-либо ошибки при пересылке дейтаграмм WDP от одного объекта WDP к другому (например, пункт назначения недостижим, в указанном порту не ожидает приложение, и т.д.), **протокол управляющих сообщений в беспроводной сети (WCMP)** предоставляет механизм обработки ошибок для WDP [WAP Forum, 1998г] и, следовательно, должен быть реализован. WCMP содержит управляющие сообщения, которые аналогичны сообщениям в протоколе управляющих сообщений Internet (ICMP [Postel, 1981б] для IPv4, [Conta, 1995], для IPv6) и также могут использоваться с диагностическими и информационными целями. WCMP может использоваться узлами и шлюзами WDP для сообщения об ошибках. Однако, сообщения об ошибках WCMP не должны отправляться как отклики на другие сообщения об ошибках WCMP. В сети на основе IP ICMP будет использоваться как WCMP (например, CDPD, GPRS). Типичными сообщениями WCMP являются **недостижимость пункта назначения** (недостижимы маршрут, порт, адрес), **проблемы с параметрами** (ошибка в заголовке пакета), **сообщение слишком большое, сбой повторной сборки** или **запрос отклика/ответа**.

Дополнительный **объект управления WDP** поддерживает WDP и предоставляет информацию об изменениях среды, которые могут повлиять на корректную работу WDP. Важной информацией является текущая конфигурация устройства, доступные в данный момент службы-носители, ресурсы обработки и памяти и т.д. Разработка и реализация таких компонентов управления считается зависимой от поставщика и, следовательно, находится вне области действия WAP.

Если носитель уже предлагает передачу IP, WDP (в данном случае — UDP) зависит от сегментации и возможностей повторной сборки уровня IP, как определено в [Postel, 1981a]. В противном случае WDP должен включать эти возможности, которые необходимы, например, для SMS GSM. Спецификация WAP предоставляет намного большую адаптацию практически ко всем службам-носителям, доступным на данный момент или планируемым в будущем [WAP Forum, 1998q], [WAP Forum, 1998б].

11.3.3. Безопасность транспортного уровня в беспроводных сетях

Если приложению требуется служба безопасности, то в архитектуру WAP поверх WDP может интегрироваться **протокол безопасности транспортного уровня в беспроводных сетях (wireless transport layer security — WTLS)**, как определено в [WAP Forum, 1998в]. WTLS может предоставлять различные уровни безопасности (для конфиденциальности, целостности данных и аутентификации) и был оптимизирован для сетей с низкой пропускной способностью и большой задержкой носителей. Более того, WTLS учитывает низкую производительность и очень ограниченные возможности памяти мобильных устройств относительно криптографических алгоритмов. WTLS поддерживает дейтаграммы и протоколы транспортного уровня, ориентированные на установление соединений. Новым по сравнению, например, с GSM, является безопасность между двумя одноранговыми объектами, а не только между мобильным устройством и базовой станцией. WTLS перенял многие свойства и механизмы TLS (ранее SSL, протокол защищенных сокетов [Dierks, 1999]), но при этом оптимизировал подтверждение установления связи между двумя одноранговыми пользователями.

Перед тем, как станет возможным обмен данными посредством WTLS, необходимо установить безопасный сеанс связи. Установление этого соединения состоит из нескольких этапов; на рис. 11.12 показана последовательность примитивов службы, необходимых для так называемого "полного квитирования" (возможны несколько оптимизаций). Инициатор соединения и одноранговый объект безопасного сеанса связи могут прервать установившийся сеанс связи в любое время, например, если предложенные параметры неприемлемы.

Первым этапом является инициация сеанса связи с помощью примитива **SEC-Create**. Параметрами являются **адрес источника (SA)**, **порт источника (SP)** инициатора соединения, **адрес назначения (DA)**, **порт назначения (DP)** одноранговой пользователя. Кроме того, инициатор предлагает **набор ключей обмена (KES)** (например, используются схемы RSA [Rivest, 1978], DH [Diffie, 1976], ECC [Certicom, 1999]), **набор шифров (CS)** (например, схемы DES, IDEA [Schneier, 1996]) и **метод сжатия (CM)** (в настоящее время далее не специфицирован). Одноранговый объект передает в ответ параметры **режима порядковых номеров (SNM)**, цикла **обновления ключей (KR)** (т.е., как часто в течение сеанса связи обновляются ключи), **идентификатора сеанса связи (SID)** (который уникален для каждого одноранговой пользователя) и выбранных **набора ключей обмена (KES')**, **набора шрифтов (CS')**, **метода сжатия (CM')**. Одноранговый пользователь выпускает также примитив **SEC-Exchange**, что указывает на желание выполнить с клиентом аутентификацию с открытым ключом, т.е. одноранговый объект пользователя запрашивает у инициатора сертификат.

Первый этап установления безопасного сеанса связи, согласование параметров и наборов безопасности указывается со стороны инициатора, далее следует запрос сертификата. Инициатор отвечает своим сертификатом и выпускает примитив **SEC-Commit.req**. Этот примитив указывает, что установление связи полностью завершено со стороны инициатора и что инициа-

тор теперь желает заново переключиться в состояние согласования условий соединения. Одноранговому объекту доставляется сертификат и указывается SEC-Commit. Уровень WTLS однорангового пользователя отправляет инициатору подтверждение. Это завершает полное квитирование для установки безопасного сеанса связи.

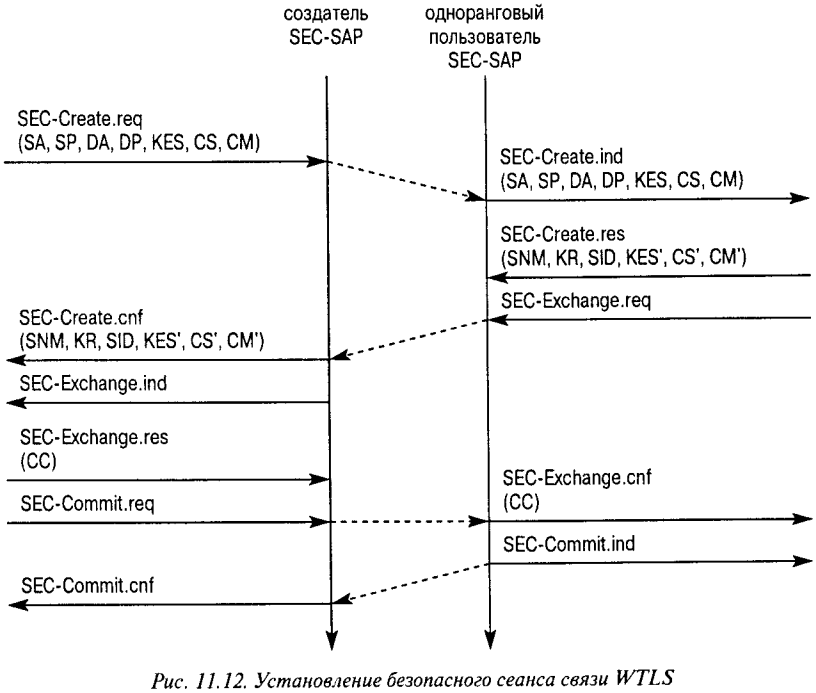

Рис. 11.12. Установление безопасного сеанса связи WTLS

После настойки безопасной связи между двумя одноранговыми пользователями возможен обмен данными, что осуществляется с помощью простого примитива **SEC-Unitdata**, как показано на рис. 11.13. SEC-Unitdata имеет ту же функцию, что и T-DUnitdata на уровне WDP, а именно: она передает дейтаграммы между отправителем и приемником. Такая передача данных ненадежна, но теперь безопасна. Это означает, что WTLS можно легко ввести в стек протоколов поверх WDP. Таким образом, параметры здесь те же: **адрес источника (SA)**, **порт источника (SP)**, **адрес назначения (DA)**, **порт назначения (DP)** и **данные пользователя (UD)**.

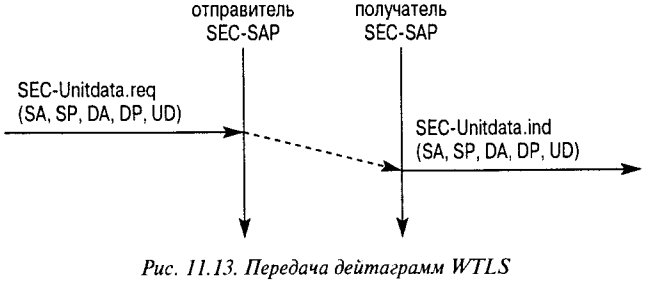

Рис. 11.13. Передача дейтаграмм WTLS

В этом разделе не будут обсуждаться ни связанные с безопасностью свойства WTLS, ни за и против различных алгоритмов шифрования. Читателю рекомендуется спецификация [WAP

Forum, 1998в] и множество прекрасной литературы по криптографии, например, [Schneier, 1996], [Kaufman, 1995].

Хотя WTLS и допускает различные механизмы шифрования с различной длиной ключа, ясно, что, учитывая вычислительную мощность портативных устройств и правила экспорта в некоторых странах, невозможно обеспечить очень сильное шифрование. В то же время приложения или пользователи, если требуется (и разрешено), вольны установить более сильное шифрование над стеком протоколов, — подходящие алгоритмы *доступны* во всем мире.

Дальнейшая работа на уровне WNLS включает последовательную поддержку безопасности уровня приложений (например, использование цифровых подписей) и выбор различных классов реализации с различными возможностями.

11.3.4. Протокол транзакций в беспроводных сетях

Протокол транзакций в беспроводных сетях (wireless transaction protocol — WTP) находится либо сразу над уровнем WDP, либо, если требуется безопасность, над уровнем WTLS [WAP Forum, 1998г]. WTP разрабатывался для работы с очень слабыми клиентами, например, мобильными телефонами. WTP предлагает некоторые преимущества для более высоких уровней, включая улучшенную надежность в службах дейтаграмм, улучшенную эффективность в службах на основе соединений и поддержку служб на основе транзакций, таких как просмотр ресурсов Web. В таком контексте транзакция определяется как запрос с соответствующим откликом, например, для Web-страниц.

WTP предлагает более высоким уровням много функций. Основой являются три **класса служб транзакции**, что объясняется далее. Класс 0 обеспечивает ненадежную передачу сообщений без какого-либо результирующего сообщения. Классы 1 и 2 обеспечивают надежную передачу сообщений, класс 1 без, а класс 2 с одним надежным результирующим сообщением (типичный сценарий запрос/отклик). Надежность WTP достигается за счет **удаления копий**, **повторной передачи**, **подтверждения приема** и использования уникальных **идентификаторов транзакций**. Ни один из классов WTP не требует фазы настройки или обрыва соединения. Это позволяет избежать в канале связи ненужных служебных издержек. WTP допускает **асинхронные транзакции**, **прекращение транзакций**, **конкатенацию сообщений** и может **сообщать** об **успехе** или **сбое** надежных сообщений (например, сервер не может обработать запрос).

Согласно спецификации термин "инициатор" используется для обозначения объекта WTP, инициирующего транзакцию (известного под именем *клиент*), а термин "ответчик" — объекта WTP, отвечающего на транзакцию (известного как *сервер*). Тремя примитивами служб, предлагаемыми WTP, являются **TR-Invoke**, инициирующий новую транзакцию, **TR-Result**, отправляющий назад результаты инициированной ранее транзакции, и **TR-Abort**, прекращающий существующую транзакцию. При нормальных транзакциях между двумя объектами WTP происходит обмен следующими модулями PDU: **активизирующий PDU**, **подтверждающий получение PDU** и **результирующий PDU**. Использование примитивов служб, PDU, связанных с ними параметров и классы служб транзакции будут объяснены ниже.

Специальным свойством WTP является способность обеспечить **подтверждение приема пользователем** или, в качестве альтернативы, автоматическое подтверждение приема объектом WTP. Если требуется подтверждение приема пользователем, пользователь WTP должен подтверждать каждое сообщение, полученное от объекта WTP. Отметим, что подтверждение сообщения пользователем предоставляет более сильную версию службы подтверждения, гарантирующую, что отклик идет от пользователя WTP, а не от самого объекта WTP. Примеры будут даны далее.

11.3.4.1. Класс WTP 0

Класс 0 предлагает службу ненадежной транзакции без результирующего сообщения. Транзакция не меняет своего состояния в процессе исполнения и не может быть прервана. Служба запрашивается

примитивом **TR-Invoke.req**, как показано на рис. 11.14. Параметрами являются **адрес источника (SA)**, **порт источника (SP)**, **адрес назначения (DA)**, **порт назначения (DP)**, как уже объяснялось в разделе 11.3.2. Помимо этого с помощью метки **A** пользователь службы может определить, генерировалось ли подтверждение приема отвечающим объектом WTP или использовалось подтверждение приема пользователем. Уровень WTP будет передавать **данные пользователя (UD)** независимо от адресата. Тип класса **C** здесь указывает класс 0. Наконец, **дескриптор** транзакции **H** предоставляет простой индекс для однозначной идентификации транзакции и является альтернативным именем кортежа (SA, SP, DA, DP), т.е. сокета с локальным значением.

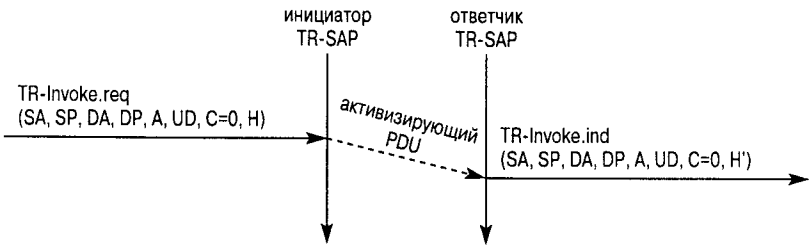

Рис. 11.14. Базовая транзакция, класс WTP 0

Объект WTP инициатора отправляет активизирующий PDU, который получает ответчик. Затем объект WTP ответчика генерирует примитив **TR-Invoke.ind** с теми же параметрами, что и у инициатора, за исключением поля H′, которое теперь является локальным дескриптором транзакции со стороны ответчика.

В этом классе ответчик не подтверждает прием сообщения, и инициатор не осуществляет повторной передачи. Хотя описанный сценарий сходен с простой службой передачи дейтаграмм, WDP рекомендуется использовать тогда, когда требуется только служба по передаче дейтаграмм. Класс WTP пополняет службу транзакций простыми дейтаграммами подобно службе, предназначенной для эпизодического использования более высокими уровнями.

11.3.4.2. Класс WTP 1

Класс 1 предлагает службу надежной транзакции, но без результирующего сообщения. Снова инициатор отправляет активизирующий PDU после **TR-Invoke.req** с более высокого уровня. В данном случае класс равен 1, и, как показано на рис. 11.15, не выбирается подтверждение приема сообщений пользователем. Ответчик подает на более высокий уровень сигнал о поступившем активизирующем PDU посредством примитива **TR-Invoke.ind**, и прием подтверждается автоматически без вмешательства пользователя. Спецификация также допускает подтверждение приема со стороны ответчика, однако, такое подтверждение не является необходимым. Для инициатора транзакция заканчивается после получения подтверждения приема, ответчик сохраняет состояние транзакции еще некоторое время, чтобы быть в состоянии повторить подтверждение приема, если он снова получит такой же активизирующий PDU (что указывает на потерю подтверждения).

Если пользователь службы класса WTP 1 со стороны инициатора запрашивает подтверждение приема со стороны ответчика, дейтаграмма выглядит, как показано на рис. 11.16. Теперь объект WTP со стороны ответчика не отправляет подтверждение приема автоматически, а ждет примитив службы **TR-Invoke.res** от пользователя. Этот примитив службы должен иметь подходящий локальный абстрактный идентификатор H′ для идентификации правильной транзакции. Теперь объект WTP может отправить подтверждающий получение PDU. Типичным применением такого класса транзакций являются службы надежной принудительной накачки информации.

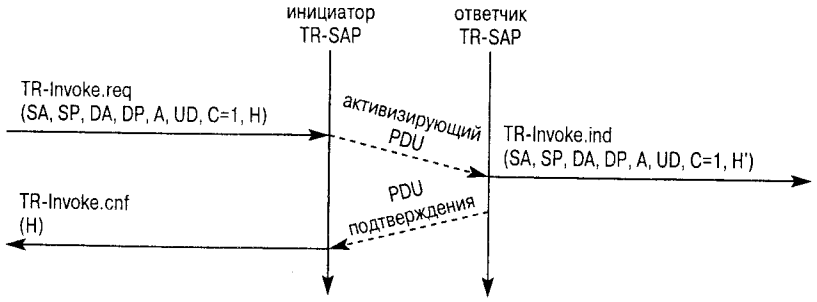

Рис. 11.15. Базовая транзакция, класс WTP 1, нет подтверждения приема

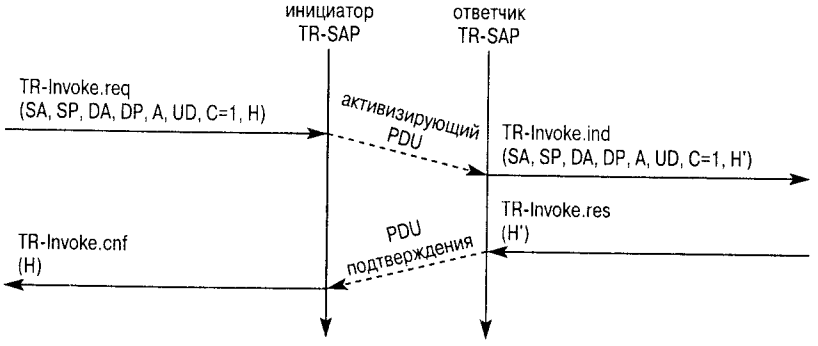

Рис. 11.16. Базовая транзакция класса WTP 1 с подтверждением приема

11.3.4.3. Класс WTP 2

Наконец, служба транзакций класса 2 обеспечивает классическую надежную транзакцию запрос/отклик, известную из многих сценариев клиент/сервер. В зависимости от требований пользователя для взаимодействия инициатор/ответчик возможно много различных сценариев. Далее представлены три примера.

На рис. 11.17 показана базовая транзакция класса 2 без подтверждения приема. Здесь пользователь со стороны инициатора запрашивает службу, объект WTP отправляет активизирующий PDU ответчику. Объект WTP со стороны ответчика указывает запрос с помощью примитива **TR-Invoke.ind**. Далее ответчик ожидает обработки запроса, пользователь со стороны ответчика может в заключение дать объекту WTP со стороны ответчика результирующие пользовательские данные, используя **TR-Result.req**. Теперь результирующий PDU можно отправить назад к инициатору, причем этот PDU неявно подтверждает прием активизирующего PDU. Таким образом, инициатор может обозначить удачную передачу активизирующего сообщения и результат двумя примитивами службы **TR-Invoke.cnf** и **TR-Result.ind**. Пользователь может ответить, используя **TR-Result.res**, затем генерируется подтверждающий прием PDU, который, в конечном счете, запускает примитив **TR-Result.cnf** на стороне ответчика.

Как объяснялось выше, более надежную службу можно предоставить, используя подтверждение приема пользователем. Теперь диаграмма в координатах "время/последовательность" выглядит иначе (см. рис. 11.18). Пользователь со стороны ответчика неявно отвечает на активизирующий PDU, используя примитив **TR-Invoke.res**, который запускает **TR-Invoke.cnf** на стороне инициатора посредством подтверждающего получение PDU. Передача результата также является службой с подтверждением, что указано следующими четырьмя примитивами

службы. Эта служба, вероятно, будет наиболее распространенной в стандартном сценарии запрос/отклик, таком, например, как распределенные вычисления.

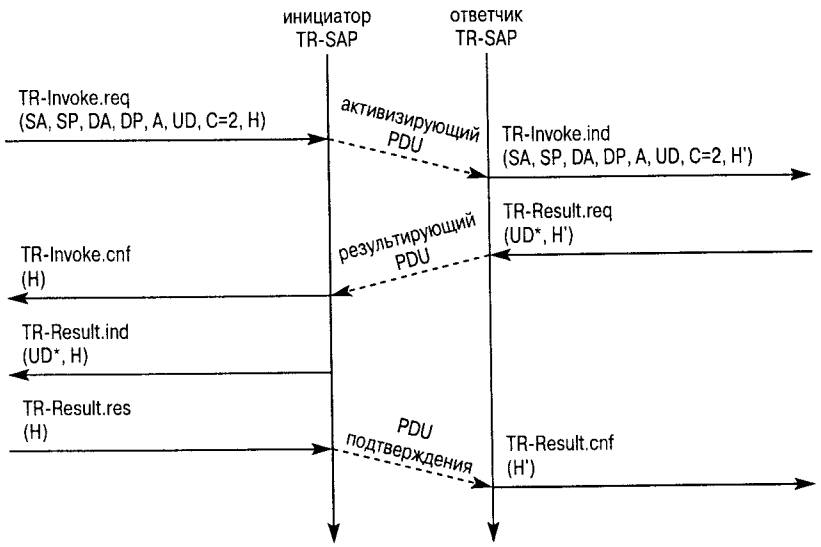

Рис. 11.17. Базовая транзакция класса WTP 2 без подтверждения приема

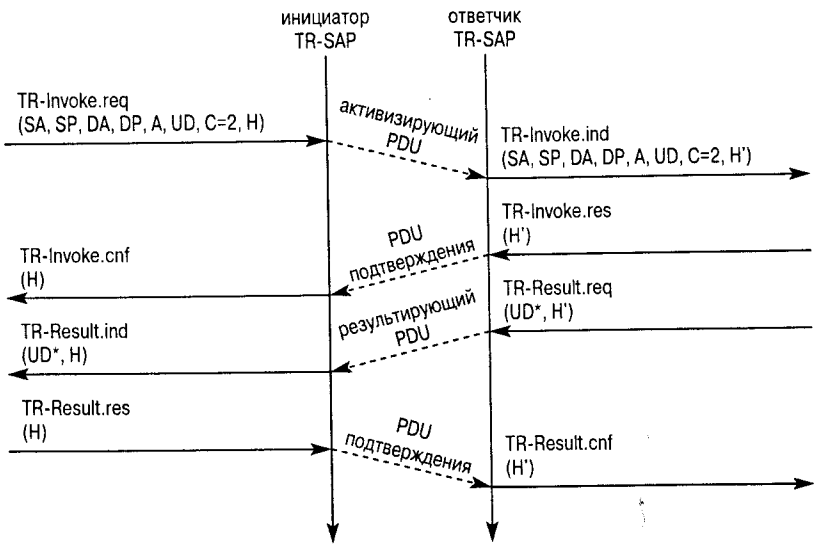

Рис. 11.18. Базовая транзакция класса WTP 2 с подтверждением приема

Если вычисления результата занимают некоторое время, ответчик может перевести инициатора в режим "удержания", чтобы предотвратить повторную передачу активизирующего PDU, поскольку инициатор может предположить, что пакет утерян, если в течение заданного интервала времени не будет отправлен отклик. Это можно сделать так, как показано на рис. 11.19. После истечения времени ожидания ответчик автоматически генерирует подтвер-

ждение приема для активизирующего PDU. Это показывает инициатору, что ответчик действует и в данный момент занят обработкой запроса. Спустя некоторое время инициатору может быть отправлен результирующий PDU.

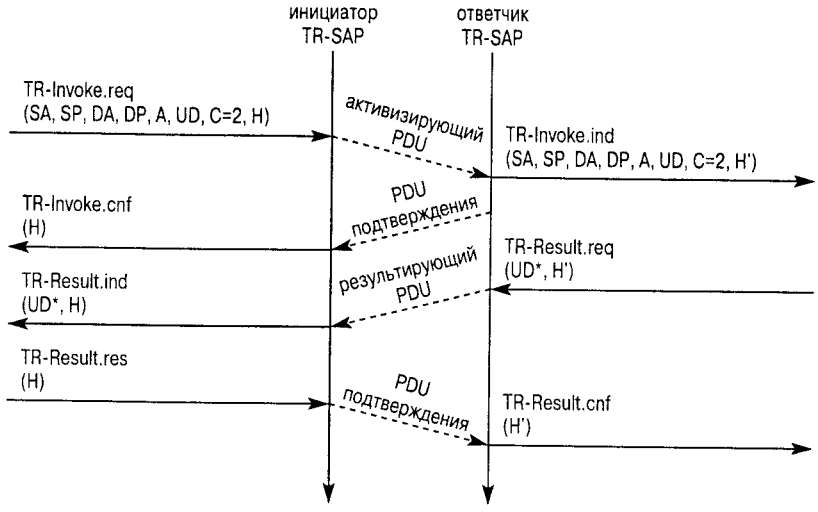

Рис. 11.19. Базовая транзакция класса WTP 2 с "удержанием" без подтверждения приема

WTP предоставляет многие другие функции, не рассмотренные здесь, такие как конкатенация и разделение сообщений, асинхронные транзакции до 2^{15} транзакций, ожидающих выполнения, т.е. затребованных, но до настоящего момента не удовлетворенных, и сегментация/повторная сборка сообщений [WAP Forum, 1998г].

11.3.5. Протокол сеансов связи для беспроводной сети

Протокол сеансов связи для беспроводной сети (wireless session protocol — WSP) разрабатывался для функционирования службой передачи дейтаграмм WDP или службой транзакций WTP [WAP Forum, 1998д]. В любом случае, если требуется, безопасность можно обеспечить на уровне безопасности WTLS. По сути, WSP предоставляет совместно используемое состояние между клиентом и сервером для оптимизации передачи содержимого. HTTP — протокол, который WSP пытается заменить в беспроводных системах, не меняет своего состояния в процессе исполнения, что уже вызвало много проблем в стационарных сетях (см. раздел 11.2.1). По этой причине многие поставщики содержимого Web для хранения состояний на машинах клиентов используют фрагменты cookie, что нельзя назвать элегантным решением (см. раздел 11.2.3). При использовании Web состояния нужны, например, для возобновления просмотра в точно таком же контексте, в котором просмотр был приостановлен. Это очень важно для клиентов и серверов. Пользователи-клиенты могут продолжать работу после приостановки броузера или при разрыве сети или же могут получать выбранную конфигурацию каждый раз при начале просмотра. Поставщики содержимого могут изготавливать на заказ страницы согласно нуждам клиентов и не должны повторно передавать одни и те же страницы. WSP предлагает следующие общие функции, необходимые для обмена содержимым между взаимодействующими клиентами и серверами.

- **Управление сеансом связи**. WSP вводит сеансы связи, которые могут **устанавливаться** от клиента к серверу и могут быть длительными. Сеансы связи могут также надлежащим образом **прерываться**. Для мобильных приложений важной является способность временного

прекращения и **восстановления** сеанса связи. Предположим, что мобильное устройство отключено — пользователю было бы желательно продолжить работу с той точки, где было выключено устройство. Время жизни сеанса связи не зависит от времени жизни транспортного соединения или непрерывной работы сети-носителя.

- **Способность согласования**. Клиент и сервер могут договориться относительно общего уровня функционирования протокола во время установления сеанса связи. Примерными параметрами согласования являются: максимальный размер SDU клиента, максимальное количество ожидающих выполнения запросов, опции протокола, размер SDU сервера.
- **Кодирование содержимого**. WSP также определяет эффективное двоичное кодирование для передаваемого им содержимого. WSP предлагает типизацию содержимого и составные объекты, как объяснялось для использования Web.

Тогда как WSP — это сеансовый протокол общего назначения, WAP задал **беспроводный сеансовый протокол/просмотр (WSP/B)**, который содержит наиболее подходящие для приложений типа просмотра протоколы и службы. В дополнение к основным свойствам WSP WSP/B предлагает следующие, адаптированные к просмотру Web-ресурсов.

- **Функциональные возможности HTTP/1.1**. WSP/B поддерживает функции, HTTP/1.1 [Fielding, 1997], такие как расширенный метод запрос/отклик, составные объекты и согласование типов содержимого. По сути, WSP/B является двоичной формой HTTP/1.1. Таким образом, заголовки содержимого HTTP/1.1 используются для определения типа содержимого, алфавита кодирования, языков и т.д., но для снижения служебных издержек протокола для распространенных заголовков определено двоичное кодирование.
- **Обмен заголовками сеанса связи**. Клиент и сервер могут обмениваться заголовками запросов/откликов, которые остаются неизменными в течение существования сеанса связи. Эти заголовки могут включать типы содержимого, алфавиты, языки, возможности устройств и другие статические параметры. WSP/B не будет интерпретировать информацию заголовков, а будет передавать все заголовки пользователям службы.
- **Передача данных посредством принудительной накачки и извлечения**. Извлечение данных из службы является традиционным механизмом Web. Этот подход также поддерживается WSP/B, где использован механизм запроса/отклика HTTP/1.1. Кроме того, WSP/B поддерживает три механизма извлечения для передачи данных: подтвержденные данные извлекаются в существующем контексте сеанса связи, неподтвержденные данные извлекаются в существующем контексте сеанса связи и без существующего контекста сеанса связи.
- **Асинхронные запросы.** WSP/B может поддерживать клиента, который может отправлять на сервер множественные запросы. Это улучшает эффективность запросов, и ответы могут теперь объединяться в меньшее количество сообщений. Улучшение затрагивает и опоздавшие сообщения, которые могут быть отправлены к клиенту, как только представится возможность.

Как уже отмечалось, WSP/B может запускаться над службами транзакций WTP или службами передачи дейтаграмм WDP. Далее показано несколько последовательностей протоколов, типичных для управления сеансом связи, — метод активации и службы по принудительной накачке.

11.3.5.1. WSP/B над WTP

WSP/B использует три класса служб WTP, представленных в разделах с 11.3.4.1 по 11.3.4.3. Класс 0 используется для неподтвержденной принудительной накачки, возобновления сеанса связи и управления сеансом связи. Подтвержденная принудительная накачка использует класс 1, метод активации, возобновление сеанса связи и управление сеансом связи класса 2. Ниже в качестве примеров приводятся несколько временны*х* диаграмм.

Первый пример (рис. 11.20) показывает установление сеанса связи WSP/B с использованием транзакций WTP класса 2. С помощью примитива **S-Connect.req** клиент может затребовать новый сеанс связи. Параметрами являются **адрес сервера (SA)**, **адрес клиента (CA)** и необязательные **заголовок клиента (CH)** и **запрошенные возможности (RC)**. Сеансовый уровень непосредственно использует схему адресации нижестоящего уровня. TR-SAP и S-SAP могут отображаться непосредственно. Заголовок клиента может включать в себя информацию, передаваемую от пользователя к пользователю, совместимую с заголовками сообщений HTTP согласно [Fielding, 1997]. Эти заголовки могут использоваться, например, для кэширования, если они постоянны в течение всего сеанса связи. Интерпретация заголовков выполняется пользователем службы. Кроме того, требуется согласование возможностей, как описывалось выше для функций WSP.

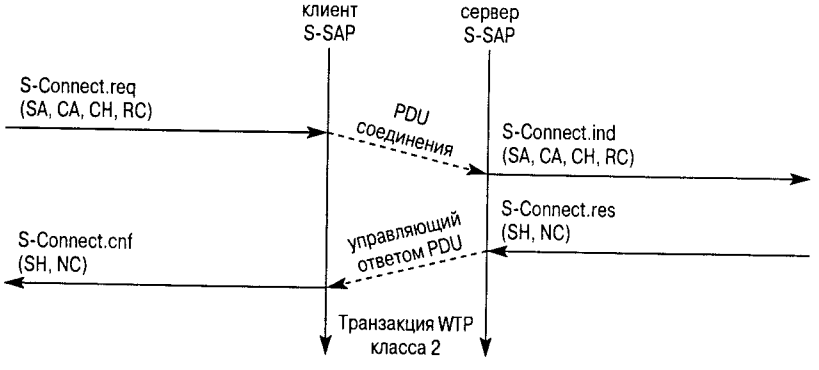

Рис. 11.20. Установление сеанса связи WSP/B

WTP передает **PDU соединения** в точку S-SAP серверу, где примитив **S-Connect.ind** указывает новый сеанс связи. Параметры те же, но теперь возможности являются обязательными. Если сервер принимает новый сеанс связи, он отвечает примитивом **S-Connect.res**, параметрами которого являются необязательный **заголовок сервера (SH)** с теми же функциями, что и заголовок клиента, и **согласованные возможности (NC)**, необходимые для согласования возможностей.

Теперь WTP передает назад клиенту **управляющий ответом PDU**, **S-Connect.cnf** подтверждает установление сеанса связи и включает в себя **заголовок сервера** (если он представлен) и **согласованные возможности** сервера. WSP/B включает несколько процедур для отказа от сеанса связи или прекращения установления сеанса.

На рис. 11.21 показана очень полезная функция WSP/B, **приостановка сеанса связи** и **возобновление сеанса связи**. Если клиент замечает, что скоро он будет недоступен (например, сеть-носитель будет недоступна вследствие роуминга в другую сеть, или пользователь отключает устройство), клиент может приостановить сеанс связи. Приостановка сеанса связи автоматически прервет всю передачу данных и заморозит текущее состояние сеанса связи со стороны клиента и сервера. Клиент прерывает сеанс связи, используя примитив **S-Suspend.req**, WTP передает **приостанавливающий PDU** серверу с использованием транзакции класса 0, т.е. неподтвержденной и ненадежной. WSP/B сообщает о приостановке со стороны клиента и сервера с помощью примитива **S-Suspend.ind**. Единственным параметром является **причина R** приостановки. Причинами могут быть запрос пользователя или приостановка, инициирования поставщиком службы.

Как показано на рис. 11.21, клиент может позже возобновить приостановленный сеанс связи с помощью **S-Resume.req**. Параметрами являются **адрес сервера (SA)** и **адрес клиента (CA)**. Если SA и CA отличаются от использованных перед приостановкой этого сеанса связи, пользователь службы должен соответствующим образом отобразить адреса, чтобы установить

связь с тем же экземпляром сервера. Восстановление сеанса связи является операцией с подтверждением.

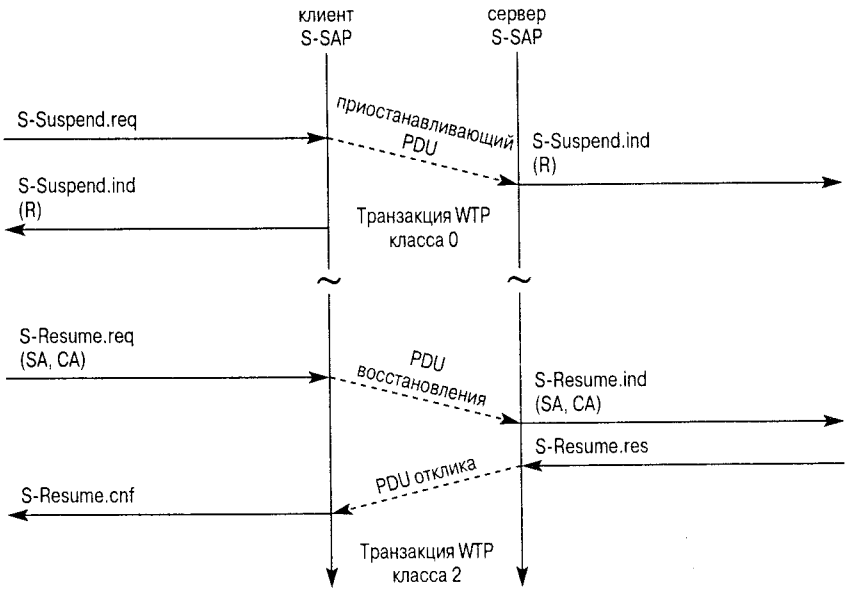

Рис. 11.21. Приостановка и возобновление сеанса связи в WSP/B

Завершение сеанса связи выполняется с помощью примитива службы **S-Disconnect.req** (рис. 11.22). Этот примитив прерывает все текущие транзакции метода или принудительной накачки, используемые для передачи данных. Отключение с обеих сторон указывается с помощью **S-Disconnect.ind**. **Причиной R** разрыва соединения может быть, например, ошибка сети, ошибка протокола, запрос однорангового пользователя, перегрузка, превышение максимального размера SDU. S-Disconnect.ind может также включать параметры, переориентирующие сеанс связи на другой сервер, где сеанс связи может продолжаться.

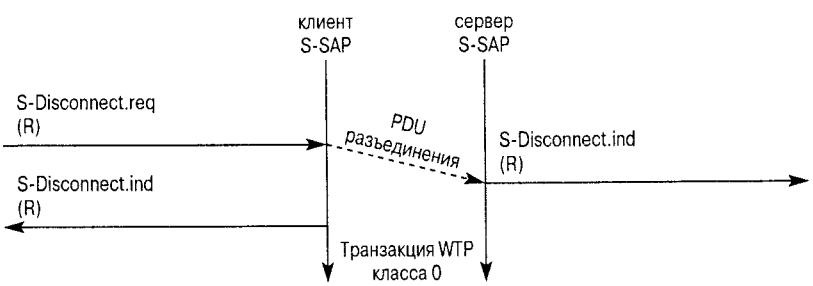

Рис. 11.22. Завершение сеанса связи WSP/B

Примитив **S-MethodInvoke** используется для запроса операций, выполняемых сервером. Результат, если таковой имеется, отправляется назад, используя примитив **S-MethodResult** (рис. 11.23). Клиент запрашивает операцию с помощью **S-MethodInvoke.req**. Параметры — **идентификатор транзакции клиента CTID**, отличающий незаконченные транзакции, **метод M**, идентифицирующий затребованные у сервера операции, и **URI** (Uniform Recourse Identifier —

универсальный идентификатор ресурса) **запроса (RU)** [Berners-Lee, 1994a]. Примерами URI являются идентификаторы URL, такие как http://www.xyz.int/ [Berners-Lee, 1994б]. Используя те же примитивы, могут быть отправлены дополнительные заголовки и тела.

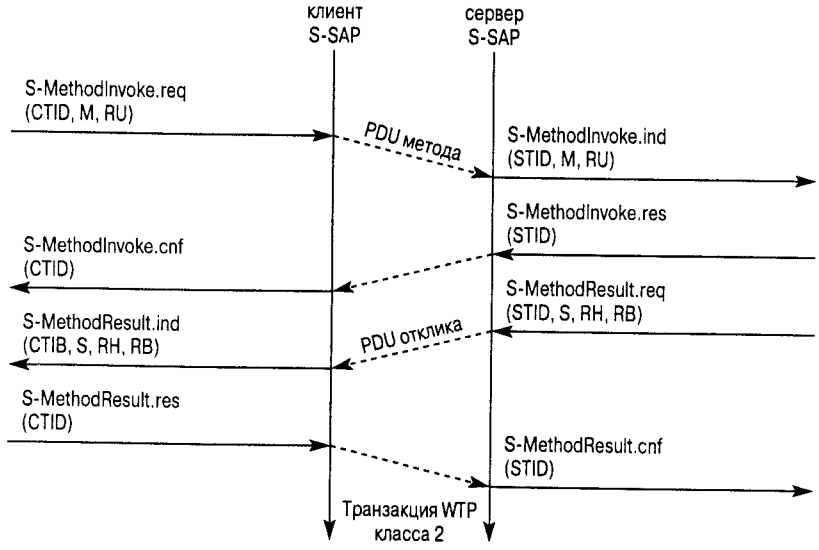

Рис. 11.23. Завершенная транзакция WSP/B

Служба транзакции WTP класса 2 сейчас предает на сервер **PDU метода**. PDU метода может быть либо PDU получения, либо PDU отправки, как определено в [Fielding, 1997]. **PDU получения** используются для методов GET, OPTIONS, HEAD, DELETE и TRACE, а также других методов HTTP/1.1, которые не посылают содержимое на сервер. **PDU отправки** — для методов POST и PUT, а также других методов HTTP/1.1, которые посылают содержимое на сервер.

Со стороны сервера примитив **S-MethodInvoke.ind** указывает запрос. В этом случае незаконченные транзакции различаются **идентификатором транзакций сервера STID**. Сервер подтверждает запрос, и теперь WSP/B не генерирует новый PDU, а полагается на нижние уровни WTP (для большей информации см. рис. 11.24).

Подобным образом результат запроса отправляется назад клиенту (примитив **S-MethodResult**). Дополнительными параметрами теперь являются **состояние (S) заголовок отклика (RH)** и **тело отклика (RB)**. Снова WSP/B весьма подобен HTTP/1.1, так что состояние S соответствует кодам состояния HTTP в работе [Fielding, 1997]. Одним из известных примеров **кода состояния** является **404**, указывающий, что сервер не смог найти Web-страницу, обозначенную в запросе, обычно это указывает на устаревшую закладку или отложенное управление Web-серверов. Все же, как правило, сервер возвращает код **200**, обозначающий, что все нормально. Заголовок и тело также эквивалентны заголовку и телу HTTP, и, следовательно, тело отклика обычно несет код Web-страницы, если состояние равно 200.

WSP не вводит PDU или примитивы службы только ради симметричности и эстетичного вида архитектуры протокола. На рис. 11.24 показано, как WSP (соответственно, и WSP/B) для своих целей использует лежащие в основе службы WTP. Примитив **S-MethodResult.req** запускает примитив **TR-Invoke.req**, параметрами уровня WSP являются данные пользователя уровня WTP. Таким образом, **активирующий PDU** уровня WTP содержит **PDU метода** уровня WSP. В отличие от моделей связи с чисто многоуровневой структурой, здесь более низкий уровень

WTP задействован в семантике примитивов более высокого уровня и не рассматривает их только как чистые данные.

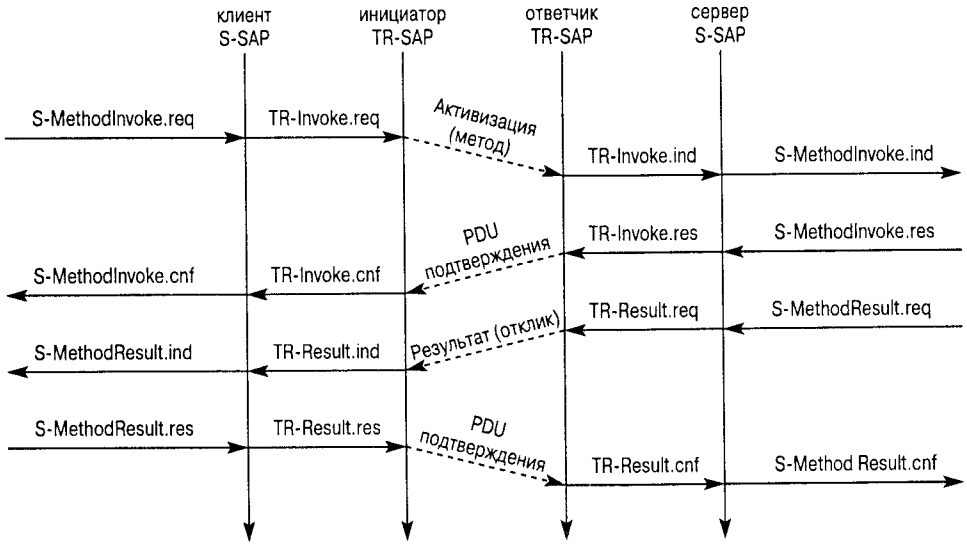

Рис. 11.24. Использование уровнем WSP уровня WTP как нижестоящего

Как показано, для подтверждения примитивов служб уровень WSP не имеет ни одного собственного PDU, он использует **PDU подтверждения приема** уровня WTP. **S-MethodInvoke.res** запускает **TR-Invoke.res**, **подтверждающий получение PDU** передается инициатору, здесь **TR-Invoke.cnf** подтверждает службу активации и запускает примитив **S-MethodInvoke.cnf**, который подтверждает службу активации метода. Это смешение уровней поддерживает большой поток избыточных данных, но также допускает разделение задач между двумя уровнями.

WSP не предоставляет упорядочения различных запросов или восстановления упорядочения откликов. На рис. 11.25 показано четыре запроса со стороны клиента (**S-MethodInvoke_i.req**). WSP может доставить их на сторону сервера в любом порядке, как указано **S-MethodInvoke.ind** (для ясности опущены подтверждения примитивов **S-MethodInvoke.res** и **S-MethodInvoke.cnf**). Кроме того, пользователю со стороны сервера могут требоваться разные периоды времени для ответа на запросы, например, если для каких-либо данных необходимо проводить выборку с диска, а другие данные уже доступны в памяти. Следовательно, отклики **S-MethodInvoke_i.req** могут следовать в произвольном порядке, служба WSP только доставляет их в точку S-SAP клиента, где они, в конечном итоге, появляются как примитив **S-MethodInvoke_i.ind** совершенно независимо от порядка запросов.

До сих пор все примитивы служб позволяли клиенту извлекать данные с сервера. С помощью примитивов принудительной накачки сервер может, если разрешено, принудительно закачивать данные клиенту. Самым простым механизмом принудительной накачки является принудительная накачка без подтверждения, как показано на рис. 11.26. Сервер отправляет клиенту незатребованные данные с помощью примитива **S-Push.req**. Параметрами при этом являются **заголовок принудительной накачки** (**PH**) и **тело принудительной накачки** (**PB**), снова имеем заголовок и тело, известные из HTTP. Ненадежная служба транзакций без подтверждения WTP класса 0 передает **PDU принудительной накачки** клиенту, где **S-Push.ind** указывает, что произошла принудительная накачка.

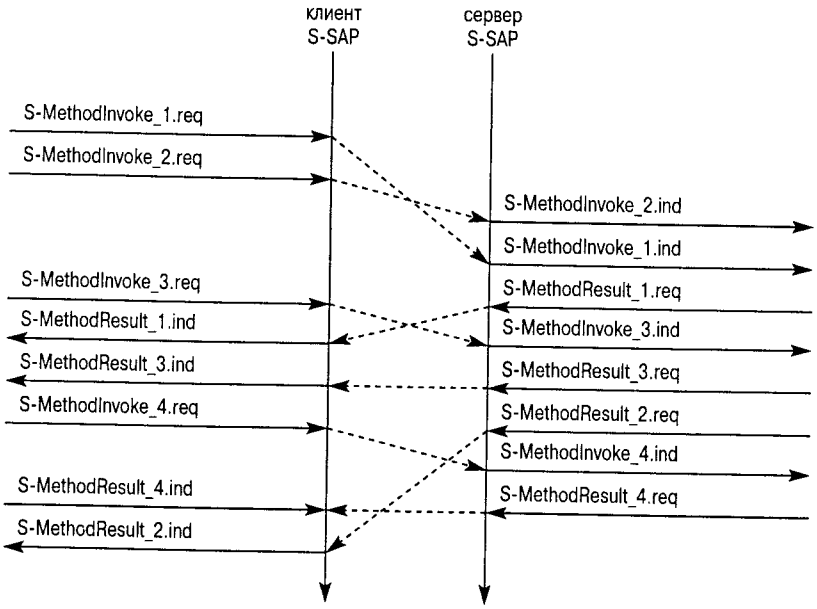

Рис. 11.25. Асинхронные неупорядоченные запросы WSP/B

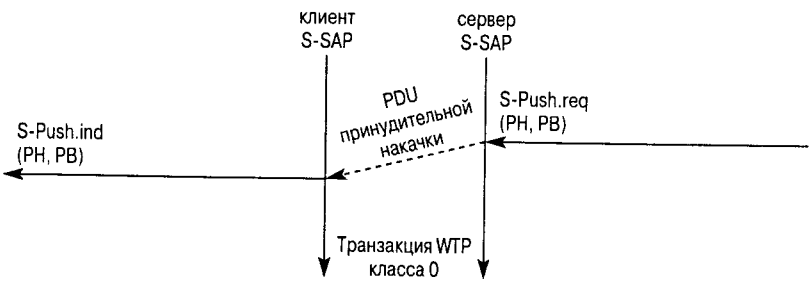

Рис. 11.26. Принудительная накачка без подтверждения WSP/B

Более надежная служба принудительной накачки предлагает примитив **S-ConfirmedPush**, как показано на рис. 11.27. Здесь сервер должен определить принудительную накачку, используя **идентификатор принудительной накачки сервера (SPID)**. Это помогает различать различные незаконченные сеансы принудительной накачки. Теперь надежная служба транзакций WTP класса 1 используется для передачи клиенту **PDU подтверждения принудительной накачки**. Со стороны клиента для различения незаконченных сеансов принудительной накачки используется **идентификатор принудительной накачки клиента (CPID)**.

Кроме того, WSP/B предоставляет множество способов прекращения любых операций, таких как установление сеанса связи, активация метода или принудительная накачка данных.

11.3.5.2. WSP/B как служба сеанса связи без установления соединения

Бывают случаи, когда служебные издержки по установлению сеанса связи и его отключению, согласованию активации метода и всех ассоциированных состояний слишком велики, а высокая надежность не требуется. В этих случаях над ненадежной службой WDP без установления соединения можно запустить WSP/B. В качестве альтернативы WDP, если нужно обеспечить

безопасность, всегда можно использовать WTLS. Примитивы службы непосредственно отображаются друг в друга. Примером области использования может быть периодическая принудительная накачка данных о погоде от уделенного датчика клиенту.

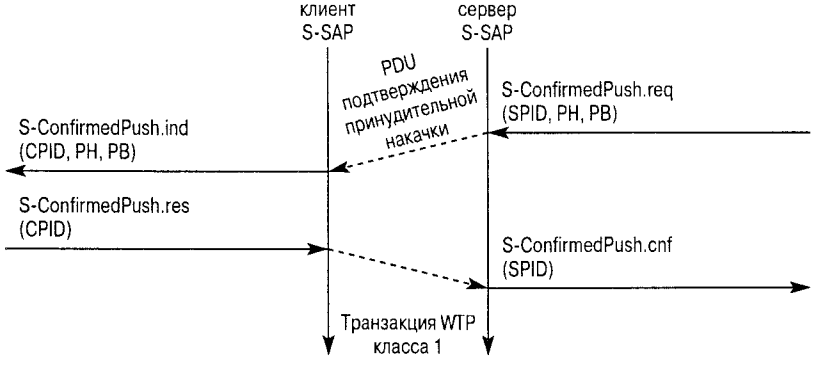

Рис. 11.27. Принудительная накачка с подтверждением WSP/B

На рис. 11.28 показано три примитива служб, доступных для сеансовой службы без установления соединения: **S-Unit-MethodInvoke.req** для запроса операции на сервере, **S-Unit-MethodResult.req** для возврата результата клиенту и **S-Unit-Push.req** для принудительной накачки данных клиенту. Передача модулей PDU (**метод**, **отклик** и **принудительная накачка**) выполняется с помощью стандартной ненадежной службы передачи дейтаграмм WDP. Помимо **адреса сервера (SA)**, **метода (M)** и **URI запроса (RU)**, которые уже объяснялись выше, пользователь примитива **S-Unit-MethodInvode.req** может определить **идентификатор транзакции (TID)** для различения транзакций на уровне пользователя. TID независимо сообщается с двумя пользователями службы.

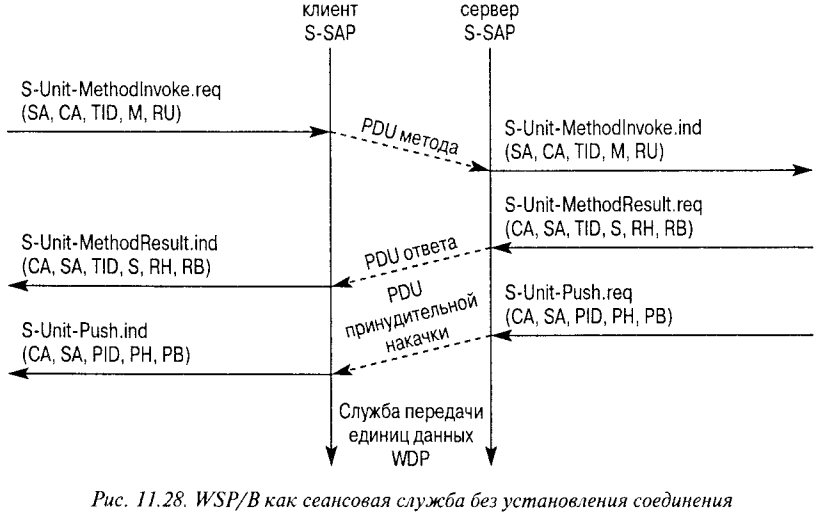

Рис. 11.28. WSP/B как сеансовая служба без установления соединения

Функция примитива **S-Unit-MethodResul** остается той же, что и выше: **состояние (S)**, **заголовок отклика (RH)** и **тело отклика (RB)** представляют результат операции. Наконец, примитив **S-Unit-Push** имеет такие параметры: **адрес клиента (CA)**, **адрес сервера (SA)**, **идентификатор**

принудительной накачки (PID), заголовок принудительной накачки (PH) и тело принудительной накачки (PB).

Хотя WSP уже предлагает много служб, задает двоичное кодирование для заголовков и содержимого и т.д., все же существует множество нерешенных проблем. Примеры: предоставление QoS, т.е. как определенные параметры качества можно применить к транзакциям и сеансам связи, поддержка широковещания, т.е., как можно создать широковещательный сеанс связи, или изохронные мультимедийные объекты, т.е. поддержка очень четких временных рамок службы сеансовых служб. Кроме того, остается открытым вопрос управления всеми службами.

11.3.6. Среда беспроводных приложений

Основной идеей, лежащей в основе **среды беспроводных приложений (wireless application environment — WAE)**, является создание универсальной прикладной среды, базирующейся, главным образом, на существующих технологиях и философии World Wide Web [WAP Forum, 1998ж]. Эта среда должна позволять поставщикам услуг, производителям программного обеспечения или поставщикам аппаратного обеспечения интегрировать свои приложения так, чтобы они могли эффективно распространяться на все многообразие различных беспроводных платформ. Однако, среда WAE не предписывает и не предполагает какой-либо специфической модели интерфейса человек-машина, а допускает разнообразие устройств, каждое со своими собственными возможностями и, возможно, дополнительным оборудованием, зависящим от поставщика (т.е. каждый поставщик может иметь свое собственное мнение). В WEA уже интегрированы нижеследующие технологии, адаптированные для использования в беспроводной среде с портативными устройствами низкой мощности. HTML [Raggett, 1998], JavaScript [Flanagan, 1997] и язык описания для портативных устройств HDML [King, 1997] — это основа для **языка разметки для беспроводных сетей (wireless markup language — WML)** и языка подготовки сценариев **WMLScript**. Были включены форматы информационного обмена для визитных карточек и телефонных книг **vCard** [IMC, 1996a], и для календарей vCalendar [IMC, 1996б]. Можно использовать идентификаторы URL, известные из Web. Кроме того, в **приложения беспроводной телефонии (wireless telephone application — WTA)** был адаптирован и интегрирован широкий круг технологий мобильных телекоммуникаций [WAP Forum, 1998e].

WAE не только полагается на развитые и установившиеся технологии, эта среда фокусируется на устройствах с очень ограниченными возможностями, узкополосными средами и специальными функциями безопасности и контроля доступа. На первом этапе для спецификации WAE был разработан полный набор приложений, в особенности для беспроводных клиентов, как описывается в следующих разделах. Дальнейшее развитие WAE будет включать расширение на большее количество форматов содержимого, интеграцию дополнительных существующих и развивающихся технологий, уделение большего внимания серверам и интеграцию интеллектуальных телефонных сетей.

Одной глобальной целью WAE является минимизация радиопередачи и потребления ресурсов портативными устройствами. Эта цель также отражена в логической модели, лежащей в основе WAE (рис. 11.29) и показанной более подробно, чем общий обзор на рис. 11.10. WAE принимает модель, которая близка к моделям WWW, однако, предполагает дополнительные шлюзы, которые могут увеличить эффективность передачи. **Клиент** выдает на отдаленный сервер закодированный запрос на операцию. Кодирование необходимо для минимизации данных радиопередачи и для экономии ресурсов портативного устройства, что объясняется при обсуждении языков WML и WMLScript.

Теперь декодеры в **шлюзах** переводят данный закодированный запрос в стандартный запрос, который понятен **исходному серверу**. Это может быть запрос на получение Web-страницы или запрос настройки вызова. Шлюз передает этот запрос подходящему исходному серверу так, как если бы он пришел от обычного клиента. Исходным сервером может быть стандартный

Web-сервер, запускающий HTTP и генерирующий содержимое, используя сценарии, предоставляя страницы из базы данных или применяя любые другие (частные) технологии. WAE не определяет стандартных генераторов содержимого или серверов, но предполагает, что большинство будет использовать стандартные технологии, существующие в современной WWW.

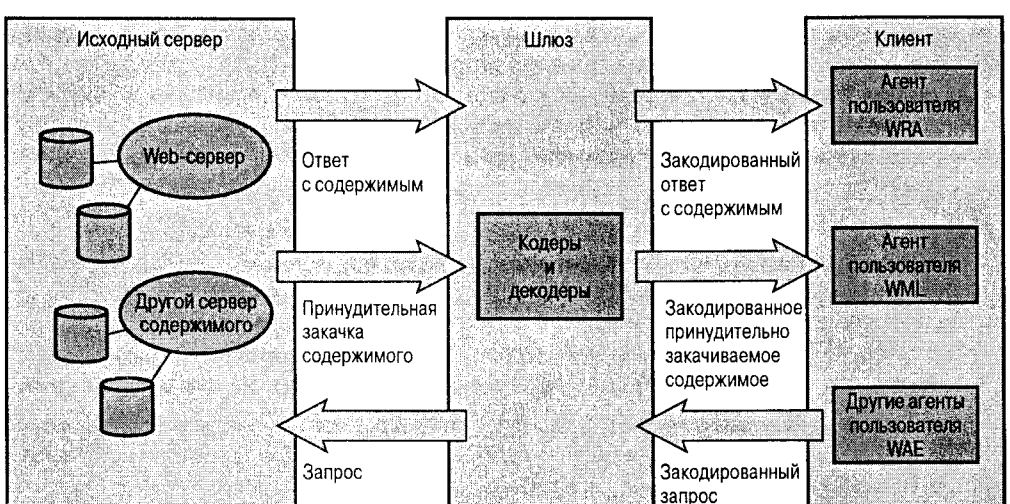

Рис. 11.29. Логическая модель WAE

Исходный сервер отвечает на запрос. Теперь шлюз кодирует отклик и его содержимое (если таковое имеется) и передает закодированный отклик с содержимым клиенту. Логическая модель WAE включает в себя не только эту стандартную схему запрос/отклик, но также и службы принудительной накачки. Затем исходный сервер закачивает содержимое в шлюз. Шлюз кодирует закачанное содержимое и закачивает его клиенту.

В пределах одного клиента может существовать несколько пользовательских агентов. Эти агенты включают такие элементы, как броузеры, телефонные книги, редакторы сообщений и т.д. WAE не определяет ни количества агентов пользователя, ни их функциональных возможностей, однако предполагает наличие базового **пользовательского агента WML**, который поддерживает WML, WMLscript или оба языка (т.е. является "WML-броузером"). Могут реализовываться дополнительные агенты с различными архитектурами, отражающими специфику домена. Этим также занимаются поставщики. В то же время задан еще один пользовательский агент с его основными службами, **пользовательский агент WTA**. Этот агент занимается доступом к мобильным телефонам и взаимодействием с ними (например, управление вызовами). Базовые языки WML и WMLscript, а также WTA будут описаны в следующих трех разделах.

11.3.7. Язык разметки для беспроводных сетей

Язык разметки для беспроводных сетей (wireless markup language — WML) (WAP Forum, 1998ж) основан на стандартном HTML [Raggett, 1998], известном из WWW, и на HDML [King, 1997]. WML специфицирован как тип документа XML [W3C, 1998]. При разработке WML были учтены несколько ограничений беспроводных портативных устройств. Прежде всего, беспроводная линия связи всегда будет иметь крайне ограниченные возможности по сравнению с проводной. Более того, существующие на данный момент портативные устройства обладают маленькими дисплеями, ограниченными пользовательскими средствами ввода, ограниченной памятью и исключительно низкими вычислительными возможностями.

В WML использована модель колоды карт. Документ WML составлен из множества **карт**. Карты могут группироваться в **колоду**. Колода WML подобна странице HTML в том, что она определяется идентификатором URL и является единицей передачи содержимого. Пользователь перемещается по последовательности карт WML с помощью WML-броузера, просматривает содержимое, записывает затребованные данные, делает выборы и т.д. WML-броузер делает выборку колод по запросу со стороны исходных серверов. Эти колоды могут быть либо статичными файлами на сервере, либо генерироваться динамически.

Важно отметить, что WML не задает, как реализация WML-броузера должна взаимодействовать с пользователем. Напротив, WML абстрактно описывает смысл взаимодействия. Пользовательский агент на портативном устройстве решает, как лучше всего представить все элементы карты. Это представление сильно зависит от возможностей устройства.

WML включает в себя несколько базовых элементов.

- **Текст и изображения**. WML, как и другие языки разметки, подсказывает, как текст и изображения можно представить пользователю. В то же время, за точное представление данных пользователю отвечает пользовательский агент, запускаемый на портативном устройстве. WML обеспечивает только набор элементов разметки, определяющих элементы визуального выделения для текстов (полужирный, курсив и т.д.) или выравнивание столбцов таблицы.
- **Взаимодействие пользователя**. WML поддерживает различные элементы для пользовательского ввода. Примерами являются ввод текста или пароля, выбор опций или управление вызовом задач. Как и ранее, агент пользователя свободен выбирать, как этот ввод будет реализован. Выбор может ограничиваться, например, физическими ключами, программными ключами или голосовым вводом.
- **Навигация**. Как и HTML-броузеры, WML предлагает механизм предыстории с навигацией по истории просмотра, гиперссылки и другие элементы навигации между картами.
- **Управление контекстом**. WML позволяет сохранять состояния между различными колодами без взаимодействия с сервером, т.е. переменное состояние может захватывать несколько колод. Карты могут иметь параметры, определяемые использованием этого состояния без доступа к серверу по узкополосному беспроводному каналу.

В следующем абзаце дан простой пример WML. После ключевого слова `WML` определена первая карта `CARD`. Эта карта колоды "отображает" текст после загрузки ("отображение" может означать голосовой вывод и т.п.). Как только пользователь активирует элемент `DO` (кнопка или голосовая команда), агент пользователя отображает вторую карту. На второй карте пользователь может выбрать одну из трех типов пиццы. В зависимости от выбора пользователя `PIZZA` может иметь одно из значений `Mar`, `Fun` или `Vul`. Повторимся, описание этих опций с помощью WML не означает автоматически, что они отображаются в виде текста. Возможно, агент пользователя прочтет опцию посредством голосового вывода, а пользователь ответит путем голосового ввода. WML описывает только суть выбора.

```
<WML>
    <CARD>
     <DO TYPE = "ACCEPT">
           <GO UTL = "#card_two"/>
     </DO>
     Это просто первая карта!
     На следующей вы сможете выбрать ...
    </CARD>
```

```
    <CARD NAME = "card_two">
    ... вашу любимую пиццу:
     <SELECT KEY = "PIZZA">
           <OPTION VALUE = "Mar">
               "Маргарита"</OPNION>
           <OPTION VALUE = "Fun">"Фунжи"
              </OPTION>
           <OPTION VALUE = "Vul">"Вулкано"
              </OPTION>
     </SELECT>
    </CARD>
</WML>
```

WML можно закодировать, используя сжатое двоичное представление для экономии пропускной способности беспроводной линии связи. Это сжатое представление основано на двоичном формате содержимого XML, как определено в [WAP Forum, 1998л]. Двоичное кодирование WML является только одной из частных версий этого формата, сжатое представление вообще характерно для содержимого XML. Сжатый формат допускает передачу без потери функциональных возможностей или семантической информации. Например, префикс URL `http://`, который очень часто входит в идентификатор URL, будет закодирован как 4B. Ключевое слово `CARD` со следующими за ним содержимым и атрибутами будет закодировано как E9. Кодом для ключевого слова `SELECT` будет 37, а для `OPTION` — 35. Эти однобайтовые коды намного эффективнее открытого текста, используемого в HTML.

11.3.8. WMLScript

WMLScript служит дополнением WML и предоставляет общую возможность создания сценариев в архитектуре WAP [WAP Forum, 1998з]. Если все содержимое WML статично, WMLScript предлагает несколько возможностей, которые не поддерживает WML.

- **Проверка достоверности данных, введенных пользователем**. Перед отправлением на сервер данных, введенных пользователем, WMLScript может проверить достоверность и сэкономить пропускную способность и время ожидания при наличии ошибки. В противном случае проверку должен проводить сервер, что включает, по крайней мере, время на передачу и подтверждение приема.
- **Возможность доступа к устройствам**. WMLScript предлагает функции доступа к аппаратным компонентам и программным функциям устройств. По телефону пользователь может, например, сделать телефонный звонок, получить доступ к адресной книге или отправить сообщение посредством службы сообщений мобильного телефона.
- **Локальное взаимодействие с пользователем**. WMLScript может непосредственно и локально взаимодействовать с пользователем, показывая сообщения или приглашая ввести данные, не порождая задержки вследствие передачи и подтверждения приема. На сервер могут передаваться только, например, результаты нескольких взаимодействий.
- **Расширение программного обеспечения устройства**. С помощью WMLScript можно сконфигурировать и добавить новые функциональные возможности даже после внедрения устройства. Пользователь может легко загрузить программное обеспечение от производителей и, таким образом, легко заменить свое программное обеспечение на новую версию.

Язык WMLScript основан на языке JavaScript [Flanagan, 1997], однако, он адаптирован для беспроводной среды. Это включает в себя небольшую память простого интерпретатора байт-кода WMLScript и эффективную радиопередачу посредством пространственно-эффективного байт-кода. Для генерации байт-кода используется компилятор WMLScript. Данный компилятор может размещаться в шлюзе (см. рис. 11.29), или же исходные серверы хранят предварительно скомпилированный байт-код WMLScript.

Кроме того, WMLScript основан на событиях, т.е. сценарий может активизироваться как реакция на запрос определенного пользователя или событие среды. WMLScript имеет также полный доступ к модели состояний WML, т.е. WMLScript может устанавливать и читать переменные WML.

WMLScript предоставляется много возможностей, известных из типовых языков программирования, — функции, выражения или операторы `while`, `if`, `for`, `return` и т.д. Этот язык слабо типизирован, т.е. любые переменные могут содержать любые типы (такие как целый, с плавающей запятой, строчный, булевский) — нет необходимости в явной типизации. WMLScript предоставляет автоматический переход между различными типами, если это возможно. Параметры функциям всегда передаются значениями.

Ниже приводится простой пример нескольких строк WMLScript: функция `pizza_test` принимает как вход одно значение. Локальная переменная `taste` (англ. "вкус") изначально имеет значение "`unknown`" ("неизвестен"). Затем сценарий проверяет, имеет ли введенный параметр `pizza_type` значение "`Mar`". Если да, то переменной `taste` присваивается значение "`well...`" ("хорошо!"), в противном случае программа проверяет, будет ли `pizza_type` иметь значение "`Vul`". Если да, `taste` присваивается значение "`quite hot`" ("достаточно пикантно"). Наконец, текущее значение `taste` возвращается в виде значения функции `pizza_test`.

```
function pizza_test (pizza_type) {
    var taste = "unknown";
    if (pizza_type = "Mar") {
     taste = "well...";
    }
    else {
     if (pizza_type = "Vul") {
           taste = "quite hot";
     };
    };
    return taste;
};
```

Компилятор WMLScript может компилировать один или несколько таких сценариев в **единицу компиляции WMLScript**. Портативные беспроводные устройства могут теперь делать выборку таких единиц компиляции, используя стандартный протокол с семантикой HTTP, такой как WSP (см. раздел 11.3.5). В пределах единицы компиляции пользователь может вызвать определенную функцию, используя стандартные указатели **URL** с **анкером фрагмента**. Анкер фрагмента определяется указателем URL, диезом (#) и идентификатором фрагмента. Если бы URL единицы компиляции сценария был `http://www.xyz.int/myscr`, пользователь мог бы вызвать сценарий и передать параметр "Vul" посредством `http://www.xyz.int/myscr#pizza_test('Vul')`.

WAP Forum определил спецификацию нескольких **стандартных библиотек** для WMLScript [WAP Forum, 1998и]. Эти библиотеки обеспечивают доступ к основным функциональным возможностям клиента WAP и, следовательно, должны быть доступны в клиентской среде подготовки сценариев. Одно исключение составляют библиотеки с плавающей запятой, которые не

являются обязательными и полезны, только если клиент поддерживает операции с плавающей запятой. На настоящий момент определены следующие шесть библиотек.

- **Lang**. Эта библиотека предоставляет функции, тесно связанные с самим WMLScript. Примером является `isInt`, проверяющая, можно ли преобразовать значение в целое, или `float`, проверяющая, поддерживаются ли операции с плавающей запятой.
- **Float**. В этой библиотеке (которая является необязательной, как отмечалось выше) имеется множество типичных арифметических операций с плавающей запятой. Примерами функций являются `round` для округления чисел и `sqrt` для вычисления квадратного корня данного значения.
- **String**. В этой библиотеке есть много функций для работы со строками. Примерами являются общеизвестные функции, такие как `length`, возвращающая длину строки, или `subString`, возвращающая подстроку данной строки. Однако, эта библиотека предоставляет также усовершенствованные функции, такие как `find`, находящая подстроку внутри строки, или `squeeze`, заменяющая несколько последовательных пробелов только одним.
- **URL**. Эта библиотека предоставляет множество функций работы с идентификаторами URL с синтаксисом, определенным в [Fielding, 1995]:

  `<схема>://<хост>:<порт>/<путь>;<параметры>?<запрос>#<фрагмент>`

Например, `http://www/xyz.int:8080/mypages;5;2?j=2&p=1#crd`. Функция `getPath` может теперь извлечь путь к этому идентификатору URL, т.е. `"mypages"`, `getQuery` в качестве возвращаемого значения содержит запрос `"j=2&p=1`, а `getFragment` выдает фрагмент, используемый в URL, т.е. `"crd"`.

- **WMLBrowser**. Эта библиотека предоставляет несколько функций, типичных для броузера, таких как `prev`, возвращающая назад на одну карту, или `refresh`, обновляющая контекст пользовательского интерфейса. Функция `go` загружает содержимое, представленное как параметры:

```
var my_card = "http://www.xyz.int/pizzamatic/apps.dck#
start";
var my_vars = "j=4&k=7";
WMLBrowser.go(my_card, my_vars);
```

- **Dialogs**. Эта библиотека была определена для взаимодействия с пользователем. Примером функции является `prompt`, которая отображает данное сообщение и предлагает пользователю произвести ввод.

11.3.9. Приложения беспроводной телефонии

Просмотр Web-ресурсов с использованием WML-броузера является лишь одним из приложений для пользователей портативных устройств. Кроме того, пользователь по-прежнему желает как обычно совершать телефонные звонки и получать доступ ко всем функциям мобильной телефонной сети с помощью мобильного телефона. Именно здесь оказываются полезными **приложение беспроводной телефонии (wireless telephone application — WTA)**, **пользовательский агент WTA** (как показано на рис. 11.29) и **интерфейс приложений беспроводной телефонии WTAI**. WTA — это совокупность специфических расширений звонков и механизмов управления опциями, объединяющих сети передачи данных и сети телефонной связи [WAP Forum, 1998].

Контур WTA интегрирует усовершенствованные службы телефонии, используя совместимый пользовательский интерфейс (например, WML-броузер), и позволяет операторам сети

увеличивать доступность различных специальных служб в своей сети. Более того, операторы сети, используя WTA, могут достигать большего количества конечных устройств, поскольку они интегрированы в среду беспроводных приложений (wireless application environment — WAE), которая поддерживает различные устройства и среды. WTA также должна давать возможность сторонним разработчикам и операторам сети создавать содержимое, не зависящее от сети, которое имело бы доступ к основным функциям сети-носителя. В то же время большинство функциональных возможностей WTA зарезервированы для сетевых операторов из соображений безопасности и устойчивости.

WTA расширяет базовую модель приложений WAE по трем пунктам.

- **Принудительная закачка содержимого**. Исходный сервер WTA может принудительно закачивать клиенту содержимое, т.е. колоды WML, сценарии WMLScript. Это содержимое может, например, давать возможность клиенту подстраиваться под новые события в сети, которые до этого были неизвестны. Пример приводится на рис. 11.31.
- **Подстройка под события сети**. Устройства могут иметь таблицу, указывающую, как следует реагировать на определенные события в мобильной сети. Событием может быть входящий звонок или текстовое сообщение. Эти устройства могут искать способ реагирования, например, выполнять поиск в личной телефонной книге для преобразования входящего телефонного номера в имя.
- **Доступ к функциям телефонии**. Приложения, запускаемые клиентом, могут очень просто получать доступ к функциям телефонии из WML или WMLScript. В библиотеках имеются множество функций настройки звонка, создания записей в телефонной книге и т.д.

Эти библиотеки были определены в спецификации WTAI [WAP Forum, 1998м] и допускают создание приложений телефонии с использованием пользовательского агента WTA. Можно использовать библиотечные функции из колод WML или WMLScript. Пользовательский агент WTA отображает карты колоды или выполняет сценарий и получает события из мобильной сети или производит принудительную закачку содержимого из исходного сервера WTA.

Было определено три класса библиотек.

- **Общие сетевые службы**. Этот класс содержит библиотеки для служб, общих для всех мобильных сетей. Библиотека **call control** (управление звонками) содержит, например, функции настройки звонков, принятия звонков и отказа от них. Библиотека **network text** (сетевой текст) содержит функции отправки, чтения и удаления текстовых сообщений. **Phonebook** (телефонная книга) позволяет манипулировать записями локальной телефонной книги (например, чтение, запись, удаление). Наконец, библиотека **miscellaneous** (разное) содержит, например, функции показа входящих электронной почты, фототелеграмм, информационных или голосовых сообщения.
- **Специфические службы сети**. Библиотеки этого класса зависят от возможностей мобильной сети. Кроме того, этот класс может содержать библиотеки, зависящие от операторов.
- **Общественные службы**. Этот класс библиотек содержит функции, допускающие открытость, т.е. функции, которые могут использовать не только операторы сети, но и сторонние поставщики. Одним из примеров является "вызов" для настройки телефонного звонка.

Все функции в этих библиотеках используют одинаковый синтаксис. Для использования в WML используется следующий формат URI [Berners-Lee, 1994a].

```
wtai://<библиотека>,<функция>;<параметры>;!<результат>
```

Первый параметр <библиотека> указывает имя библиотеки, например, cc для WTACallControl, wp для WTAPublic. Далее следует название функции <функция>, например, cs для "настройки звонка", mc для "вызова". Если требуется, далее могут следовать параметры. Это могут быть телефонные номера и т.д. Наконец, может быть возвращен один или несколько результатов. Эти результаты присваиваются переменным в контексте пользовательского агента.

В пределах карты WML идентификатор URI для звонка по определенному номеру может быть следующим:

```
wtai://wp/mc;07216086415
```

Те же функции можно использовать и в WMLScript. Здесь вызов функции использует ту же схему, что и вызов какой-либо другой функции в WMLScript.

```
<возвращаемоезначение> = имяфункции (параметры);
```

Параметр возвращаемоезначение может понадобиться, если приложению нужно это значение для дальнейших операций. Параметр имяфункции извлекается из имени библиотеки и имени функции в библиотеке. Наконец, функции можно передать одно или более значений.

Тот же пример звонка по номеру теперь может быть таким:

```
WTAPublic.makeCall("07216088415");
```

Архитектура WTA включает в себя детально разработанный механизм **обработки событий**. Клиент, т.е. пользовательский агент WTA, должен быть в состоянии обрабатывать события телефонии, которые близки к событиям реального времени. Это означает, что отреагировать нужно за строго ограниченный промежуток времени. Существует два пути поступления события клиенту. Либо операционная система мобильного устройства обнаруживает событие, либо сервер WTA поставляет событие клиенту. Первый случай означает, что традиционное событие прибывает на устройство посредством мобильной телефонной сети, а затем преобразуется в событие WTA. Во втором случае событие приходит с сервера WTA, используя службу поставки WSP.

Вне зависимости от того, откуда пришло событие, пользовательский агент должен в первую очередь сопоставить его с каким-либо существующим **связыванием события WML** в текущем контексте. Если связывание с этим событием существует, то должен быть загружен указанный URL. Это означает, что событие только инициирует загрузку новой карты. Если связывание события WML отсутствует, то событие сопоставляется с локальной **таблицей событий**. Таблица должна указывать, какие URL должны загружаться при появлении этого события. Затем пользовательский агент WTA загружает этот новый URL. Приведенный пример показывает, что события всегда отображаются в URL.

Если событие поставлено сервером WTA, сервер может дополнительно поставить содержимое, указывающее, как обрабатывается это событие. Если используется только этот метод (т.е. все события приходят с сервера WTA и мобильная сеть также посылает свои события в первую очередь на сервер), тогда события WTA используются в так называемом **режиме центрального сервера**. Если клиент обрабатывает события из мобильной сети, а сервер предоставляет, например, только обновления связываний событий, события WTA используются в так называемом **режиме центрального клиента**. Эти два противоположных режима могут объединяться в одну архитектуру, позволяя операторам выбирать режим, который наилучшим образом подходит для каждого события.

Обзор логической архитектуры WTA представлен на рис. 11.30. Показанные компоненты не являются обязательными в этой архитектуре, однако, брандмауэр или другие исходные серверы могут оказаться полезными. **Клиент** связан посредством мобильной сети с **исходным сервером WTA**, с другими телефонными сетями (например, стационарной PSTN) и **шлюзом WAP**. Здесь не показан агент пользователя WML, запускаемый на клиенте, или другие агенты пользовате-

ля. Клиент может использовать голосовую связь по сети или осуществлять обычный обмен данными. К другим исходным серверам внутри доверяемого домена можно получать доступ посредством шлюза WAP. Брандмауэр полезен при связи с посторонним исходным сервером за пределами доверяемого домена.

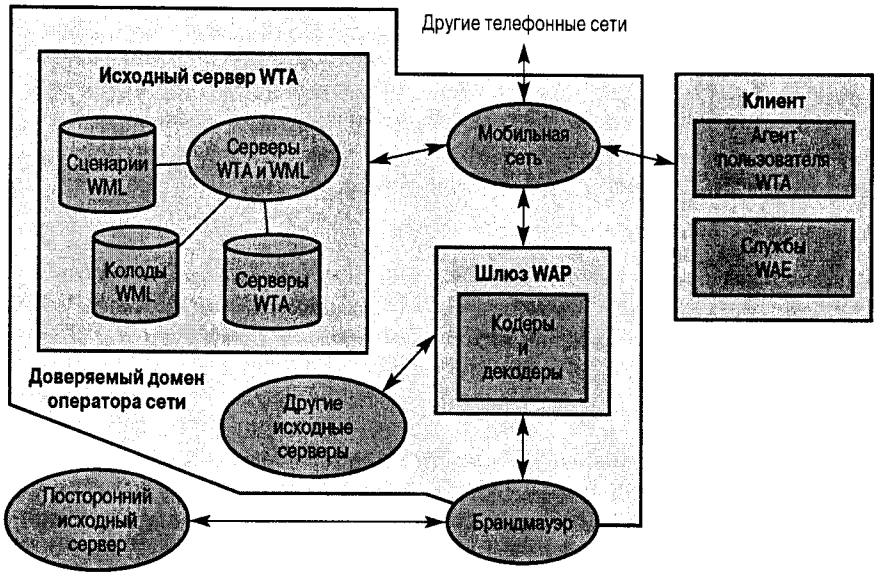

Рис. 11.30. Логическая архитектура WTA

Помимо безопасности одним из отличий исходных серверов WTA от других серверов является более строгий контроль QoS. Оператор сети точно знает задержку, надежность и пропускную способность мобильной сети и, таким образом, может лучше контролировать поведение служб. Другие серверы, которые, вероятно, находятся в Internet, могут оказаться не в состоянии гарантировать такое же хорошее QoS, как сетевой оператор. Аналогично, агент пользователя управляет содержимым очень жестко, в режиме реального времени, по сравнению со стандартным агентом пользователя WML, используемым для просмотра Web-ресурсов.

На рис. 11.31 приведен пример взаимодействия между клиентом WTA, сервером WTA, мобильной сетью (с, вероятно, большим количеством серверов) и голосовым сервером. Как показано, кто-либо может оставить сообщение на голосовом сервере. Это инициирует передачу сообщения с голосового сервера на сервер WTA, указывающего на прибытие нового сообщения. Затем сервер WTA может динамически сгенерировать новую колоду, содержащую, например, ссылки на все голосовые сообщения, которые хранятся в данный момент на голосовом сервере. Затем он поставляет эту колоду клиенту WTA.

Клиент WTA отображает новую колоду, и пользователь может выбирать, какое голосовое сообщение он хочет прослушать. Этот запрос посылается назад на сервер WTA, и клиент WTA отображает новую карту для пользователя, например, показывая сообщение, подобное "Слушайте". Сервер WTA переводит запрос определенного голосового сообщения в запрос для голосового сервера, требующий воспроизвести выбранное голосовое сообщение. Теперь этот запрос имеет формат, подходящий голосовому серверу.

Далее голосовой сервер настраивает вызов клиента WTA, используя стандартную процедуру настройки вызова в мобильной телефонной сети. Сеть указывает клиенту на вызов, и без взаимодействия с клиентом производиться автоматический ответ на вызов, поскольку клиент уже

подготовлен к этому событию. Вызов принят, и, наконец, устанавливается голосовая связь между голосовым сервером и клиентом WTA. Пунктирная линия на рисунке указывает стандартное взаимодействие мобильной сети, нехарактерное для WTA.

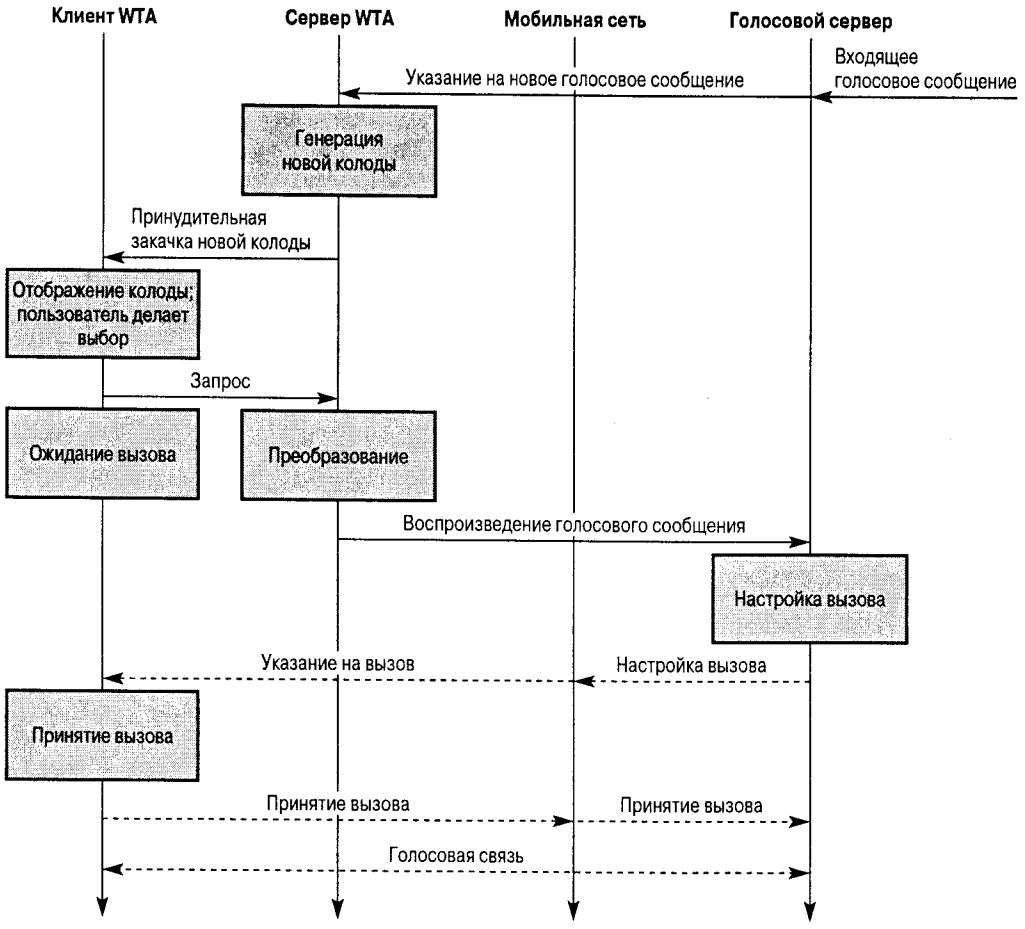

Рис. 11.31. Пример WTA: голосовое сообщение

В следующих примерах проиллюстрированы интеграция WTA, вызовов библиотеки WTAI, WML и WMLScript и то, как разработчики могут использовать функции WTA [WAP Forum, 1998м]. Представим, что по телевизору транслируется некоторое шоу. В конце шоу можно проголосовать за одного из участников. Традиционный способ — каждый кандидат объявляет соотнесенный с ним телефонный номер и для голосования необходимо его набрать (такой метод до некоторой степени подвержен ошибкам, если используется длинный номер). Используя WAP с WTA, оператор сети может поставить портативному устройству колоду из нескольких карт и предоставить простой выбор для голосования, как объясняется в следующих абзацах.

Первый пример состоит только из WML и включает две карты. Первая карта предлагает проголосовать. В случае согласия будет отображена вторая карта. На этой карте с именем `voteChamp` на роль победителя можно выбрать Микки, Дональда и Плуто. Не нужно набирать номер, выбрать можно непосредственно имя выбранного кандидата. Переменной `voteNo` будет присвоено значение (т.е. телефонный номер), связанное с выбранным претендентом. Отметим,

что URL используется после того, как пользователь сделает выбор. URL указывает систему использовать функцию sc (настройка вызова) из библиотеки cc (управление вызовом) в WTAI. Эта функция вызывает определенный номер (здесь — voteNo) и хранит запрос (указывается параметром "1") даже после удаления текущего контекста.

В целом, эта колода WML позволяет пользователю выбирать победителя и затем автоматически звонить по верному номеру. Пользователь не должен беспокоиться о том, какой номер является правильным. Общей проблемой WML является отсутствие динамического поведения. Что произойдет, если что-то пойдет не так с вызовом, например, номер будет занят? Для выполнения проверки на наличие ошибок и сообщении о них пользователю в дополнение к WML должен использоваться WMLScript.

```
<WML>
  <CARD>
    <DO TYPE="ACCEPT" TASK="GO"URL="#voteChamp"/>
    Проголосуйте за вашего победителя!
  </CARD>
  <CARD NAME="voteChemp">
    <DO TYPE="ACCEPT" TASK="GO"
       URL="wtai://cc/sc;$voteNo;1"/>
    Претенденты:
    <SELECT KEY="voteNo">
      <OPTION VALUE="6086415">Mickey</OPTION>
      <OPTION VALUE="6086416">Donald</OPTION>
      <OPTION VALUE="6086417">Pluto</OPTION>
    </SELECT>
  </CARD>
</WML>
```

Ниже приведен тот же пример, голосование за победителя, но с использованием WMLScript и WML для улучшения обработки ошибок и сообщения о них. Как и ранее, оператор сети может поставлять этот код в портативное устройство. Прежде всего, определяется функция voteCall. Данная функция как вход принимает номер, а затем, во второй строке, устанавливает вызов этого номера. Теперь для настройки вызова используется библиотечная функция WTAControl.setup (не идентификатор URL, как в предыдущем примере). Этой функции передаются два значения — номер телефона для набора и значение, предписывающее сохранить вызов после удаления текущего контекста. Преимуществом использования описанной функции, а не URL, является возможность простой обработки возвращенного значения, здесь сохраненного в переменной j.

В зависимости от значения j следующие строки могут подготовить сообщение, которое WML-броузер сможет отобразить для пользователя. Функция настройки вызова определена таким образом, чтобы возвращенные отрицательные значения обозначались как ошибки. Само значение представляет код ошибки WTAI. Предопределенными кодами ошибки являются, например, –5 для сообщения "вызываемая сторона занята", –6 для сообщения "сеть занята" или –7 для сообщения "нет ответа" (т.е. настройка вызова временно прерывается). Если возвращенное значение не отрицательно, переменной броузера Message присваивается строка "Called", переменной No — значение Nr, т.е. вызываемый номер. В противном случае происходит ошибка, и Message присваивается строка "Error", а No присваивается код ошибки, сохраненный в j.

Далее следует колода WML, подобная приведенной в первом примере, но с некоторыми важными отличиями. Как и ранее, первая колода отображает некоторый текст. После согласия на голосование WML-броузер отображает имена трех кандидатов. Пользователь может сделать выбор, и телефонный номер, соотнесенный с кандидатом, хранится в переменной `voteNo`. В отличие от предыдущего примера не загружаются никакие URL для WTAI, а броузер загружает URL, указывающий на сценарий WMLScript. Сценарий WMLScript размещен в единице компиляции `script`; имя сценария — `voteCall`. Затем этой функции передается значение переменной `voteNo`.

Важно отметить, что управление теперь находится в сценарии. Вторая карта не передает управление третьей карте. В нашем примере третья карта, которая должна отобразить сообщение, вызывается функцией WMLScript со строкой `WMLBrowser.go ("showResult")`. Это загружает карту `showResult` в броузер WML. Данная третья карта отображает некоторый текст и значения переменных `Message` и `No`. Эти значения были присвоены ранее, в сценарии WMLScript.

```
Function voteCall(Nr) {
  var j = WTACallControl.setup(Nr,1);
  if (j>=0) {
    WMLBrowser.setVar("Message","Called");
    WMLBrowser.setVar("No", Nr);
  }
  else {
    WMLBrowser.setVar("Message","Error");
    WMLBrowser.setVar("No", j);
  }
  WMLBrowser.go("showResult");
}

<WML>
  <CARD>

    <DO TYPE="ACCEPT" TASK="GO" URL="#voteChamp"/>
    Проголосуйте за вашего победителя!
  </CARD>
  <CARD NAME="voteChemp">
    <DO TYPE="ACCEPT" TASK="GO"
       URL="/script#voteCall($voteNo)"/>
    Претенденты:
    <SELECT KEY="voteNo">
      <OPTION VALUE="6086415">Mickey</OPTION>
      <OPTION VALUE="6086416">Donald</OPTION>
      <OPTION VALUE="6086417">Pluto</OPTION>
    </SELECT>
  </CARD>
  <CARD NAME="showResult">
    Состояние вызова: $Message $No
  </CARD>
</WML>
```

В этом очень простом примере показано взаимодействие WML и WMLSrcipt вместе с WTAI. Конечно, коды ошибки или телефонные номера не должны отображаться у заказчика, эти коды и номера должны преобразовываться в открытый текст с помощью WMLScript.

Хотя WTAI приемлем для множества различных мобильных сетей, WAP Forum установил несколько дополнений WTAI, действительных только для специфических сетей. В настоящее время дополнения определены для Тихоокеанской цифровой сотовой системы (PDC, [WAP Forum, 1998н]), Глобальной системы мобильной связи (GSM, [WAP Forum, 1998о]) и сотовой сети IS-136 TDMA [WAP Forum, 1998п]. Типичной функцией GSM является, например, участие в групповых вызовах. Таким образом, дополнение WTAI для GSM в своей библиотеке имеет функцию `WTAGSM.multiparty` для WMLScript и добавляет URI `wtai://gsm.jm`.

11.3.10. Примеры стеков WAP

После представления различных аспектов WAP в данном последнем разделе приводится некоторая дополнительная информация о работах по стандартизации с использованием примерных конфигураций, показанных на рис. 11.32. WAP пытается, насколько это возможно, использовать существующие в Internet технологии и философию. Таким образом, простейший стек протоколов (стек 3) не требует новых протоколов или реализаций. Если какому-либо приложению требуется только ненадежная служба передачи дейтаграмм без обеспечения безопасности, WAP предлагает возможность использования UDP, если сеть-носитель предоставляет службы IP (это справедливо, например, для GPRS). Основываясь на этом очень простом стеке, можно конфигурировать все более и более сложные стеки, добавляя безопасность с WTLS или надежные службы транзакций с WTP. Приложения для распределенных вычислений, такие как CORBA, могут использовать надежные службы передачи данных. В настоящее время эти приложения по большей части используют TCP. Однако, TCP может оказаться далеко не лучшим выбором в беспроводной мобильной среде, как это показано в главе 10. Альтернативное решение может обеспечить WAP.

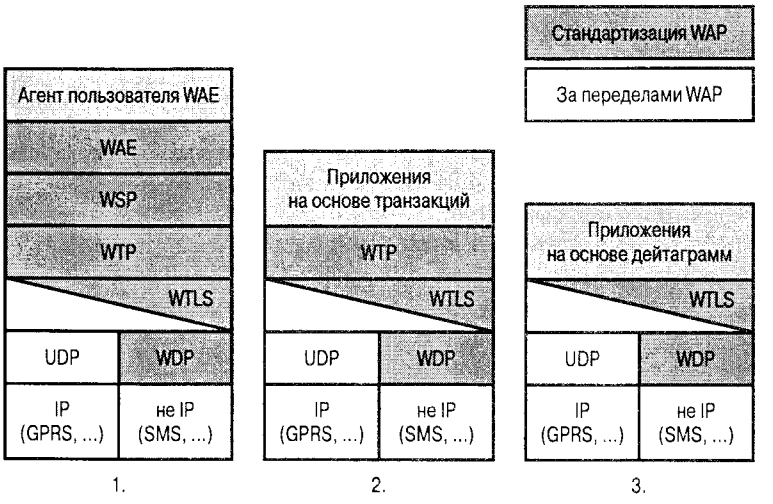

Рис. 11.32. Пример стека протоколов согласно WAP

Типичному приложению WAP, т.е. пользовательскому агенту WAP, такому как агент WML или WTA, может потребоваться полный стек протоколов, как показано в стеке 1. Пользовательский агент запускает WAE и полагается, например, на службу поставки WSP для предоставления клиенту событий WTA с сервера WTA.

11.4. Резюме

WWW — это приложение Internet, привлекающее все возрастающее количество людей. Как показано в главах 1 и 4, по всему миру растет количество абонентов мобильной беспроводной связи. Таким образом, совершенно логично объединить две сферы — WWW и мобильную связь. В этой главе рассматриваются несколько проблем, к которым приводит сочетание этих двух областей, например, низкая доступная пропускная способность в сочетании с неэффективными протоколами или HTML в качестве языка описания, не адаптированного к требованиям портативных устройств.

Основная часть главы посвящена WAP — контуру, состоящему из нескольких уровней связи, языков разметки и подготовки сценариев и соединения с телефонной сетью. Этот контур все еще находится в стадии развития. В то же время, многие компании и организации поддерживают такой подход. Важными вопросами для всех архитектур и исследовательских проектов, поддерживающих мобильный доступ к Web, являются предварительная выборка и кэширование содержимого [Jaing, 1998], преобразование содержимого и адаптация к возможностям беспроводных устройств [Brewer, 1998], [Fox, 1998], [Han, 1998] и использование броузера во время разъединения [Mazer, 1998].

Следующим этапом в поддержке мобильности является предоставление завершенных структур программного обеспечения для развития новых мобильных приложений. Примерами являются набор инструментальных средств Rover [Joseph, 1997] и MobiWare [Angin, 1998]. Rover — это инструмент для разработки оптимизированных под мобильные устройства приложений, надежных при сбоях. WobiWare — это мобильная промежуточная среда, использующая CORBA и Java. Этот инструмент поддерживает создание адаптированных к мобильным сетям приложений с помощью открытых программируемых интерфейсов.

11.5. Упражнения

1. Почему сильная согласованность файловой системы является проблемой в беспроводной и мобильной среде? Какие есть альтернативы?
2. Как обычная файловая система реагирует на разъединение? Попробуйте отключить компьютер, который был смонтирован в файловую систему через сеть.
3. Каковы преимущества неизменности состояния HTTP? В каких ситуациях полезны состояния, и как они обеспечиваются в настоящее время? Где хранятся длительные состояния, где краткосрочные?
4. Какие свойства HTTP неэкономно расходуют пропускную способность? Какие дополнительные проблемы возникают при использовании HTTP/1.0 совместно с TCP?
5. Как кэширование уменьшает время доступа и снижает требуемую пропускную способность? Где может размещаться кэш и каковы специфические преимущества каждого местоположения?
6. Какие проблемы связаны с кэшированием? Какой тип содержимого может быть кэширован, какое содержимое вызывает проблемы? К каким дополнительным проблемам приводит мобильность клиента?
7. Какие существуют противоречия между возможностями HTML и реалиями беспроводных портативных устройств? Какие решения предлагаются? Какова роль дополнительных модулей и как они влияют на удобство использования Web-страниц?
8. Назовите механизмы улучшения доступа к Web-ресурсам для портативных устройств. Какая общая проблема характерна для них и что привело, в конечном счете, к развитию WAP?

9. Каковы типичные усовершенствования базовой архитектуры Web клиент/сервер? Пересмотрите эти модернизации для мобильных беспроводных пользователей с доступом к Web по мобильной телефонной сети. Где эффективнее провести модернизацию?
10. Каковы основные цели работ WAP Forum и как они отразились на протокольной архитектуре WAP?
11. Какие пути перемещений предлагает в своих протоколах WAP для приложений Internet и телефонии, а также для их протоколов?
12. Является ли WDP стационарным протоколом и почему WAP не определил SAP, которую может использовать WDP?
13. Почему WAP определяет свой собственный уровень безопасности, а не полагается на безопасность, которую обеспечивает мобильная телефонная сеть?
14. Назовите преимущества и недостатки использования пользовательского подтверждения приема в WTP. Какие приложения типичны для обоих случаев?
15. Какой класс WTP наилучшим образом отражает типичную схему доступа к Web? Как избежать ненужных служебных издержек при использовании WSP над уровнем этого класса при просмотре Web-ресурсов?
16. Какие проблемы HTTP может решить WSP? Почему эти решения особенно необходимы в мобильных беспроводных средах?
17. Почему WSP/B не располагает отклики в том же порядке, что и запросы? Подумайте, например, о запросах к различным объектам Web-страницы.
18. Какие преимущества предлагает служба сеанса связи без установления соединения по сравнению с простой службой передачи дейтаграмм?
19. Какие усовершенствования в классическую модель Web клиент/сервер вводит WAE? Каковы функции этой модернизации?
20. Каково фундаментальное различие между WML и HTML? Почему это различие важно при рассмотрении портативных устройств? Что еще задается, помимо сохранения пропускной способности?
21. Почему к WML был добавлен язык подготовки сценариев? Как этот язык может помочь сэкономить пропускную способность и снизить задержку?
22. Какие существуют типичные события телефонии и как они интегрированы в WAP? Каким образом пользователь, используя Web-броузер, может получить доступ к функциям мобильных телефонов?
23. Какова роль сервера WTA? Какие существуют способы интеграции сервера WTA в архитектуру WAP?

11.6. Рекомендуемая литература

Adobe (1999) `http://www.adobe.com/`.

Angin, O.; Cambell, A. T.; Kounavis, M. E.; Liao, R. R. E. (1998) "The MobiWare toolkit: Programmable support for adaptive mobile networking", *IEEE Personal Communications*, 5, (4).

Berners-Lee, T. (1994a) *Universal Resource Identifiers in WWW, a unifying syntax for the expression of names and addresses of objects on the network as used in the world wide Web*, RFC 1630.

Berners-Lee, T. (1994) *Uniform Resource Locators (URL)*, RFC 1738

Berners-Lee, T.; Fielding, R.; Frystyk, H. (1996) *Hypertext Transfer Protocol — HTTP/1.0*, RFC 1945.

Bickmore, T.; Schilit, B. (1997) "Digestor: device independent access to the world wide Web", proc. 6th International World Wide Web Conference, Santa Clara, CA, US.

Blue Squirrel Corporation (1999) *WebWhacker v3.2*, `http://www.bluesquirrel.com./`.

Bluetooth (1999) `http://www/bluetooth.com/`.

Brewer, E. A.; Katz, R. H.; Chawathe, Y.; Gribble, S. D.; H.; Hodes, T.; Nguyen, G.; Stemm, M.; Henderson, T.; Amit, E.; Balakrishnen, H.; Fox, A.; Padmanabhan, V.; Seshan, S. (1998) "A network architecture for heterogeneous mobile computing", *IEEE Personal Communications*, 5, (5).

Certicom Corporation (1999) `http://www.certicom.com/`.

Conta. A.; Deering, S. (1999) *Internet Control Massage Protocol (ICMPv6) for Internet Protocol Version 6*, RFC 1885.

Dierks, T.; Allen, C. (1999) *The TLS protocol version 1.0*, RFC 2246.

Diffie, W.; Hellman, M. (1976) "New directions in cryptography", *IEEE Transactions on Information Theory*, 22, (6).

European Telecommunications Standards Institute (1999). `http://www.etsi.org/`.

Fielding, R. (1995) *Relative Uniform Resource Locators*, RFC 1808.

Fielding, R.; Gettys, J.; Mogul, J.; Frystyk, H.,; Berners-Lee, T. (1997) *Hypertext Transfer Protocol — HTTP/1.1*, RFC 2068.

Flanagan, D. (1997) *JavaScript: the definitive guide*. O'Reilly.

Floyd, R.; Housel, B.; Tait, C. (1998) "Mobile Web access using eNetwork Web Express", *IEEE Personal Communications*, (5).

Fox, A.; Brewer, E. (1996) "Reducing WWW latency and bandwidth requirements by real-time distillation", proc. 5th International World Wide Web Conference, Paris, France.

Fox. A.; Gribble, S.; Brewer, E. A.; Amir, E. (1998) "Adapting to network and client variability via on-demand dynamic distillation", proc. ASPLOS'96, Cambridge, MA, USA.

Fox, A.; Gribble, S.D.; Chawathe, Y.; Brewer, E. A. (1998) "Adapting to network and client variation using infrastructure proxies: Lessons and perspectives", *IEEE Personal Communications*, 5, (4).

Guedes, V.; Moura, F. (1995) "Replica Control in Mio-NFS", proceedings of the ECOOP"95 Workshop on Mobility and Replication, Aarhus, Denmark, Auguest.

Han, R.; Bhagwat, P.; LaMaire, R.; Mummert, T.; Perret, V.; Rubas, J. (1998) "Dynamic adaptation in an image transcoding proxy for mobile Web browsing", *IEEE Personal Communications*, 5, (6).

Hiedemann, J. S.; Page, T. W.; Guy, R. G.; Popeck, G. J. (1992) "Primarily disconnected operation: experiences with Ficus", proceedings of the 2nd Workshop on the Management of Replicated Data, Monterey, IEEE Computer Society Press, CA, USA, November.

Honeyman, P.; Huston, L. B. (1995) "Communications and consistency in mobile file systems", *IEEE Personal Communications 2(6)*, December.

Housel, B.;Lindquist, D. (1996) "WebExpress: A system for optimizing Web browsing in a wireless environment", proc. ACM/IEEE MobiCom'96 conference, Rye, NY, USA.

Howard, J. H.; Kazar, M. L.; Menees, S. G.; Nichols, D. A.; Satyanarayanan, M.; Sidebothman, R. N., West, M. J. (1988) "Scale and performance in a distributed file system", *ACM Transactions on Computer Systems*, 6(1), February.

Huston, L.B.; Honeyman, P. (1993) "Disconnected operation for AFS", proceedings of USENIX Symposium on Mobile and Location-Independent Computing, Cambridge, MA, USA, August.

Internet Engineering Task Force (1999) `http://www.ietf.org/`.

Internet Mail Consortium (1996) *vCard — the electronic business card*, Internet Mail Consortium.

Internet Mail Consortium (1996) *vCalendar — the electronic calendaring and scheduling format*, Internet Mail Consortium.

Jiang, Z.; Kleinrock, L. (1998) "Web prefetching in a mobile environment", *IEEE Personal Commuvications*, vol. 5, no. 5.

Joseph, A. D.; Kaashoek, M. F. (1997) "Building reliable mobile-aware applications using the Rover toolkit", *Wireless Networks*, J. C. Baltzer, 3.

Kaufman, C.; Perlman, R.; Speciner, M. (1995) *Network security — private communication in a public world*, Prentice Hall.

King, P.; Hyland, T. (1997) *Handheld device markup language specification*. Unwired Planet.

Kistler, J. J.; Satyanarayanan, M. (1992) "Disconnected operation in the Coda file system", *ACM Transactions on Computer Systems*, 10(1), February.

Kristol, D.; Montulli, L. (1997) *HTTP state management mechanism*, RFC 2109.

Kumar, P.; Satyanarayanan, M. (1993) "Spporting application-specific resolution in an optimistically replicated file system", proceedings of 4th Workshop on Workstation Operation Systems, Napa, CA USA, October.

Liljeberg, M.; Alanko, T., Kojo, M.; Laamanen, H.; Raatikainen, K. (1995) "Optimizing world wide Web for weakly connected mobile workstations: An indirect approach", proc. 2nd International Workshop on Services in Distributed and Networked Environments, SDNE'95, Whistler, B. C., Canada.

Liljeberg, M.; Helin, H.; Kojo, M.; Raatikainen, K.: Mowgli WWW (1996) "Improved usability of WWW in mobile WAN environments", proc. IEEE Global Internet 1996, London, England.

Lotus Corporation (1999) *Weblicator 1.03*, `http://www.lotus.com/`.

LoVerso, J.; Mazer, M. (1997) "CaubWeb: detaching the Web with Tcl", proc. 5th Annual USENIX Tcl/Tk Workshop, Boston, MA, USA.

Lu, Q.; Satyanarayanan, M. (1994) "Isolation-only transactions for mobile computing", *Operation Systems Review*, 28(2), April.

Mazer, M. S.; Brooks, C. L. (1998) "Writing the Web while disconnected", *IEEE Personal Communications*, 5, (5).

Microsoft Corporation (1999) *Microsoft Internet Explorer*, `http://www.microsoft.com/`.

Mummert, L. B.; Ebling, M. R.; Satyanarayanan, M. (1995) "Exploiting weak connectivity for mobile file access", proceedings of the 15th Symposium on Operating System Principles, Copper Mountain Resort, CO USA, December.

Netscape Corporation (1999) *Netscape Communicator*, `http://www.netscape.com/`.

Popek, G. J.; Guy, R. G.; Page, T. W. (1990) "Replication in the Ficus distributed file system", proceedings of the Workshop on the Management of Replicated Data, Los Alamitos, IEEE Computer Society Press, CA USA, November.

Postel, J. (1980) *User Datagram Protocol*, RFC 768.

Postel, J. (1981a) *Internet Protocol*, RFC 791.

Postel, J.(1981б) *Internet Control Message Protocol*, RFC 792.

Raggett, D.; LeHors, A.; Jacobs, I. (1998) *HTML 4.0 specification*, W3C recommendation, REC-html40-1998024.

Rivest, R.; Shamir, A.; Adleman, L. (1978) "A method for obtaining digital signatures and public-key cryptosystems", *Communications of the ACM*, 21, (2).

Satyanarayanan, M.; Kistler, J. J.; Mummert, L. B.; Ebling, M. R.; Kumar, P.; Lu, Q. (1993) "Experiences with disconnected operation in a mobile computing environment", proceeding of USENIX Symposium on Mobile and Location-Independent Computing, Cambridge, MA, USA, August.

Schilit, B.; Douglis, F.; Kristol, D. M.; Krzyzanowski, P.; Sienicki, J.; Trotter, J.A. (1996) "TeleWeb: loosely connected access to the world wide Web", proc. 5th International World Wide Web Conference, Paris, France.

Schneier, B. (1996) *Applied cryptography second edition: Protocol, Algorithms and source code in C*, John Wiley & Sons.

World Wide Web Consortium (1998) *Extensible Markup Language (XML) 1.0 Specification*, W3C recommendation, REC-xml-19980210.

World Wide Web Consortium (1999) `http://www.w3c.org/`.

WAP Forum (1998а) *Wireless application protocol architecture spcification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998б) *Wireless datagram protocol specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998в) *Wireless transport layer security protocol*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998г) *Wireless transaction protocol specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998д) *Wireless session protocol specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998е) *Wireless application environment specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998ж) *Wireless application environment overview*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998з) *WMLscript language specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998и) *WMLscript standard libraries specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998к) *Wireless markup language specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998л) *Binary XML content format specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998м) *Wireless telephony application specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998н) *Wireless telephony application interface specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998о) *Wireless telephony application interface specification*, PDC specific addendum, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998п) *Wireless telephony application interface specification*, GSM specific addendum, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998р) *Wireless telephony application interface specification*, IS-136 specific addendum, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998с) *WAP over GSM USSD specification*, WAP Forum, http://www.wapforum.org/, April.
WAP Forum (1998т) *Wireless control message protocol specification*, WAP Forum, http://www.wapforum.org/, April.
Wireless Application Protocol Forum (1998s) http://www.wapforum.org/.

Сокращения

AAL	(ATM Adaptation Layer) Уровень адаптации ATM
ABR	(Available BitRate) Доступная скорость передачи
ACF	(Association Control Function) Функция управления ассоциациями
ACK	(ACKnowledgement) Подтверждение приема
ACL	(Asynchronous Connectionless Link) Асинхронный канал связи без установления соединения
ACT	(Ad-hoc Controller Terminal) Терминал эпизодического контроллера
ADA	(Alias DA) Альтернативный адрес назначения
ADSL	(Asymmetric Digital Subscriber Line) Асимметричная цифровая абонентская линия
AESA	(ATM End System Address) Адрес конечной системы ATM
AFS	(Andrew File System) "Файловая система Эндрю"
AGCH	(Access Grant CHannel) Канал предоставления доступа
AIB	(Alias Information Base) Информационная база альтернативных имен
AID	(Acknowledgement Identifier) Идентификатор подтверждения приема
AIDCS	(AID CheckSum) Контрольная сумма AID
AK-HCPDU	(AcKnowledgement HCPDU) HCPDU подтверждения приема
AM	(Amplitude Modulation) Амплитудная модуляция
AMES	(ATM Mobility Extension Service) Служба ATM с поддержкой мобильности
AMPS	(Advanced Mobile Phone System) Усовершенствованная система мобильной радиотелефонной связи
ANSI	(American National Standards Institute) Национальный институт стандартизации США
AP	(Access Point) Точка доступа
APCF	(AP Control Function) Функция управления точкой доступа
APCM	(AP Connection Management) Управление соединением точки доступа
APCP	(AP Control Protocol) Протокол управления AP
ARIB	(Association of Radio Industries and Broadcasting) Ассоциация радиоиндустрии и широковещания
ARQ	(Automatic Repeat reQuest) Автоматический запрос повторной передачи
ARQN	(ARQ sequence Number) Порядковый номер ARQ
ASA	(Alias SA) Альтернативный адрес источника
ASCII	(American Standard Code for Information Interchange) Американский стандартный код для обмена информацией
ASK	(Amplitude Shift Keying) Амплитудная манипуляция
ASP	(Active Server Page) Активная страница сервера
Assoc	(Association) Ассоциация
ATIM	(Ad-hoc TIM) Карта индикации эпизодических транспортных потоков
ATM	(Asynchronous Transfer Mode) Асинхронный режим передачи

ATMC (ATM Connection function) Функция связи ATM
AuC (Authentication Centre) Центр аутентификации
AUS (Authentication Server) Сервер аутентификации
Auth (Authentication) Аутентификация
BCA (Borrowing Channel Allocation) Распределение каналов с заимствованием
BCCH (Broadcast CCH) Широковещательный канал управления
BCH (Bose-Chaudhuri-Hocquenghem) Код Боуза-Чоудхури-Хоквенгема
BER (Bit Error Rate) Частота появления ошибочных битов
BFSK (Binary FSK) Двоичная FSK
B-ISDN (Broadband ISDN) Широкополосная сеть ISDN
BLI (Block Length Indicator) Индикатор длины блока
BLIR (Block Length Indicator Replica) Копия индикатора длины блока
BLIRCS (BLIR CheckSum) Контрольная сумма BLIR
BMP (BitMap) Битовое отображение
BPSK (Binary PSK) Двоичная PSK
BRAN (Broadband Radio Access Network) Широкополосная сеть радиодоступа
BSC (Base Station Controller) Контроллер базовых станций
BSS (Base Station (Sub)system) (Под)система базовых станций
BSS (Basic Service Set) Базисный набор служб
BSSAP (BSS Application Part) Прикладная часть BSS
BSSGP (BSS GPRS Protocol) Протокол (под)системы базовых станций GPRS
BTS (Base Transceiver Station) Базовая трансиверная станция
BTSM (BTS Management) Управление BTS
BW (BackWard) Обратный
CAC (Channel Access Control) Управление доступом к каналу
CAC (Connection Admission Control) Управление доступом соединения
CAMEL (Customized Application for Mobile Enhanced Logic) Индивидуальные приложения с мобильной расширенной логикой
CATV (Community Antenna Television) Абонентское телевидение
CBR (Constant BitRate) Постоянная скорость передачи
CC (Call Control) Управление вызовами
CC (Country Code) Код страны
CCA (Clear Channel Assessment) Оценка незанятости канала
CCCH (Common CCH) Общий канал CCH
CCF (Call control and Connection Function) Функция управления вызовом и соединением
CCH (Control Channel) Канал управления
CCIR (Consultative Committee for International Radiocommunication) Международный консультативный комитет по радиовещанию
CD (Compact Disc) Компакт-диск
CDM (Code Division Multiplexing) Уплотнение с кодовым разделением
CDMA (Code Division Multiplexing Access) Множественный доступ с кодовым разделением каналов
CDPD (Cellular Digital Packet Data) Сотовая система передачи пакетов цифровых данных
CDV (Cell Delay Variation) Разброс задержек в ячейке
CEPT (European Conference for Posts and Telecommunications) Европейская конференция почтовых и телекоммуникационных ведомств
CGI (Common Gateway Interface) Общий шлюзовый интерфейс
CIDR (Classless InterDomain Routing) Бесклассовая междоменная маршрутизация
CIF (Common Interleaved Frame) Общий чередующийся кадр

CKSN	(Ciphering Key Sequence Number) Порядковый номер ключа шифрования
CLMS	(Connectionless Message Service) Служба сообщений без установления соединения
CM	(Call Management) Управление соединениями
CN	(Core Network) Внутренняя сеть
CN	(Correspondent Node) Корреспондентский узел
Cnf	(Confirmation) Подтверждение
COA	(Care-Of Address) Адрес передачи
Codec	(Coder/Decoder) Кодер/декодер, кодек
COFDM	(Coded OFDM) Кодированная OFDM
COMS	(Connection Oriented Message Service) Служба передачи сообщений на основе соединений
CORBA	(Common Object Request Broker Architecture) Обобщенная архитектура посредника объектных запросов
COS	(Cross Over Switch) Разделительный коммутатор
CPM	(Continuous Phase Modulation) Модуляция без разрыва фазы
CPU	(Central Processing Unit) Центральный процессор
CRC	(Cyclic Redundancy Check) Циклическая проверка четности с избыточностью
CS	(CheckSum) Контрольная сумма
CSCW	(Computer Supported Cooperative Work) Совместная работа на базе компьютеров
CSMA	(Carrier Sense Multiple Access) Множественный доступ к среде с детектированием несущей
CSMA/CA	(CSMA with Collision Avoidance) CSMA с предотвращением конфликтов
CSMA/CD	(CSMA with Collision Detection) CSMA с обнаружением конфликтов
CT	(Cordless Telephone) Радиотелефон
CTS	(Clear To Send) Готов к передаче
CU	(Capacity Unit) Модуль емкости
CVSD	(Continuous Variable Slope Delta) Непрерывная дельта-модуляция с переменной огибающей
CW	(Contention Window) Окно состязания
DA	(Destination Address) Адрес назначения
DAB	(Digital Audio Broadcasting) Цифровое аудиовещание
DAMA	(Demand Assigned Multiple Access) Множественный доступ с распределением по запросу
D-AMPS	(Digital-AMPS) Цифровая AMPS
DBPSK	(Differential Binary PSK) Дифференциальная двоичная фазовая манипуляция
DC	(Direct Current) Постоянный ток
DCA	(Dynamic Channel Allocation) Динамическое распределение каналов
DCCH	(Dedicated CCH) Выделенный канал управления
DCF	(Distributed Coordination Function) Распределенная координационная функция
DCS	(Digital Cellular System) Цифровая сотовая система
DDIB	(Duplicate Detection Information Base) Информационная база обнаружения копий
DECT	(Digital Enhanced Cordless Telecommunications) Цифровые расширенные беспроводные телекоммуникации
DFWMAC	(Distributed Foundation Wireless MAC) Распределенное беспроводное управление доступом к среде
DH	(Diffie, Hellman) Диффи, Хеллман
DHCP	(Dynamic Host Configuration Protocol) Протокол динамической конфигурации хоста
DIFS	(DCF IFS) Распределенная координационная функция IFS

DLC	(Data Link Control) Управление каналом передачи данных
DNS	(Domain Name System) Служба имен доменов
DPCCH	(Dedicated Physical Control Channel) Выделенный физический канал управления
DPCH	(Dedicated Physical CHannel) Выделенный физический канал
DQPSK	(Differential QPSK) Дифференциальная QPSLK
DS	(Distribution System) Распределительная система
DSDV	(Destination Sequence Distance Vector) Дистанционно-векторная маршрутизация с нумерацией и оценкой пути к пункту назначения
DSL	(Digital Subscriber Loop) Цифровой абонентский шлейф
DSMA	(Digital Sense Multiple Access) Множественный доступ с цифровым детектированием
DSSS	(Direct Sequence Spread Spectrum) Спектр, расширенный методом прямой последовательности
DT-HCPDU	(DaTa-HCPDU) Информационный HCPDU
DTIM	(Delivery TIM) Карта индикации целевого транспортного потока
DTMF	(Dual Tone Multiple Frequency) Двухтональный многочастотный набор
DVB	(Digital Video Broadcasting) Цифровое видеовещание
DVB-C	(DVB-Cable) Кабельное DVB
DVB-S	(DVB-Satellite) Спутниковое DVB
DVB-T	(DVB-Terrestrial) Наземное DVB
DVD	(Digital Versatile Disk) Универсальный цифровой диск
DVTR	(Digital Video Tape Recorder) Цифровой видеомагнитофон
EDGE	(Enhanced Data rates for GSM Evolution) Ускоренная передача данных для развития GSM
EDTV	(Enhanced Definition TV) Телевидение улучшенной четкости
EHF	(Extremely High Frequency) Крайне высокая частота
EIR	(Equipment Identity Register) Регистр опознавания оборудования
EIRP	(Equivalent Isotropically Radiated Power) Эквивалентная изотропно-излучаемая мощность
EIT	(Event Information Table) Таблица информации о событиях
EMAS	(End-user Mobility-supporting ATM Switch) Коммутатор ATM, поддерживающий мобильность конечного пользователя
EMAS-E	(EMAS-Entry) Объект EMAS
EMAS-N	(EMAS-Network) Сеть EMAS
ESS	(Extended Service Set) Расширенный набор служб
ETSI	(European Telecommunications Standards Institute) Европейский институт стандартов по телекоммуникациям
EY-NPMA	(Elimination-Yield Non-preemptive Priority Multiple Access) Исключение, порождающее приоритетный множественный доступ
FA	(Foreign Agent) Чужой агент
FACCH	(Fast Associated Dedicated CCH) Быстрый ассоциированный выделенный канал управления
FCA	(Fixed Channel Allocation) Стационарное распределение каналов
FCCH	(Frequency Correction CHannel) Канал коррекции частоты
FDD	(Frequency Division Duplex) Дуплекс с частотным разделением каналов
FDM	(Frequency Division Multiplexing) Уплотнение с частотным разделением
FDMA	(Frequency Division Multiple Access) Множественный доступ с частотным разделением
FEC	(Forward Error Correction) Прямое исправление ошибок

FHSS	(Frequency Hopping Spread Spectrum) Расширенный спектр со скачкообразной перестройкой частоты
FIB	(Fast Information Block) Блок быстрой информации
FIC	(Fast Information Channel) Канал быстрой информации
FM	(Frequency Modulation) Частотная модуляция
FPLMTS	(Future Public Land Mobile Telecommunication System) Система будущих мобильных телекоммуникаций общего пользования
FR	(Frame Relay) Ретрансляция кадров
FSK	(Frequency Shift Keying) Частотная манипуляция
FT	(Fixed Radio Termination) Стационарное оконечное радиоустройство
GEO	(Geostationary (or Geosynchronous) Earth Orbit) Геостационарная (или геосинхронная) околоземная орбита
GFSK	(Gaussian FSK) Гауссова FSK
GGSN	(Gateway GSN) Шлюзовый GSN
GIF	(Graphics Interchange Format) Формат графического обмена
GMM	(Global Multimedia Mobility) Глобальная мультимедийная мобильность
GMSC	(Gateway MSC) Шлюз MSC
GMSK	(Gaussian MSK) Гауссова MSK
GPRS	(General Packet Radio Service) Общая пакетная радиослужба
GPS	(Global Positioning System) Глобальная система навигации и определения положения
GR	(GPRS Register) Регистр GPRS
GRE	(Generic Routing Encapsulation) Общая маршрутная инкапсуляция
GSM	(Groupe Spéciale Mobile, Global System for Mobile Communications) Глобальная система мобильной связи
GSN	(GPRS Support Node) Узел поддержки GPRS
GTP	(GPRS Tunnelling Protocol) Протокол туннелирования GPRS
GWL	(GateWay Link) Шлюзовый канал
HA	(Home Agent) Базовый агент
HBR	(High Bit-Rate) Высокая скорость передачи данных
HC	(HIPERLAN CAC) Управление доступом к каналу HIPERLAN
HCPDU	(HIPERLAN CAC PDU) PDU управления доступом к каналу HIPERLAN
HCQoS	(HIPERLAN CAC QoS) Качество обслуживания HIPERLAN CAC
HCSAP	(HIPERLAN CAC SAP) Точка доступа к службе HIPERLAN CAC
HCSDU	(HIPERLAN CAC SDU) Модуль данных службы уровня CAC сети HIPERLAN
HDA	(Hashed Destination HCSAP Address) Хэшированный адрес назначения HCSAP
HDACS	(HDA CheckSum) Контрольная сумма HDA
HDB	(Home Data Base) База данных об исходном положении
HDLC	(High level Data Link Control) Высокоуровневый протокол управления каналом
HDML	(Handheld Device Markup Language) Язык разметки портативных устройств
HDTP	(Handheld Device Transport Protocol) Транспортный протокол портативных устройств
HDTV	(High Definition TV) Телевидение высокой четкости
HEC	(Header Error Check) Проверка заголовка на наличие ошибок
HEO	(Highly Elliptical Orbit) Сильно эллиптическая орбита
HF	(High Frequency) Высокая частота
HI	(HBR-part Indicator) Индикатор компонента HBR
HIB	(Hello Information Base) Информационная база приветствий
HID	(HIPERLAN Identifier) Идентификатор HIPERLAN
HIPERLAN	(High-PERformance LAN) Высокопроизводительная ЛВС
HLR	(Home Location Register) Регистр исходного положения

HM	(HIPERLAN MAC) Уровень MAC сети HIPERLAN
HMPDU	(HIPERLAN MAC PDU) PDU уровня MAC сети HIPERLAN
HMQoS	(HIPERLAN MAC QoS) Качество обслуживания HIPERLAN MAC
HO	(HandOver) Переключение
HO-HMPDU	(HellO-HMPDU) HMPDU приветствия
HP	(HIPERLAN PHY) Физический уровень HIPERLAN
HSCSD	(High Speed Circuit Switched Data) Высокоскоростная передача данных с коммутацией каналов
HTML	(HyperText Markup Language) Язык гипертекстовой разметки
HTTP	(HyperText Transfer Protocol) Протокол передачи гипертекстовых файлов
ICMP	(Internet Control Message Protocol) Протокол управляющих сообщений Internet
ICO	(Intermediate Circular Orbit) Промежуточная круговая орбита
ID	(Identifier) Идентификатор
IEEE	(Institute of Electrical and Electronics Engineers) Институт инженеров по электротехнике и электронике
IETF	(Internet Engineering Task Force) Проблемная группа проектирования Internet
IFS	(Inter Frame Spacing) Межкадровый интервал
ILR	(Interworking Location Register) Межсетевой регистр местонахождения
IMEI	(International Mobile Equipment Identity) Международная идентификация мобильного оборудования
IMF	(Identity Management Function) Функция управления тождественностью
IMSI	(International Mobile Subscriber Identity) Международная идентификация мобильных абонентов
IMT	(International Mobile Telecommunications) Международные мобильные телекоммуникации
IN	(Intelligent Network) Интеллектуальная сеть
IOT	(Isolation Only Transactions) Только разделяющие транзакции
IP	(Internet Protocol) Протокол Internet
IR	(InfraRed) Инфракрасный
IrDA	(Infrared Data Association) Ассоциация передачи данных в инфракрасном диапазоне
IS	(Interim Standard) Временный стандарт
ISDN	(Integrated Services Digital Network) Цифровая сеть связи с интеграцией услуг
ISI	(InterSymbol Interference) Межсимвольная интерференция
ISL	(Inter Satellite Link) Межспутниковый канал
ISM	(Industrial, Scientific, Medical) Полоса частот для промышленного, научного и медицинского использования
ISMA	(Inhibit Sense Multiple Access) Множественный доступ с детектированием подавления
ISO	(International Organization for Standardization[1]) Международная организация по стандартизации
I-TCP	(Indirect TCP) Косвенный TCP
ITU	(International Telecommunication Union) Международный телекоммуникационный союз
ITU-R	(ITU Radiocommunication sector) Сектор радиосвязи ITU
ITU-T	(ITU Telecommunication sector) Сектор телекоммуникаций ITU

[1] Это *не* "Международная организация стандартизации" ("International Standardization Organization") или "Международная организация по стандартам" ("International Standards Organization"), как любят писать некоторые авторы. ISO — это не акроним, это название происходит от греческого слова *isos*, которое означает "равно-" в том смысле, в котором оно употребляется в словах изометрический, изоморфный и т.п.

IV	(Initialization Vector) Вектор инициализации
IWF	(InterWorking Function) Межсетевая функция обмена
JCT	(Japanese Cordless Telephone) Японская система беспроводных телефонов
JDC	(Japanese Digital Cellular) Японская цифровая сотовая система
JPEG	(Joint Photographic Experts Group) Объединенная группа экспертов в области фотографии
KID	(Key IDentifier) Идентификатор ключа
LA	(Location Area) Зона местонахождения
LAI	(Location Area Identification) Идентификация положения
LAN	(Local Area Network) Локальная сеть, ЛВС
LAPD	(Link Access Procedure for the D-channel) Процедура доступа для D-канала
$LAPD_m$	(LAPD for mobile) LAPD для мобильной связи
LBR	(Low Bit-Rate) Низкая скорость передачи данных
LC	(Link Controller) Контроллер канала связи
LED	(Light Emitting Diode) Светодиод
LEO	(Low Earth Orbit) Низкая околоземная орбита
LF	(Low Frequency) Низкая частота
LI	(Length Indicator) Индикатор длины
LIR	(Least Interference Routing) Маршрутизация с минимальной интерференцией
LLC	(Logical Link Control) Управление логическим каналом
LM	(Link Manager) Администратор связей
LMP	(Link Manager Protocol) Протокол управления каналом связи
Loc	(Location) Расположение
LOS	(Line-of-Sight) Линия визирования
LRU	(Last Recently Used) Наиболее позднее использование
LS	(Location Server) Сервер местоположения
MAC	(Medium Access Control) Управление доступом к среде
MACA	(Multiple Access with Collision Avoidance) Множественный доступ с избежанием конфликтов
MATM	(Mobile ATM) Мобильная ATM
MBS	(Mobile Broadband System) Мобильная широкополосная система
MCC	(Mobile County Code) Мобильный код страны
MCI	(Multiplex Configuration Information) Информация о конфигурации мультиплексора
MCM	(MultiCarrier Modulation) Модуляция с множественными несущими
MEO	(Medium Earth Orbit) Средняя околоземная орбита
MF	(Medium Frequency) Средняя частота
MH	(Mobile Host) Мобильный хост
MHEG	(Multimedia and Hypermedia information coding Experts Group) Группа экспертов по кодированию мультимедийной и гипермедийной информации
MIB	(Management Information Base) База управляющей информации
ML	(MSDU Lifetime) Время жизни MSDU
MM	(Mobility Management) Управление мобильностью
MMF	(Mobility Management Function) Функция управления мобильностью
MN	(Mobile Node) Мобильный узел
MNC	(Mobile Network Code) Мобильный код сети
MOC	(Mobile Originated Call) Звонок с мобильного устройства
MOT	(Multimedia Object Transfer) Протокол передачи мультимедийных объектов
MPEG	(Moving Pictures Expert Group) Экспертная группа по вопросам движущегося изображения

M-QoS	(Mobile QoS) Мобильное качество обслуживания
MS	(Mobile Station) Мобильная станция
MS	(Mobile Switch) Мобильный коммутатор
MSAP	(MAC SAP) Точка доступа к службе MAC
MSC	(Mobile (Services) Switching Center) Центр коммутации мобильных служб
MSC	(Main Service Channel) Канал главной службы
MSDU	(MAC SDU) Модуль данных службы MAC
MSIN	(Mobile Subscriber Identification Number) Идентификационный номер мобильного абонента
MSISDN	(Mobile Subscriber ISDN Number; Mobile Station International ISDN Number) Номер ISDN мобильного абонента; международный номер ISDN мобильной станции (см сноску на стр. 112)
MSK	(Minimum Shift Keying) Манипуляция с минимальным сдвигом
MSRN	(Mobile [Station|Subscriber] Roaming Number) Номер роуминга мобильной [станции|абонента]
MT	(Mobile Termination) Мобильное оконечное устройство
MTC	(Mobile Terminated Call) Вызов мобильного оконечного устройства
M-TCP	(Mobile TCP) Мобильный TCP
MTSA	(Mobile Terminal Security Agent) Агент безопасности мобильного терминала
MUL	(Mobile User Link) Канал мобильных пользователей
NA-TDMA	(North American-TDMA) Североамериканская система TDMA
NAV	(Net Allocation Vector) Вектор распределения сети
NDC	(National Destination Code) Национальный код адресата
NFS	(Network File System) Сетевая файловая система
NIB	(Neighbour Information Base) База данных соседей
NIT	(Network Information Table) Таблица сетевой информации
NMAS	(Network Mobility-supporting ATM Switch) Коммутатор ATM с поддержкой мобильности сети
NMT	(Nordic Mobile Telephone System) Скандинавская мобильная телефонная система
NNI	(Network-to-Network Interface) Межсетевой интерфейс
NNI+M	(NNI+Mobility) NNI с поддержкой мобильности
NRL	(Normalized Residual HMPDU Lifetime) Нормированное остаточное время жизни HMPDU
NSA	(Network Security Agent) Агент безопасности сети
NSS	(Network and Switching Subsystem) Подсистема сетей и коммутации
NTSC	(National Television Standards Committee) Национальный комитет по телевизионным стандартам США
OFDM	(Orthogonal FDM) Ортогональная FDM
OMC	(Operation and Maintenance Centre) Центр эксплуатации и обслуживания
OSI	(Open Systems Interconnection) Взаимодействие открытых систем
OSS	(Operation Subsystem) Операционная подсистема
PACS	(Personal Access Communications System) Система связи персонального доступа
PACS-UB	(PACS-Unlicensed Band) Система связи персонального доступа в нелицензируемой полосе частот
PAD	(PADding) Заполнение
PAD	(Program Associated Data) Ассоциированные с программой данные
PC	(Personal Computer) Персональный компьютер, ПК
PCF	(Point Coordination Function) Точечная координационная функция
PCH	(Paging Channel) Поисковый канал

PCM	(Pulse Code Modulation) Импульсно-кодовая модуляция
PCS	(Personal Cellular System) Персональная сотовая система
PCS	(Personal Communication Service) Персональная служба связи
PDA	(Personal Digital Assistant) Персональный цифровой помощник
PDC	(Personal Digital Cellular) Персональная цифровая система сотовой связи
PDF	(Portable Document Format) Переносимый формат документов
PDN	(Public Data Network) Сеть передачи данных общего пользования
PDO	(Packet Data Optimized) Служба с оптимизацией пакетов
PDTCH	(Packet Data TCH) Канал обмена пакетными данными
PDU	(Protocol Data Unit) Модуль данных протокола
PHS	(Personal Handyphone System) Персональная система переносных телефонов
PHY	(PHYsical layer) Физический уровень
PIFS	(PCF IFS) Межкадровый промежуток PCF
PIN	(Personal Identity Number) Персональный идентификационный номер
PLCP	(Physical Layer Convergence Protocol) Протокол сходимости физического уровня
PLI	(Padding Length Indicator) Индикатор длины заполнения
PLL	(Phase Lock Loop) Контур фазовой автоподстройки частоты, ФАПЧ
PLMN	(Public Land Mobile Network) Мобильная наземная сеть общего пользования
PLW	(PLCP-PDU Length Word) Длина слова PDU протокола PLCP
PM	(Phase Modulation) Фазовая модуляция
PMD	(Physical Medium Dependent) Зависимый от физической среды
POTS	(Plain Old Telephone Service) Обычная телефонная сеть
PRACH	(Physical Random Access Channel) Физический канал случайного доступа
PRMA	(Packet Reservation Multiple Access) Множественный доступ с резервированием пакетов
PS	(Power Saving) Энергосбережение
PSF	(PLCP Signalling Field) Сигнальное поле PLCP
PSK	(Phase Shift Keying) Фазовая манипуляция
PSN	(PDU Sequence Number) Порядковый номер PDU
PSPDN	(Public Switched Packet Data Network) Коммутируемая телефонная сеть общего пользования
PSTN	(Public Switched Telephone Network) Общественная коммутируемая телефонная сеть
PT	(Portable radio Termination) Портативное оконечное радиоустройство
PTP	(Point-to-Point) Двухточечный
PTP-CLNS	(PTP-ConnectionLess Network Service) Двухточечная сетевая служба без установления соединения
PTP-CONS	(PTP-Connection Oriented Network Service) Двухточечная сетевая служба, ориентированная на установление соединения
PUK	(PIN Unblocking Key) Ключ снятия блокировки PIN
QAM	(Quadrature AM) Квадратурная амплитудная модуляция
QoS	(Quality of Service) Качество обслуживания
QPSK	(Quadrature PSK) Квадратурная фазовая манипуляция
RA	(Receiver Address) Адрес получателя
RACH	(Random Access Channel) Канал произвольного доступа
RAL	(Radio Access Layer) Уровень радиодоступа
RAND	(Random number) Случайное число
Req	(Request) Запрос
Res	(Response) Отклик
RFC	(Request For Comments) Серия документов IETF (букв. "Запрос на комментарии")

RIB	(Route Information Base) База данных маршрутизации
RIP	(Routing Information Protocol) Протокол маршрутной информации
RL	(Residual Lifetime) Остаточное время жизни
RLC	(Radio Link Control) Управление радиоканалом
RLP	(Radio Link Protocol) Протокол радиоканала
RM	(Resource Management) Управление ресурсами
RNS	(Radio Network Subsystem) Сетевая радиоподсистема
RR	(Radio Resource) Радиоресурсы
RRC	(Radio Resource Control) Контроль радиоресурсов
RRM	(Radio Resource Management) Управление радиоресурсами
RSA	(Rivest, Shamir, Adleman) Алгоритм Ривеста-Шамира-Адельмана
RSS	(Radio SubSystem) Радиоподсистема
RT	(Radio Transceiver) Радиотрансивер
RTR	(Radio Transmission and Reception) Радиопередача и прием
RTS	(Request To Send) Запрос на передачу
RTT	(Radio Transmission Technologies) Технологии радиопередачи
RTT	(Round Trip Time) Период кругового обращения (время на передачу и подтверждение приема)
SA	(Source Address) Адрес источника
SAAL	(Signalling AAL) AAL передачи сигналов
SACCH	(Slow Associated Dedicated CCH) Медленный ассоциированный выделенный канал управления
SAMA	(Spread Aloha Multiple Access) Множественный доступ с расширением по схеме Aloha
SAP	(Service Access Point) Точка доступа к службе
SATM	(Satellite ATM Services) Спутниковые службы ATM
SC	(Sanity Check) Санитарная проверка
SC	(Synchronization Channel) Канал синхронизации
SCF	(Service Control Function) Функция контроля служб
SCH	(Synchronous Channel) Синхронный канал
SCO	(Synchronous Connection-Oriented link) Ориентированный на установление соединения синхронный канал
SCPS	(Space Communication Protocol Standards) Стандарты протоколов космической связи
SCPAS-TP	(SCPS-Transport Protocol) Транспортный протокол SCPS
SDCCH	(Stand-alone Dedicated CCH) Автономный выделенный канал управления
SDM	(Space Division Multiplexing) Пространственное уплотнение сигналов
SDMA	(Space Division Multiple Access) Множественный доступ с пространственным разделением
SDT	(Service Description Table) Таблица описания служб
SDTV	(Standard Definition TV) Телевидение стандартной четкости
SDU	(Service Data Unit) Модуль данных службы
SEC-SAP	(Security SAP) Точка доступа к службе безопасности
SEQN	(SEQuence Number) Порядковый номер
SFD	(Start Frame Delimiter) Стартовый разделитель кадра
SFN	(Single Frequency Network) Сеть единой частоты
SGSN	(Serving GSN) Обслуживающий GSN
SH	(Supervisory Host) Хост супервизора
SHF	(Super High Frequency) Сверхвысокая частота, СВЧ
SIFS	(Short IFS) Короткий межкадровый интервал
SIG	(Signalling) Передача сигналов

SIM	(Subscriber Identity Module) Модуль идентификации абонента
SMRIB	(Source Multipoint Relay Information Base) Информационная база исходных многоточечных ретрансляторов
SMS	(Short Message Service) Служба коротких сообщений
SN	(Subscriber Number) Номер абонента
SNACK	(Selective Negative ACKnowledgement) Выборочное отрицательное подтверждение
SNAP	(Sub-Network Access Protocol) Протокол доступа к подсети
SNDCP	(Subnetwork Dependent Convergence protocol) Зависящий от подсети протокол сближения
SRES	(Signed Response) Отмеченный отклик
SS	(Supplementary Service) Дополнительная служба
SS7 (SSN7)	(Signalling System No. 7) Система передачи сигналов 7
S-SAP	(Session-SAP) Точка доступа к службе сеансового уровня
SSL	(Secure Sockets layer) Протокол защищенных сокетов
STA	(STAtion) Станция
SUMR	(Satellite User Mapping Register) Регистр соответствия пользователей и спутников
SW	(Short Wave) Коротковолновый
SwMI	(Switching and Management Infrastructure) Инфраструктура коммутации и управления
T	(Terminal) Терминал
TA	(Transmitter Address) Адрес передатчика
TCH	(Traffic CHannel) Канал информационного обмена
TCH/F	(TCH Full rate) Полноскоростной TCH
TCH/FS	(TCH/F Speech) TCH/F передачи речи
TCH/H	(TCH Half rate) Полускоростной TCH
TCH/HS	(TCH/H Speech) TCH/H передачи речи
TC-HMPDU	(Topology Control-HMPDU) HMPDU управления топологией
TCP	(Transmission Control Protocol) Протокол управления передачей
TD-CDMA	(Time Division-CDMA) CDMA с временным разделением
TDD	(Time Division Duplex) Дуплекс с временным разделением
TDM	(Time Division Multiple Access) Множественный доступ с временным разделением
TDT	(Time and Data Table) Таблица времени и дат
TE	(Terminal Equipment) Терминальное оборудование
TETRA	(Terrestrial Trunked Radio) Наземная радиомагистраль
TFI	(Transport Format Identifier) Идентификатор транспортного формата
TFTS	(Terrestrial Flight Telephone System) Полетная телефонная система
TI	(Type Indicator) Индикатор типа
TIB	(Topology Information Base) Информационная база топологии
TIM	(Traffic Indication Map) Карта индикации транспортного потока
TLLI	(Temporary Logical Link Identity) Временная идентификация логического канала
TLS	(Transport Layer Security) Протокол безопасности транспортного уровня
TM	(Traffic Management) Управление информационным обменом
TMN	(Telecommunication Management Network) Сеть управления телекоммуникациями
TMSI	(Temporary Mobile Subscriber Identity) Временная идентификация мобильного абонента
TOS	(Type Of Service) Вид службы
TPC	(Transmit Power Control) Контроль мощности передачи
TR-SAP	(Transaction SAP) Точка доступа к службе транзакций
T-SAP	(Transport SAP) Точка доступа к транспортной службе
TSF	(Timing Synchronization Function) Функция временной синхронизации

T-TCP	(Transaction-oriented TCP) TCP, ориентированный на транзакции
TTC	(Telecommunications Technology Council) Совет телекоммуникационных технологий
TTL	(Time To Live) Время жизни
TV	(TeleVision) Телевидение
UBR	(Unspecified BitRate) Не уточненная скорость передачи
UD	(User Data) Пользовательские данные
UDP	(User Datagram Protocol) Протокол пользовательских дейтаграмм
UE	(User Equipment) Пользовательское оборудование
UHF	(Ultra High Frequency) Ультравысокая частота, УВЧ
UIM	(User Identification Module) Модуль идентификации пользователя
UMTS	(Universal Mobile Telecommunication system) Универсальная мобильная телекоммуникационная система
UN	(United Nations) Организация Объединенных Наций, ООН
UNI	(User-to-Network Interface) Интерфейс "пользователь-сеть"
UNI+M	(UNI+Mobility) UNI с поддержкой мобильности
UP	(User Priority) Приоритет пользователя
UPT	(Universal Personal Telecommunications) Универсальные личные соединения
URI	(Uniform Resource Identifier) Универсальный идентификатор ресурса
UTRA	(UMTS Terrestrial Radio Access) Наземный радиодоступ UMTS
UTRAN	(UTRA Network) Сеть UTRA
UWC	(Universal Wireless Communications) Универсальная беспроводная связь
V+D	(Voice and Data) Речь и данные
VAD	(Voice Activity Detection) Детектирование активности речи
VBR	(Variable BitRate) Переменная скорость передачи
VBR-nrt	(VBR non real-time) VBR нереального времени
VBR-rt	(VBR real-time) VBR реального времени
VC	(Virtual Circuit) Виртуальный канал
VCC	(Visitor Country Code) Гостевой код страны
VDB	(Visitor Data Base) База данных посетителей
VHE	(Virtual Home Environment) Виртуальная исходная среда
VHF	(Very High Frequency) Очень высокая частота, ОВЧ
VLF	(Very Low Frequency) Очень низкая частота, ОНЧ
VLR	(Visitor Location Register) Регистр местонахождения посетителей
VNDC	(Visitor National Destination Code) Гостевой государственный код адресата
W3C	(World Wide Web Consortium) Консорциум World Wide Web
WAE	(Wireless Application Environment) Среда беспроводных приложений
WAN	(Wide Area Network) Глобальная сеть
WAP	(Wireless Application Protocol) Протокол беспроводных приложений
WATM	(Wireless ATM) Беспроводная ATM
WCAC	(Wireless Connection Admission Control) Контроль доступа беспроводного соединения
W-CDMA	(Wideband-CDMA) Широкополосная схема CDMA
WCMP	(Wireless Control Massage Protocol) Беспроводный протокол управляющих сообщений
W-CTRL	(Wireless ConTRoL) Беспроводное управление
WDP	(Wireless Datagram Protocol) Беспроводный протокол дейтаграмм
WLAN	(Wireless LAN) Беспроводная ЛВС
WLL	(Wireless Local Loop) Беспроводная местная линия связи
WML	(Wireless Markup Language) Язык разметки для беспроводных приложений
WMLScript	(Wireless Markup Language Script) Сценарий WML
WMT	(Wireless Mobile Terminal) Беспроводный мобильный терминал

WPAN	(Wireless Personal Area Network) Беспроводная персональная сеть
WP-CDMA	(Wideband Packet-CDMA) Широкополосная сеть CDMA с коммутацией пакетов
WRC	(World Radio Conference) Всемирная конференция по радиосвязи
WSP	(Wireless Session Protocol) Беспроводный сеансовый протокол
WSP/B	(Wireless Session Protocol/Browsing) Беспроводный сеансовый протокол/просмотр
WTA	(Wireless Telephony Application) Приложение беспроводной телефонии
WTAI	(Wireless Telephony Application Interface) Интерфейс приложения беспроводной телефонной связи
WTLS	(Wireless Transport Layer Security) Протокол безопасности транспортного уровня беспроводных приложений
WTP	(Wireless Transaction Protocol) Беспроводный протокол транзакций
WWW	(World Wide Web) Глобальная система гипертекстовой связи, досл. "Всемирная паутина"
XML	(eXtensible Markup Language) Расширяемый язык разметки
XOR	(eXclusive OR) Исключающее ИЛИ

A

ABR, *228*
ACF, *227*; *228*
AESA, *244*
AFS, *305*
AGCH, *108*
AIUR, *121*
Aloha, *81*; *92*; *131*
  с разделением, *81*
  с резервированием, *83*
AMES, *219*; *222*
AMPS, *25*; *26*; *40*; *63*; *87*; *97*
A-Netz, *25*
ANSI, *133*
APCF, *228*
APCM, *249*
APCP, *223*; *248*
ARIB, *133*
ARQ, *230*
ASK, *57*
ATIM, *194*
ATM, *33*; *79*
ATM Forum, *220*
ATMC, *227*; *228*
AuC, *105*
AUS, *244*

B

BCA, *70*
BCCH, *107*
BER, *186*; *229*
Bluetooth, *19*; *33*; *67*; *87*; *177*; *209*
BMP, *163*
B-Netz, *25*
BPSK, *58*
BRAN, *229*; *230*
BSC, *103*
BSS, *101*; *177*
BSSAP, *112*
BSSGP, *125*
BTS, *101*; *102*
BTSM, *111*

C

CAC, *200*; *249*
CAMEL, *133*
CBR, *227*
CC, *111*; *113*; *130*; *249*
CCA, *178*; *179*
CCCH, *107*
CCF, *226*; *228*
CCH, *107*
CDM, *54*; *63*; *71*
CDMA, 26; 27; 40; 87 – 94; 133 – 138; 154; 234
CDV, *230*
CEPT, *39*
CIF, *160*
CKSN, *125*
CLMS, *130*
CM, *111*
C-Netz, *25*
Coda, *305*
COFDM, *60*
COMS, *130*
COS, *236*
CPM, *57*
CSMA, *82*
  CSMA/CA, *82*
CSMA/CA, *183*
CSMA/CD, *76*
CU, *160*

D

DAB, *18*; *19*; *33*; *38*; *160*
DAMA, *83*
DCCH, *108*
DCF, *182*
DCS 1800, *26*
DECT, *25*; *26*; *32*; *38*; *41*; *70*; *80*; *93*; *126*
DFWMAC, *182*
DHCP, *33*; *255*; *271*
DIFS, *182*
DLC, *75*
  в сетях IEEE 802.11, *178*

в системе DECT, *130*
на уровне RAL, *230*
DQPSK, *60*
DSDV, *277*
DSMA, *87*
DSSS, *64*; *66*; *89*; *180*
DTIM, *193*
DTMF, *112*
DVB, *33*; *164*
DVD, *165*
DVTR, *165*

E

EDGE, *133*
EDTV, *165*
EIR, *105*
EIRP, *179*
EIT, *165*
EMAS, *244*
EMAS-E, *223*
EMAS-N, *223*
ESS, *178*
Ethernet, *76 – 82*; *174 – 178*
ETSI, *39*; *98*; *133*
EY-NMPA, *82*

F

FACCH, *108*
FCCH, *107*
FDD, *78*
FDM, *52*; *63*; *66*; *70*; *71*; *75*; *78*; *79*
FDMA, *78*; *93*; *94*
FEC, *100*; *110*
FHSS, *66*; *179*
FIB, *160*
FIC, *160*; *161*
Ficus, *308*
FPLMTS, *132*
FSK, *56*; *57*

G

GGSN, *124*
GIF, *163*
Globalstar, *153*
GMM, *133*
GMSC, *104*
GPRS, *122*
GPS, *18*; *20*; *144*
GR, *124*
GRE, *266*
GSM, *18*; *25 – 27*; *38*; *40*; *97 –126*
GSN, *124*
GTP, *125*
GWL, *145*

H

HBR, *199*
HDLC, *100*
HDTV, *165*
HIPERACCESS, *231*
HIPERLAN, *26*; *33*; *41*; *177*; *196*; *231*
HIPERLINK, *231*
HLR, *104*; *152*
HSCSD, *121*
HTML, *311*
HTTP, *308 – 310*

I

ICO, *93*; *153*
IEEE 802.11, *26*; *33*; *41*; *177*; *178*
IMEI, *103*
IMF, *227*
IMSI, *103*; *111*; *113*
IMT-2000, *27*; *132*
IrDA, *39*
Iridium, *26*; *93*; *153*
IS-54, *63*, *93*
IS-95, *63*
ISDN, *99*
ISL, *145*
ISMA, *87*
I-TCP, *290*; *291*
ITU, *27*; *39*
IWF, *104*

J

JPEG, *163*

L

LA, *104*
LAI, *104*

LAPD, *111*
LBR, *199*
LIR, *281*
Little Work, *307*
LLC, *75*; *178*; *197*; *198*; *227*; *230*
LRU, *306*
LS, *244*

M

MAC, *75*; *76*; *178*; *182*; *189*; *190*; *198*; *204*; *212*; *227*
в системе DECT, *130*
на уровне RAL, *229*
MACA, *85*
MANET, *283*
MATM, *224*
MBS, *27*
MCC, *113*
MCI, *162*
MCM, *60*
MHEG, *163*
MIB, *179*; *191*
MIo-NFS, *308*
MM, *111*; *130*; *249*
MMF, *226*; *228*
MNC, *113*
Mobile IP, *21*; *33*; *255 – 285*
MOC, *114*
MOT, *162*
MPEG, *163*
M-QoS, *247*
MSC, *104*; *160*; *161*
MSIN, *113*
MSISDN, *104*; *112*
MSK, *58*
MSRN, *113*
MTC, *113*
M-TCP, *294*
MTSA, *227*
MUL, *145*

N

NAV, *186*
NDC, *113*
NFS, *304*
NIT, *165*
NMT, *25*; *39*
NNI+M, *223*; *226*
Nrt-VBR, *228*
NSA, *228*
NSS, *101*; *104*

O

OFDM, *60*
OMC, *105*
OSS, *101*; *104*

P

PACS, *41*
PAD, *161*
PCF, *182*; *187*
PCH, *108*
PCM, *112*
PCS, *222*
PCS 1900, *26*
PDA, *22*; *30*
PDC, *40*; *97*
PDN, *104*; *124*
PDTCH, *126*
PHS, *80*
PIFS, *183*
PIN, *103*
PLCP, *178*
PLMN, *99*
PMD, *178*
PRMA, *83*
PSK, *58*
PSTN, *99*
PUK, *103*

Q

QAM, *60*
QoS, *14*; *33*; *171*; *197*; *220*; *221*; *230*; *237*; *336*; *344*
QoS, *247*
QPSK, *59*

R

RACH, *108*
RAL, *221*; *225*; *229*
RLC, *126*
RLP, *100*

RM, *249*
RNS, *134*
RR, *111*
RRC, *227*; *228*
RRM, *249*
RSS, *101*
RTR, *227*; *228*
RTT, *110*; *288*
Rt-VBR, *227*

S

SACCH, *108*
SAMA, *92*
SAP, *178*
SATM, *223*
SC, *161*
SCF, *228*
SCH, *107*
SCPS, *300*
SCPS-TP, *300*
SDM, *51*; *68*; *75*; *78*
SDMA, *78*; *93*; *94*
  в GSM, *105*
SDT, *165*
SDTV, *165*
SDU, *123*
SFN, *160*
SGSN, *124*
SIFS, *183*
SIM, *103*
SMS, *101*; *111*
SN, *113*
SNACK, *300*
SNDCP, *125*
SS, *111*
SS7, *101*; *104*
SUMR, *152*

T

TCH, *107*
TCP, *34*
TCP, *287*; *290*; *293*
TDD, *80*
TDM, *52*; *63*; *66*; *75*; *78*; *79*; *80*; *83*
TDMA, *26*; *40*; *79*; *81*; *93*; *94*
  с резервированием, *84*
TDT, *165*
Teledesic, *153*
TETRA, *19*; *130*
TLLI, *125*
TMN, *105*
TMSI, *104*; *111*; *113*
Token Ring, *79*; *177*
TTC, *133*
T-TCP, *298*

U

UBR, *228*
UDP, *287*
UIM, *226*
UMTS, *27*; *132*; *133*
UNI, *225*
UNI+M, *223*; *225*
UPT, *132*
UTMS, *33*
UTRA, *133*
UTRAN, *134*

V

VAD, *111*
VCC, *113*
VHE, *133*
VLR, *104*; *152*
VNDC, *113*

W

WAP, *34*
WARC, *24*
WATM, *33*; *219*; *221 – 223*
WCAC, *249*
W-CDMA, *134*
WLAN, *20*; *41*; *171*
WLL, *127*
WMT, *223*
WRC, *39*
WTM, *226*
WWW, *34*; *308*

А

Адресация, *31*; *190*; *247*
Амплитуда сигнала, *41*
Антенны, *43*

диаграмма направленности, *43*
изотропный излучатель, *43*
интеллектуальные, *45*
многоэлементные решетки, *45*
направленные, *44*
разнесенные, *45*
секторные, *44*; *70*
эффект направленности, *43*
Асимметричные системы связи, *158*
Аутентификация, *119*
в Bluetooth, *215*
в GSM, *119*

Б

Базовая сеть, *245*
Базовая станция, *116*; *127*; *175*
Безопасность
в Bluetooth, *215*
в GSM, *118*
в Mobile IP, *258*
в WAP, *319*
в беспроводных сетях, *322*
в БЛВС, *172*; *173*
в системах WATM, *243*
Беспроводная передача, *37*
Беспроводная связь
история, *23*
БЛВС, *171*

Г

Гармоника, *42*

Д

Декапсуляция, *264*
Демодуляция, *57*
Детектирование, *57*
Диаграмма направленности, *43*
Дифракция, *48*
Дуплекс, *78*; *80*

З

Затухание, *49*
Защитный интервал, *105*
Зона обслуживания, *147*

И

Изотропный излучатель, *43*
Инкапсуляция, *264*; *265*; *266*; *269*
Инфракрасная передача, *39*
Инфракрасный диапазон, *173*

К

Канальный уровень, *31*
Код Баркера, *64*; *88*; *180*
Код состояния, *309*
Конфликт, *76*; *77*; *82*; *84*; *184*; *305*; *306*
Коррелятор, *65*
Кэш, *310*
Кэширование, *310*

М

Манипуляция
амплитудная, *56*; *57*
с минимальным сдвигом, *58*
фазовая, *56*; *58*
частотная, *56*; *58*
Маршрутизация, *221*; *278*
DSDV, *277*
в спутниковых системах, *151*
в эпизодических сетях, *274*
с минимальной интерференцией, *281*
треугольная, *268*
Межсимвольная интерференция, *49*
Многолучевое распространение, *48*
Множественный доступ
CDMA, *88*
DSMA, *87*
ISMA, *87*
MACA, *85*
SAMA, *92*
Мобильный телефон, *22*
Модуляция, *55*
ASK, *56*; *57*
BPSK, *58*
CPM, *57*
DQPSK, *60*
FSK, *56*; *57*
PSK, *56*; *58*
QAM, *60*
QPSK, *59*
амплитудная, *56*

аналоговая, *56*
с несколькими несущими, *60*
спектральная эффективность, *57*
устойчивость к помехам, *57*
фазовая, *56*
цифровая, *55*; *64*
частотная, *56*
энергетическая эффективность, *57*

Н

Накопление, *305*
Настроечная последовательность, *49*
Ноутбук, *23*

О

Окно состязания, *184*
Опрос, *87*
Оптическая передача, *38*
Основная эталонная модель, *30*
Откат, *183*; *185*

П

Пейджер, *22*
Перегрузка, *287*
Переключение, *115*; *221*; *269*
в WATM, *232*
в спутниковых системах, *152*
обратное, *236*
прямое, *241*
Переносной компьютер, *22*; *23*
Перескок частоты, *53*
Пикосеть, *211*
Плавный старт, *288*
Подтверждение, *185*
Портал, *178*
Посещаемая сеть, *245*
Привязка, *268*

Р

Радиомагистраль, *41*; *131*
Радиопередача, *38*
уплотнение, *50*
Разложение Фурье, *42*
Рассеивание сигнала, *47*
Расширение спектра, *61 – 66*
чип, *64*
Роуминг, *17*; *112*; *179*; *195*

С

Сетевой уровень, *31*
в мобильной связи, *255*
в системе DECT, *130*
Сети с инфраструктурой, *174*
Сеть
эпизодическая, *273*
Сигналы, *41*
временная диаграмма, *42*
гармоника, *42*
дифракция, *48*
затухание, *47*; *49*
квадратурная составляющая, *43*
основная частота, *42*
отражение, *47*
параметры, *41*
постоянная составляющая, *42*
потери, *46*
распространение, *42*
рассеивание, *47*
расширение, *48*
синфазная составляющая, *43*
фазовая диаграмма, *42*
частотная диаграмма, *42*
частотный спектр, *42*
экранирование, *47*
Симметричные системы связи, *157*
Симметричный вибратор, *43*
Синхронизация, *79*; *191*
Согласованность, *304*
Сотовые сети, *97*
Сотовые системы, *68*
базовая станция, *68*; *69*
кластер, *70*
распределение каналов, *70*
ячейка, *68*
Сотовый телефон, *19*
Спектр
расширенный, *61*
Спутники, *146*
GEO, *150*
LEO, *150*
MEO, *151*
геостационарные, *146*
типы орбит, *148*

Спутниковые системы, *93*
Суперкадр, *135*

Т

Телеграф, *24*
Телефоны
  беспроводные, *40*
  мобильные, *40*
Терминал, *76*; *77*
Топология
  динамическая, *275*
Точка доступа, *174*; *177*
Транзакция, *308*
Транспортный уровень, *31*
  в мобильной связи, *287*
Туннелирование, *264 – 270*

У

Угол возвышения, *147*
Угол склонения, *147*
Уплотнение, *50*; *87*
  CDM, *54*
  COFDM, *60*
  FDM, *52*
  OFDM, *60*
  SDM, *51*
  TDM, *52*
  временное, *52*
  пространственное, *50*
  частотное, *52*
Управление местоположением, *242*
Уровень приложений, *31*

Ф

Фазовая диаграмма, *42*
Фазовый сдвиг, *41*
Файловые системы, *304*
Физический уровень, *30*; *109*; *128*
  в Bluetooth, *211*
  в сетях HIPERLAN, *198*
  в сетях IEEE 802.11, *179*
  в системе DECT, *128*
  на уровне RAL, *229*
Фрагментация, *187*

Ц

Цифровая модуляция, *55*; *64*
Цифровые сотовые сети, *97*

Ч

Частота, *41*
  перескок, *78*
  распределение спектра, *37*
  регулирование, *39*
Частотная диаграмма, *42*
Чиповая последовательность, *64*
Чужая сеть, *245*

Ш

Широковещательные системы, *157*
Шифрование
  в Bluetooth, *215*
  в GSM, *120*
  в HIPERLAN, *205*

Э

Эквалайзер, *49*
Экранирование, *47*
Электронная записная книжка, *22*; *30*
Эмуляция, *306*
Энергосбережение, *172*; *192*; *197*
Эпизодическая сеть, *174*;*190*; *273*
Эффект Доплера, *50*

Я

Ячейка, *46*; *102*; *115*; *127*; *145*

Научно-популярное издание

Йоган Шиллер

Мобильные коммуникации

Литературный редактор	*Т.Т. Шматко*
Верстка	*О.В. Линник*
Художественный редактор	*Г.В. Базылев*
Обложка	*Е.П. Дынник*
Корректоры	*З.В. Александрова, Л.А. Гордиенко, Л.В. Коровкина, О.В. Мишутина*

Издательский дом "Вильямс".
101509, Москва, ул. Лесная, д. 43, стр. 1.
Изд. лиц. ЛР № 090230 от 23.06.99
Госкомитета РФ по печати.

Подписано в печать 17.09.2002. Формат 70×100/16.
Гарнитура Times. Печать офсетная.
Усл. печ. л. 30,9. Уч.-изд. л. 28,6.
Тираж 3500 экз. Заказ № 1276.

Отпечатано с диапозитивов в ФГУП "Печатный двор"
Министерства РФ по делам печати,
телерадиовещания и средств массовых коммуникаций.
197110, Санкт-Петербург, Чкаловский пр., 15.